

Aus der Klinik für Kardiologie und Pneumologie
(Prof. Dr. med. G. Hasenfuß)
der Medizinischen Fakultät der Universität Göttingen

**Lernverhalten und Lernerfolg von
Studierenden der Humanmedizin an der
Universitätsmedizin Göttingen während der
Corona-Pandemie**

INAUGURAL-DISSERTATION

zur Erlangung des Doktorgrades
der Medizinischen Fakultät der
Georg-August-Universität zu Göttingen

vorgelegt von

Charlotte Elisabeth Arkenberg [geb. Kampsen]

aus

Vechta

Göttingen 2023

Dekan: Prof. Dr. med. W. Brück

Betreuungsausschuss

Betreuer*in: Prof. Dr. med. T. Raupach

Ko-Betreuer*in: PD Dr. med. dent. S. Sennhenn-Kirchner

Prüfungskommission

Referent*in: Prof. Dr. med. T. Raupach

Ko-Referent*in: PD Dr. med. dent. S. Sennhenn-Kirchner

Drittreferent*in:

Datum der mündlichen Prüfung: 30. Juli 2024

Hiermit erkläre ich, die Dissertation mit dem Titel „Lernverhalten und Lernerfolg von Studierenden der Humanmedizin an der Universitätsmedizin Göttingen während der Corona-Pandemie“ eigenständig angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Göttingen, den
.....
(Unterschrift)

Die Daten, auf denen die vorliegende Arbeit basiert, wurden teilweise publiziert:

Seer M, **Kampsen C**, Becker T, Hobert S, Anders S, Raupach T (2022): Use of digital teaching resources and predictors of medical student performance during the pandemic: A prospective study. PLoS One 17, e0268331

Inhaltsverzeichnis

Abbildungsverzeichnis	III
Tabellenverzeichnis	IV
Abkürzungsverzeichnis.....	V
1 Einleitung	1
1.1 Wandel der universitären Lehre im Zeichen der Digitalisierung.....	1
1.2 Entwicklung der digitalen Lehre in der Humanmedizin.....	2
1.3 Beginn und Folgen der Covid-19 Pandemie auf die universitäre Lehre.....	4
1.4 Neue Lehrformate im Zuge der Covid-19 Pandemie.....	4
1.5 Erste Resultate der Lehre unter Covid-19.....	6
1.5.1 Für und Wider der digitalen Lehre	6
1.5.2 Bedeutung der Erkennung von gefährdeten Studierenden	8
1.6 Wissenschaftliche Fragestellung.....	9
2 Material und Methoden	11
2.1 Bildungseinrichtung und Studienkohorte.....	11
2.2 Änderungen im Ablauf des Moduls durch die Covid-19 Pandemie	12
2.2.1 Zeitliche Änderung.....	12
2.2.2 Inhaltliche Änderungen.....	13
2.2.3 Angewandte Lehrformate	13
2.3 Messung der Lernergebnisse.....	20
2.4 Datensammlung und statistische Analyse	21
3 Ergebnisse.....	25
3.1 Charakteristika der Studienkohorte	25
3.2 Analyse der Nutzungscharakteristika	25
3.2.1 Tutorials und fallbasierte Aufgaben	25
3.2.2 Elektronische Fallseminare (E-Fallseminare)	26
3.2.3 #clue.....	28
3.3 Überprüfung des Wissenserwerbs und der Wissensanwendung	29
3.3.1 EMERGE.....	29
3.3.2 Studentische Leistungen der Abschlussprüfungen	30
3.4 Analyse der Leistungsverläufe	31
3.5 Analyse der Nutzungscharakteristika als Einflussfaktoren auf den Lernerfolg.....	33
3.5.1 Regressionsanalyse mit den Prüfungsleistungen als abhängige Variable.....	33
3.5.2 Regressionsanalyse mit den EMERGE-Leistungen als abhängige Variable.....	35
3.6 Prädiktoren für negative Leistungsveränderungen.....	37
4 Diskussion.....	39
4.1 Darlegung der Ergebnisse und Studienlage.....	39
4.2 Beurteilung der Nutzungscharakteristika der digitalen Ressourcen	40
4.2.1 Hohe Leistungen in den fragenbasierten <i>tools</i>	40

4.2.2	Geschlechterunterschiede	43
4.2.3	Altersunterschiede.....	44
4.2.4	<i>Early-completer</i> und <i>non early-completer</i>	44
4.2.5	Bearbeitungszeit der E-Fallseminare.....	45
4.2.6	Abschließende Beurteilung der Nutzungscharakteristika	46
4.3	Beurteilung der Leistungsverläufe	47
4.4	Einflussfaktoren auf den Lernerfolg	48
4.4.1	Die studentische Prüfungsleistung	48
4.4.2	Studentische Leistung in EMERGE	49
4.5	Wertung der Prädiktoren für negative Leistungsveränderungen	52
4.6	Stärken und Limitationen.....	53
4.7	Weiterführende Forschung.....	54
5	Zusammenfassung	55
6	Literaturverzeichnis	56

Abbildungsverzeichnis

Abbildung 1: Zeitlicher Ablauf des Moduls M3.1 vor Beginn der Pandemie und im ersten Pandemie-Semester.	14
Abbildung 2: Screenshots einer EMERGE-Spielsitzung (Middeke et al. 2018).	19
Abbildung 3: Lehr- und Prüfungsformate nach Bildungszielen kategorisiert.	21
Abbildung 4: Streudiagramm Perzentilränge der Prüfungsergebnisse.	32

Tabellenverzeichnis

Tabelle 1: Einflussfaktoren auf die Prüfungsleistung der Studierenden im Sommersemester 2020 in einer univariaten linearen Regressionsanalyse.	34
Tabelle 2: Multivariate lineare Regressionsanalyse der Einflussfaktoren auf die studentische Prüfungsleistung im Sommersemester 2020.	35
Tabelle 3: Einflussfaktoren auf die Leistung der Studierenden im computergestützten Spiel EMERGE in einer univariaten linearen Regressionsanalyse.	36
Tabelle 4: Prädiktoren für die studentische Prüfungsleistung in der ersten Modulwoche: eine univariate lineare Regressionsanalyse.	38
Tabelle 5: Multivariate lineare Regressionsanalyse der Prädiktoren der studentischen Prüfungsleistung in der ersten Modulwoche.	38

Abkürzungsverzeichnis

Bundesregierung D	Bundesregierung Deutschland
BZ	Bildungsziel
Covid-19	Coronavirus disease 2019
DIPS	Digitales Prüfungs- und Schulungszentrum
E-Fallseminar	Elektronisches Fallseminar
<i>E-Learning</i>	<i>Electronic-Learning</i>
ILIAS	Integriertes Lern-, Informations- und Arbeitskooperations-System
IMPP	Institut für medizinische und pharmazeutische Prüfungsfragen
KI	Konfidenzintervall
KS	Klinisches Semester
MC	Multiple-Choice
OSCE	<i>Objective Structured Clinical Examination</i>
SARS-CoV-2	<i>Severe acute respiratory syndrome coronavirus 2</i>
VR	<i>Virtual Reality</i>

1 Einleitung

1.1 Wandel der universitären Lehre im Zeichen der Digitalisierung

Der Grundstein der medizinischen Lehre wurde schon weit vor unserer Zeit gelegt. Einer der bedeutendsten Ärzte der Antike, Hippokrates von Kos (um 460 v. Chr. geboren), war nicht nur ein außergewöhnlicher Arzt, sondern stellte auch die Lehre der Medizin in seinen Lebensmittelpunkt. So erwähnen Emmanouil und Kollegen sowie Yapijakis, dass Hippokrates als Wegbereiter der medizinischen Literatur und Lehre selbst einige Schriften des *Corpus Hippocraticum* verfasste (Emmanouil et al. 2008; Yapijakis 2009). Hierbei handelt es sich um eine zusammengestellte Schrift aus mehr als 60 einzelnen Texten, die unter anderem der Lehre von Krankheiten und deren Prognose dienten (Emmanouil et al. 2008). Gleichzeitig hielt er vor seinen Studierenden Vorträge, in denen er nicht nur Erkrankungen lehrte, sondern auch ethische Grundsätze vermittelte (Yapijakis 2009). Er lehrte die Fähigkeiten, Symptome zu erkennen, Diagnosen zu stellen und Therapien einzuleiten. Noch heute benutzen die Mediziner Begriffe, die maßgeblich von Hippokrates geprägt wurden, wie Diabetes, Gastritis oder Epilepsie (Yapijakis 2009). Damit legte er die Grundlagen für eine Vervielfältigung und Weitergabe seines Wissens.

Die Formate „Vorlesung“ und „Unterricht am Krankenbett“, wie sie zum Teil bereits von Hippokrates verwendet wurden, waren über Jahrhunderte Hauptinstrumente in der medizinischen Lehre. Auch noch heute sind diese Lehrformate in der Approbationsordnung fest verankert.

Im Laufe des 20. Jahrhunderts erzielte schließlich die technische Entwicklung große Erfolge: Mit der Geburtsstunde des Internets im Jahr 1957 sowie der Entwicklung von Film und Fernsehen zu Beginn des 20. Jahrhunderts schritt eine zunächst langsam verlaufende, aber immer weiter fortschreitende Revolution in den Bereichen der Wissensvermittlung und Kommunikation voran (Walsh 2015). Mit dem stetigen Ausbau der Technik begann auch die Entwicklung und der Ausbau neuer, medien-basierter Lehransätze in der universitären Lehre (Irby und Wilkerson 2003). Vorlesungen erhielten einen Wandel von reinen Vorträgen aus Büchern über Dias hin zu digitalen Folien (Niamtu 2001). Ebenso fanden Video-Demonstrationen Einzug in die moderne universitäre Lehre (Schittek Janda et al. 2005). Unter anderem gelang es durch den Einsatz digitaler Instrumente in den Bereichen Chirurgie

(Hoff et al. 2014; Tarpada et al. 2016), Pathologie (Kuo und Leo 2019) und Radiologie (Ackermann et al. 2010), die Fähigkeiten der Studierenden zu stärken. *Gamification-tools*, wie digitale Notaufnahme-Simulationen (Middeke et al. 2018), die darauf abzielen, dass Studierende ihr erlerntes Wissen über verschiedene Krankheitsbilder auf virtuelle Patienten anwenden, nehmen mittlerweile ebenso einen gewissen Stellenwert in der (zumeist extracurricularen) medizinischen Ausbildung ein.

Im jetzigen Zeitalter ist an eine moderne zeitgemäße Lehre ohne technische Begleitung kaum zu denken. Die medizinische Lehre geht mit dem Trend der Digitalisierung und bleibt damit attraktiv für die junge Generation (Chumley-Jones et al. 2002; Ruiz et al. 2006). Wie wichtig und erfolgreich die neuen Errungenschaften aus der Digitalisierung sind, verzeichnete auch die Bundesregierung. So setzte sich diese die Ausweitung der Digitalisierung auf die Agenda und tätigte erhebliche Investitionen in diesem Bereich (Zawacki-Richter 2021). Von jener Ausweitung ist auch die universitäre Lehre maßgeblich betroffen.

Vor allem zur jetzigen Zeit, während sich das *Severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2) pandemisch ausbreitet, ist die Digitalisierung ein hochaktuelles Thema: Mit Beginn der SARS-CoV-2 Pandemie und den nicht mehr durchführbaren Präsenzveranstaltungen infolge der Kontaktbeschränkungen ab Frühjahr 2020, nahm die Relevanz der *Electronic-Learning*(*E-Learning*)-Formate rasant zu. Um den Lehrbetrieb aufrecht zu erhalten, wurde mit Hochdruck an neuen digitalen Lehrformaten gearbeitet. Hierzu konnten die Universitäten zum Teil auf die zuvor entwickelten technischen Möglichkeiten zurückgreifen (Details siehe in Kapitel 1.4).

1.2 Entwicklung der digitalen Lehre in der Humanmedizin

Schon frühzeitig wurde das Potenzial der neu entstandenen Technologien erkannt und für eine innovative medizinische Lehre genutzt (Masys 1998). Durch die Entwicklung von transportablen Geräten und den stetigen Ausbau des Internets gelang es, an fast jedem Ort und zu jeder Zeit erreichbar zu sein und auf Wissen zuzugreifen (Walsh 2015). Im Zuge dessen gab es eine rasante Entwicklung von unterschiedlichen digitalen Lehrkonzepten, deren Resultate in diversen Studien untersucht wurden.

Ein Review von George und Kollegen (George et al. 2019) zeigte in diesem Kontext auf, dass die digitale Lehre im Vergleich zur konventionellen Präsenzlehre mindestens gleichwertig anzusehen oder ihr sogar überlegen sei. Zu dem Entschluss kamen auch Brusamento und Kollegen (Brusamento et al. 2019) sowie Peine und Kollegen (Peine et al.

2016). In der randomisierten Studie von Peine und Kollegen wurden 224 Studierende in vier Kohorten eingeteilt, die jeweils unterschiedliche Lernmethoden verwendeten (E-Learning, Selbststudium, Vorlesungen und Seminare). In einer Multiple-Choice(MC)-Prüfung, die nach der Lehrintervention durchgeführt wurde, erzielten die Studierenden der *E-Learning*-Gruppe die besten Ergebnisse (Peine et al. 2016).

Die oben genannten Studien berichten weiterhin von einer hohen Effektivität des *E-Learnings* bzw. der modernen Selbstlernmethoden: Die Studien belegen eine bessere Erinnerungsfähigkeit der Studierenden an gelerntes Wissen durch *E-Learning*. Darüber hinaus bewerteten die Nutzer digitaler Lehrformate nicht nur den Erwerb und das Abrufen von Erlerntem aus dem Gedächtnis positiv; sie gaben auch an, mit höherer Motivation zu lernen und insgesamt sehr zufrieden mit den neuen Formaten zu sein (Kang et al. 2018).

Cook und Kollegen untersuchten dazu in einer Metaanalyse die Effektivität der digital-gestützten Lehre im Vergleich zu traditionellen Lehrmethoden (Cook et al. 2008). Sie stellten fest, dass im Vergleich zu keiner Intervention die digitalen Lehrmethoden positive Auswirkungen auf das Wissen und die klinischen Fähigkeiten haben. Im Vergleich zu traditionellen Lehrmethoden zeigten sich hingegen sehr heterogene, wenig aussagekräftige Effekte unter den Studien, die in die Metaanalyse mit einbezogen wurden. Ein Einbeziehen von Diskussionen und ein hohes Maß an Interaktivität in den digitalen Lehransätzen kann zusätzlich zu einer Steigerung der Effektivität beitragen. Dementsprechend kamen Cook und Kollegen zu dem Resultat, dass sowohl die digital-gestützte Lehre als auch die traditionellen Lehrmethoden ähnliche Effekte erzielen (Cook et al. 2008). In dem Artikel von 2010 untersuchten Cook und Kollegen weiterhin die Auswirkungen der digitalen Lehre und beurteilten deren Effektivität (Cook et al. 2010). Sie kamen zu dem Entschluss, dass praktische Übungen, Feedback und angesetzte Wiederholungen sinnvoll sind und Lernerfolge erhöhen können. Allerdings erwähnten die Autoren auch, dass die Studienlage in Bezug auf die Effektivität sehr heterogen ist (Cook et al. 2010).

Studien wie die von Leong und Kollegen (Leong et al. 2012) oder Chumley-Jones und Kollegen (Chumley-Jones et al. 2002) wiesen durch *E-Learning*-Formate zwar einen verbesserten Wissenserwerb nach, sehen jedoch das größte Potenzial in einer Kombination aus traditioneller Präsenzlehre und modernen *E-Learning*-Formaten. Bereits 2006 postulierten Ruiz und Kollegen, dass zukünftig die modernen digitalen Lehransätze als erfolgsversprechende Methode Teil der *blended-learning*-Strategie sind und damit zusätzlich zur konventionellen Lehre vermehrt angeboten werden (Ruiz et al. 2006). Etwa zehn Jahre später berichteten die Arbeitsgruppe von Taveira-Gomes und Kollegen von diversen digitalen

Formaten und Softwaresystemen, die auf Basis der bis dahin vorliegenden Studien zum Thema *E-Learning* zur Verbesserung der Lehre entwickelt wurden (Taveira-Gomes et al. 2016).

Abschließend ist zu sagen, dass es bereits vor der Pandemie schon unzählige Studien im Bereich *E-Learning* gab, die die Effektivität, Zufriedenheit, Kosten und weitere Kriterien der Lehre analysierten. Dennoch wurde in einem solchen flächendeckenden Ausmaß die digitale Lehre im Medizinstudium bislang noch nicht eingesetzt.

1.3 Beginn und Folgen der Covid-19 Pandemie auf die universitäre Lehre

Mit dem Beginn der *Coronavirus disease* 2019(Covid-19)-Pandemie standen Lehrkräfte weltweit vor großen Herausforderungen. So auch an der Medizinischen Fakultät der Universitätsmedizin Göttingen.

Bis Anfang März 2020 war das SARS-CoV-2 Virus weitgehend unbekannt. Damit standen Ärzte- und Pflegschaft vor neuen Herausforderungen, um infizierte Patienten*innen nach dem aktuellen Wissensstand zu versorgen. Gleichzeitig erhielt, angesichts der globalen Pandemie, die Ausbildung der künftigen Ärzte*innen eine enorme Aufmerksamkeit und Dringlichkeit. Dennoch war genau diese Ausbildung zu Zeiten der Pandemie maßgeblich erschwert. Dozierende waren meist gleichzeitig Ärzte und Ärztinnen an den universitären Kliniken und mussten die Versorgung der Erkrankten aufrechterhalten. Hinzu kam, dass niemand absehen konnte, wie sehr die Lehrenden in die Krankenversorgung eingebunden sein würden. Dennoch musste eine Fortführung der Lehre gewährleistet werden.

Die nächste Schwierigkeit stellten die Kontaktbeschränkungen dar, die im Zuge der Pandemie erlassen wurden (Bundesregierung 2020).

Die einzige Möglichkeit, die Ausbildung der Medizinstudierenden ohne ein erhöhtes Infektionsrisiko durch Präsenzlehre fortzuführen, war die komplette Umstellung auf eine digitale Lehre. Daher wurde die Präsenzlehre weitgehend durch eine digitale Lehre ersetzt (Rose 2020).

1.4 Neue Lehrformate im Zuge der Covid-19 Pandemie

Dennoch war es für die Lehrenden an den medizinischen Fakultäten schwierig, mit digitalen Formaten die Lehre auf dem gleichen Niveau zu halten, wie vor der Pandemie - zu Zeiten,

als die medizinische Ausbildung ausschließlich in Präsenzveranstaltungen stattfand. Viele Dozierende arbeiteten unter Hochdruck an geeigneten Lösungen und brachten eine Vielzahl neuer Ansätze hervor, um eine optimale Lehre unter den Vorschriften der Pandemie zu gewährleisten. Einer der Ansätze bestand darin, dass *flipped-classroom-model* anzuwenden – Ein Modell dessen Effektivität bereits im Vorfeld in diversen Studien belegt worden ist (O’Flaherty und Phillips 2015; Tolks et al. 2016).

Auch für die Umsetzung des virtuellen Anatomieunterrichts gab es neue Ansätze und Fortschritte. Fairen und Kollegen untersuchten einen *virtual Reality*(VR)-basierten-Anatomieunterricht (Fairén et al. 2020). Mit dem Programm VR4Health konnten dabei menschliche Organe dreidimensional dargestellt werden. Lehrende konnten mit ihren Studierenden in Kontakt treten und allgemein die Lernmotivation und das räumliche Denken, bezogen auf den menschlichen Körper, verbessern.

Doch nicht nur in der Anatomie wurden solche virtuellen Formate für den Ausfall der Präsenzlehre entwickelt. Als weiteres Beispiel sind die Arbeiten von Aldeman und Kollegen zu nennen, die ein virtuelles Angebot im Fachbereich Pathologie für Studierende generierten und mit Erfolg implementierten (Aldeman et al. 2021). Sie entwickelten das Programm SmartPathk, in dem sich Studierende histologische Präparate von Nierenbiopsien, speziell mit dem Krankheitsbild Glomerulopathie, anschauen können und weiteres Fachwissen vermittelt bekommen. Dieses Training ist nicht nur zur jetzigen Zeit der weltweiten Pandemie besonders hilfreich, sondern wird auch in Zukunft einen bedeutsamen Nutzen haben: In Regionen, in denen wenige Nephrologen ansässig sind, kann durch dieses Instrument eine effektive Weiterbildung im Bereich der Pathologie betrieben werden (Aldeman et al. 2021).

Neben den erfolgreich eingebundenen digitalen Lehransätzen in Anatomie und Pathologie, boten auch andere Fachbereiche digitale Ersatzformate an. Unter anderem sind die Fachrichtungen Dermatologie (Kim et al. 2020; Patel et al. 2020) Chirurgie (Chang et al. 2020), Pädiatrie (Yang et al. 2021) und Notfallmedizin (Austin et al. 2021; Foohey et al. 2022) zu nennen. Ebenso entwarfen Knie und Kollegen eine Alternative zum Kommunikationstraining für Medizinstudierende an der Universität Witten/Herdecke (Knie et al. 2020).

1.5 Erste Resultate der Lehre unter Covid-19

1.5.1 Für und Wider der digitalen Lehre

Aufgrund der plötzlichen Entwicklung des globalen Notfalls standen die Lehrenden sowie Dekanate vor großen Herausforderungen. Vielerorts war die nötige Infrastruktur für eine rein digitale Lehre kaum bzw. unzureichend ausgebaut. Aufgrund der Dringlichkeit der Fortführung der medizinischen Ausbildung kam es jedoch dazu, dass nur unvollständig untersuchte und teils unreife virtuelle Formate eingesetzt werden mussten. Erschwerend kam hinzu, dass die Evidenz dieser digitalen Lernmethoden nur unzureichend erforscht und belegt war. Zwar sind in der Literatur vereinzelte Studien zur virtuellen Lehre zu finden, doch wurden diese nicht in einem solch breiten Spektrum durchgeführt, wie es zu Zeiten der Corona-Pandemie möglich war. Die meisten Studien waren zudem deskriptiv ausgelegt (Zuo et al. 2020).

Wie neueste Forschungsergebnisse zeigen, gibt es vorerst keinen Anhalt dafür, dass das Leistungsniveau der Studierenden im ersten Jahr durch den zwangsmäßigen Wechsel auf eine reine Online-Lehre gefallen sei (Nathaniel et al. 2021).

Eine Studie an der Universitätsmedizin Göttingen führte einen Vergleich zwischen Medizinstudierenden im klinischen Studienabschnitt durch (Seifert et al. 2021). Für den Vergleich wurden die Prüfungsergebnisse einer Kohorte vor dem Beginn der Corona-Pandemie und im ersten Pandemie-Semester (digitalisiert) herangezogen. Die Untersuchung zeigte dabei keinen statistisch signifikanten Unterschied bezogen auf die Prüfungsleistung. Dabei untersuchten Seifert und Kollegen auch die Zufriedenheit und Selbsteinschätzung (*comparative self-assessments*) der Studierenden während der Online-Lehre, die eine außerordentliche Disziplin und Selbstorganisation von den Studierenden forderte (Seifert et al. 2021). Damit war diese nicht die einzige Studie, die das Wohlbefinden und die Zufriedenheit der Studierenden mit der Online-Lehre untersuchten.

Auch die Arbeitsgruppe von Wiedenmann und Kollegen erhoben im Sommersemester 2020 bei Studierenden in der Augenheilkunde die Zufriedenheit im digitalen Semester (Wiedenmann et al. 2021). Die Studierenden erklärten, dass sie sich nicht nur gut auf die Abschlussprüfung vorbereitet fühlten, sondern auch in Zukunft die digitalen Videos für eine sinnvolle Ergänzung zum Präsenzunterricht hielten. Insgesamt gab es noch weitere Studien, in denen die digitale Lehre als Zugewinn gesehen wurde und von Studierenden durchaus positiv (vor allem in Bezug auf Flexibilität und Erlernen technischer Kompetenzen) bewertet wurde (Al-Balas et al. 2020; Dost et al. 2020; Elshami et al. 2021). Die Zufriedenheit gilt

dabei als wichtiger Einflussfaktor für eine hohe Leistungsfähigkeit, denn wie Judge und Kollegen zeigten, gibt es eine positive Korrelation zwischen Zufriedenheit und Leistungsfähigkeit (Judge et al. 2001).

Dennoch erlangte die digitale Lehre nicht nur positives Feedback. Auf der Kehrseite beklagten die Studierenden den fehlenden Kontakt zu Kommilitonen*innen und Dozenten*innen sowie eine insgesamt erschwerte Studiensituation (Dost et al. 2020; Lee et al. 2021; Wang et al. 2020). Die psychischen Belastungen einiger Studierender hätten während der Pandemie bzw. des Lockdowns zugenommen (Son et al. 2020; Zolotov et al. 2020). Fehlende soziale Interaktionen, Einsamkeit und geringe soziale Unterstützung seien dabei entscheidende Faktoren (Elmer et al. 2020). Hierbei stellten Elmer und Kollegen fest, dass besonders Studentinnen unter den fehlenden sozialen Interaktionen und weiteren Covid-19 geschuldeten Stressfaktoren psychisch litten (Elmer et al. 2020).

Eine erhöhte subjektive Arbeitslast und vermehrte notwendige Selbstkoordination sowie Überforderung mit diversen Online-Plattformen, die zur Lehre genutzt wurden, führten zu einem höheren Stresserleben (Son et al. 2020). Ebenfalls sei die Prävalenz der Depressionen, Angstzustände und Burnouts unter Studierenden (vor allem im letzten Semester des Studiums) angestiegen (Elmer et al. 2020; Islam et al. 2020; Zis et al. 2021).

Aus Studien wird zudem ersichtlich, dass eine finanziell angespannte Situation für zusätzliche Belastung sorgte: Durch die Pandemie verloren viele Studierende ihren Nebenjob (Tsurugano et al. 2021). Teils gerieten auch die sonst finanziell unterstützenden Eltern in wirtschaftliche Engpässe. Vereinzelt führte diese Belastung zu einer erniedrigten Studienmotivation und zur Beeinträchtigung der psychischen Gesundheit (Sauer et al. 2022; Tsurugano et al. 2021).

Auch die technische Ausstattung der Studierenden für eine reibungslose Umsetzung der digitalisierten Lehre wies Defizite auf. Laut Aristovnik und Kollegen mangelte es vor allem an ruhigen Arbeitsplätzen, funktionierenden Internetverbindungen und Zugängen zu technischen Geräten wie beispielsweise Druckern (Aristovnik et al. 2020). Zudem gaben einige Studierende an, dass ihre häusliche Wohnsituation für eine digitale Lehrintervention unpassend sei. Darunter falle nicht nur, dass die passende Ausstattung wie Rechner, Drucker etc. nicht vorhanden ist, sondern auch dass es zu einer vermehrten Ablenkung im häuslichen Umfeld kommt (Dost et al. 2020).

Aus den oben genannten Studien geht nur in unzufriedenstellendem Maße hervor, mit welcher Intensität sich die Studierenden mit den online zur Verfügung gestellten Ressourcen

auseinandersetzen und in welcher Weise sich dies auf die Leistungen des/der einzelnen Studenten*in der summativen Abschlussprüfung auswirkt. Besonders dringlich ist die Identifikation von Prädiktoren, die Schwierigkeiten der Studierenden mit der Anpassung an die digitale Lehre anzeigen. Solche Studierende werden als gefährdete Studierende während der Umstellung auf eine rein digitale Lehre betrachtet.

1.5.2 Bedeutung der Erkennung von gefährdeten Studierenden

Wie bereits Barber und Kollegen vor der Corona-Pandemie darlegten, ist es möglich, mit geeigneten Analysen gefährdete Studierende frühzeitig zu identifizieren (Barber et al. 2018). Zwar ist es nicht möglich, die Studierenden schon während der Zulassung auf ihre zukünftige Leistung beurteilen zu können, jedoch sollte dies innerhalb der ersten Jahre mit geeigneten Vorhersagewerten möglich sein. Barber und Kollegen erkannten, dass das Geschlecht beispielsweise ein solcher prädiktiver Wert ist (Barber et al. 2018). Dabei hat das weibliche Geschlecht ein niedrigeres Versagensrisiko als das Männliche.

Zur Zeit der Corona-Pandemie und der damit verbundenen Umstellung auf die digitale Lehre bleibt die Frage offen, inwiefern sich die Umstellung auf die rein digitale Lehre vor allem auf die Leistungsfähigkeit der männlichen Studierenden auswirkt. Denkbar wäre, dass Studenten, aufgrund der auf Technik-basierten Lehre, motivierter sind und effektiver lernen.

Bereits vor Jahren erkannten Marden und Kollegen die Bedeutung von gefährdeten Studierenden und entwickelten Quizfragen, um die sogenannten *students at risk* herauszufiltern (Marden et al. 2013). Durch die Studie entdeckten sie, dass diejenigen unter den Studierenden, die in den Quiz-Fragen schlecht abschnitten, auch eine geringere Prüfungsleistung erbrachten. Folglich sind die Quiz-Fragen ein Instrument, um gefährdete Studierende im Vorfeld zu identifizieren und ihnen eine mögliche Unterstützung anzubieten. Die Dringlichkeit liegt darin, solche Studierenden herauszufiltern und sie gesondert zu fördern. Möglicherweise können so vermehrte Studienabbrüche, Frustrationen oder gar psychische Erkrankungen wie Depressionen vermieden werden (Ludwig AB et al. 2015). Wie Heublein und Wolter feststellten, sind zu hohe Studienanforderungen und Leistungsprobleme Hauptrisikofaktor für die Entscheidung zum Studienabbruch (Heublein und Wolter 2011). Manche Studienabbrecher gaben außerdem an, dass eine mangelnde Studienmotivation und finanzielle Probleme sie zum Abbruch bewegt hätten.

Aus der Studienlage ist ersichtlich, dass gefährdete Studierende und die Erkennung solcher Studierenden seit Jahren Ziel der Forschung sind. Daher war ein weiteres Ziel der hier beschriebenen Studie, neben der Untersuchung der Nutzungscharakteristika der

Studierenden, auch Einflussfaktoren zu identifizieren, die dazu führen, dass Studierende von der digitalen Lehre profitieren. Ebenso sollten auch Faktoren ausfindig gemacht werden, die eine negative Auswirkung auf die studentische Leistung während der Online-Lehre haben. Frühe Prädiktoren für gefährdete Studierende zu erfassen, um diese möglichst rasch zu erkennen und ihnen gezielt vorbeugend Unterstützung anzubieten, war unter anderem Anspruch an die Studie.

1.6 Wissenschaftliche Fragestellung

Die weitreichenden Folgen der Corona-Pandemie forderten auch an der Universitätsmedizin Göttingen eine tiefgreifende Umstrukturierung der Lehre. Nachdem der Lehrbetrieb vorerst vom 16. März 2020 bis zum 20. April 2020 eingestellt wurde, begann das Sommersemester 2020 für jeden Studierenden mit ausschließlich digitalen Lehrveranstaltungen.

Für Lehrformate, die üblicherweise im Präsenzunterricht abgehalten wurden, musste ein digitales Ersatzformat gefunden werden. Dazu zählten unter anderem Vorlesungen, Praktika und Seminare, die mit oder ohne Patientenkontakt einhergingen. Als Ersatz für Vorlesungen, Praktika und Seminare entstanden Videoaufzeichnungen. Ergänzt wurde das Lehrprogramm durch Videokonferenzen, Chats und Webinare, in denen ein interaktiver Austausch zwischen Dozenten*innen und Studenten*innen stattfinden konnte. Während diverse Lehrformate online abgehalten wurden, sollten universitäre Prüfungen sowie der Unterricht am Krankenbett weiter in Präsenz am Ende des Semesters stattfinden (weitere Details in Kapitel 2.2).

Im Zuge der Umstrukturierung wurden Lehrpläne umgestellt und neue, alternative Ansätze zur Fortführung der Lehre entwickelt. Dabei wurden die neuen Ansätze in die Curricula aufgenommen, ohne dass ihre Wirksamkeit und Tragweite im Vorfeld ausreichend untersucht wurde. Generell gab es bis zum damaligen Standpunkt nur wenige Studien, die die Effektivität der digitalen Lehre untersucht haben. Zudem wiesen die wenigen Studien meist einen zu geringen Stichprobenumfang auf und untersuchten nur Teilaspekte der digitalen Lehre (Djermester et al. 2021; Heinzmann et al. 2021; Weissmann et al. 2021).

Erstmals werden in dieser Studie mehrere Teilaspekte des digitalen Lernens erfasst und simultan miteinander verglichen, um den Erfolg des digitalen Lernens insgesamt besser beurteilen zu können. Es wurde erhofft, aus dem beobachteten Lernverhalten Rückschlüsse auf individuelle Einflussfaktoren, die den Lernerfolg erhöhen oder gar erniedrigen, ziehen zu können. Daraus ergaben sich folgende Forschungsfragen: (I) In welcher Intensität und

mit welcher Verzögerung nach Bereitstellung nutzen Studierende der Humanmedizin digitale Ressourcen? Und: (II) Wie beeinflussen individuelle Charakteristika und Nutzungsentscheidungen den objektiv messbaren Lernerfolg im Kontext einer komplett digitalisierten Lehre?

Unter diesen Forschungsfragen war es Ziel der folgenden Analyse (1), die studentische Leistung vor der Pandemie und im ersten Semester nach Ausbruch des SARS-CoV-2 Virus unter Kontaktbeschränkungen miteinander zu vergleichen und eventuelle Unterschiede aufzudecken. Des Weiteren (2) wurde nach Zusammenhängen zwischen den Nutzungscharakteristika und den Ergebnissen der Abschlussprüfung zum Ende des Moduls gesucht, um daraus individuelle Einflussfaktoren auf den Lernerfolg abzuleiten. Eine Hypothese der Studie ist, dass eine intensivere Nutzung der digitalen Ressourcen, frühere Abgaben und höhere erzielte Prozentsätze in den unterschiedlichen Formaten auch zu einem besseren Ergebnis in der summativen Abschlussprüfung führen. Ebenso war es Ziel dieser Studie (3), besonders gefährdete Studierende zu entdecken, deren Charakteristika zu identifizieren und frühe Prädiktoren wahrzunehmen. Dies beruht auf dem Hintergrund, dass gefährdete Studierende möglichst früh erkannt werden und ihnen eine gezielte Unterstützung angeboten werden kann. Vor der Auswertung der Daten wurde angenommen, dass die Leistungserfolge und Ressourcennutzungen in der ersten Woche des Moduls einen Hinweis auf die spätere Leistung in der Abschlussklausur geben kann.

2 Material und Methoden

2.1 Bildungseinrichtung und Studienkohorte

Bei der vorliegenden Studie, die im Sommersemester 2020 an der Universitätsmedizin Göttingen durchgeführt wurde, handelt es sich um eine prospektive Kohortenstudie. Zum Sommersemester 2020 hatten sich 169 Studierende für das Modul M3.1 „Erkrankungen des Herz-Kreislauf-Systems und der Lunge“ angemeldet.

Das Studium der Humanmedizin ist in einen zweijährigen vorklinischen und einen dreijährigen klinischen Studienabschnitt unterteilt. Der klinische Studienabschnitt wird dabei durch einzelne Module aufgebaut, die in den jeweiligen Semestern abgehandelt werden. Das Modul M3.1 startet zu Beginn des dritten klinischen Semesters. Insgesamt dauert das Modul sechs Wochen, in denen wichtige Kenngrößen wie Ätiologie, Pathophysiologie, Symptome, Diagnostik, Therapie und Prognose über Erkrankungen des Herz-Kreislauf-Systems und der Lunge gelehrt werden.

Die Lehre im Modul M3.1 zielt besonders darauf ab, die künftigen Ärzte*innen bestmöglich auf ihre spätere Arbeit vorzubereiten. Daher besteht ein übergeordnetes Ziel der Dozierenden darin, den Wissenserwerb zu unterstützen und gleichzeitig ihre Fähigkeiten im *clinical reasoning* zu stärken. *Clinical reasoning* beschreibt die Anwendung von erlerntem Wissen auf neue komplexe, klinische Situationen, in denen Entscheidungen über diagnostische Verfahren und therapeutische Ziele getroffen werden müssen. Didaktisch gibt es viele Strategien, um das klinisch orientierte Denken zu fördern, unter anderem durch fallbasierten Kleingruppenunterricht oder Schlüsselmerkmalsfragen (Hrynychak et al. 2014; Raupach et al. 2016).

Für einen erfolgreichen Wissenserwerb und eine Stärkung des *clinical reasoning* waren zentrale Elemente der in Präsenz abgehaltenen Prä-Covid-19 Lehre die Vorlesungen, der fallbasierte Kleingruppenunterricht sowie der Unterricht am Krankenbett. Ergänzt wurden diese Kernelemente bereits vor der Pandemie schon durch einzelne digitale Lehrformate, zu den die Elektronische Fallseminare (E-Fallseminare), #clue und die virtuelle Notaufnahme EMERGE zählten.

In den E-Fallseminaren erhielten die Studierenden Patientenfälle, zu denen ihnen *key-feature*-Fragen gestellt wurden. In diesem Lehrkonzept wurde der direkte *testing effect* (Larsen et al. 2008) ausgenutzt und somit die Festigung des Wissens unterstützt und Fähigkeiten im *clinical reasoning* gestärkt (Ludwig S et al. 2018; Raupach et al. 2016; Wahlgren et al. 2006). Auch

#clue, eine App, die den Studierenden täglich zwei originale IMPP (Institut für medizinische und pharmazeutische Prüfungsfragen)-Fragen zur Verfügung stellte, verfolgte das Ziel, die Merkfähigkeit von Erlerntem zu erhöhen. Auf freiwilliger Basis konnten die Fragen, die sich auf die Vorlesungsinhalte des Vortages bezogen, gelöst werden. Dabei wurden zusätzlich Fragen von vorherigen Tagen wiederholt, um einen Abstand zwischen Lernen und Abfragen zu schaffen (Larsen et al. 2008). Ebenso wie die E-Fallseminare und #clue boten die Modulverantwortlichen EMERGE an; eine virtuelle Notaufnahme mit realen Patientenfällen. Durch das *serious game* versetzten sich Studierende in die Rolle eines*r Arztes*in hinein und trainierten dabei ihre Kompetenzen im *clinical reasoning* (Rutledge et al. 2018).

Die drei virtuellen Formate (E-Fallseminare, #clue und EMERGE) wurden auch im digitalen Sommersemester 2020 angeboten.

2.2 Änderungen im Ablauf des Moduls durch die Covid-19 Pandemie

2.2.1 Zeitliche Änderung

Kernziele des Moduls M3.1 „Erkrankungen des Herz-Kreislauf-Systems und der Lunge“ waren der Wissenserwerb und die Anwendung des erlernten Wissens im Sinne des *clinical reasoning*. Im Sommersemester 2020 wurde das ansonsten sechswöchige Modul M3.1 auf vier Wochen verkürzt. Diese Wochen standen zur Vermittlung der theoretischen Grundlagen durch Vorlesungen, Tutorials, E-Fallseminare und #clue zur Verfügung. Die Entscheidung zur Verkürzung des Moduls um zwei Wochen wurde bewusst getroffen, da die Dozierenden erwarteten, dass sie zum Ende des Semesters (im Juni/Juli 2020) eine Präsenzlehre und damit Wiederaufnahme des Unterrichts am Krankenbetts möglich sei. Auch die Theoriephasen der folgenden Module im dritten klinischen Semesters (Modul M3.2 „Erkrankungen der Niere und des Urogenitalsystems“ und Modul M3.3 „Erkrankungen des Blutes, des Knochenmarks und Grundlagen der Tumorerkrankungen“) wurden verkürzt. Somit entstand am Ende des Semesters eine Zeit von drei Wochen für den praktischen Unterricht am Krankenbett (siehe Abbildung 1).

Durch die insgesamt 15 Stunden *bedside*-Lehre vermittelten Dozierende klinische Kompetenzen in den Fächern Kardiologie (inklusive pädiatrische Kardiologie), Herz-Thorax- und Gefäßchirurgie sowie Pneumologie. Neben der *bedside*-Lehre erhielten die Studierenden gegen Ende des Semesters ebenfalls die Möglichkeit, ihr klinisches Wissen in der virtuellen Notaufnahme EMERGE zu prüfen.

2.2.2 Inhaltliche Änderungen

Da Lehrformate wie Vorlesungen und Kleingruppenunterricht durch die Bestimmungen des Infektionsschutzgesetzes nicht im gewohnten Maß stattfinden konnten (Bundesregierung 2020), wurde ein dezidiertes Konzept entwickelt, das im Wesentlichen auf drei Säulen aufbaute: (1) das Einführen einer asynchronen Lehre für maximale Flexibilität, (2) Änderungen anhand des *flipped-classroom-model's* und (3) die Kombination aus Kernelementen und fakultativen Zusatzangeboten.

Die Ausbildung der künftigen Ärzte*innen hatte weiterhin eine hohe Priorität. Doch die globale Pandemie stellte eine große Herausforderung für Ärzte*innen dar, die zeitgleich meist auch als Dozierende an den universitären Kliniken einen Lehrauftrag innehatten. Mit dem Beginn der Pandemie im März 2020 war noch nicht bekannt, wie sehr die Ärzte*innen in die medizinische Versorgung der Covid-19-Infizierten eingebunden sein würden. Aus diesem Grund entschieden die Lehrbeauftragten der Universitätsmedizin Göttingen, (1) die Lehre asynchron zu gestalten. Medizinstudierende und Lehrende mussten dadurch nicht gleichzeitig in einer Online-Veranstaltung anwesend sein. Die Entscheidung zu diesem Verfahren wurde zudem von der Sorge einer Überlastung kollaborativer Online-Plattformen bei Wiederaufnahme der Lehre beeinflusst.

Weiterhin griffen die Modulkoordinatoren (2) auf das *flipped-classroom-model* zurück. Hierbei erhalten Studierende wesentliche Inhalte des Curriculums durch Texte und Videos, die sie im Selbststudium weiter vertiefen. Zusätzlich stehen ihnen Aufgaben zur Vertiefung zu Verfügung. Zahlreiche Studien in diesem Bereich belegten bereits die hohe Effektivität dieses Konzepts (Hew und Lo 2018; Kühl et al. 2017; O'Flaherty und Phillips 2015; Tolks et al. 2016).

Ein weiterer Ansatz war (3) die Kombination aus Kernelementen und fakultativen Zusatzangeboten. Kernelemente dienten zur Lehre der im Curriculum verankerten fachlichen Inhalte in digitaler Form. Die fakultativen Zusatzangebote stellten eine weitere Option für Studierende dar, Wissen zu festigen und mit geeigneten Aufgaben zu überprüfen. Sie waren keine Voraussetzung für die abschließende Prüfung, stellten jedoch eine Steigerung der Studienqualität dar.

2.2.3 Angewandte Lehrformate

Um die in Kapitel 2.2.2 genannten Prinzipien umzusetzen, bot das Modul M3.1 den Studierenden im Sommersemester 2020 folgende Lehrangebote an: Vorlesungen und

Tutorials, Elektronische Fallseminare (E-Fallseminare), #clue und EMERGE. Ebenso konnten die Studierenden am Unterricht am Krankenbett teilnehmen. Während dieser Lehrinhalten wurden keine Studiendaten zur Auswertung erhoben. Die nachfolgende Abbildung verdeutlicht den zeitlichen Ablauf des Moduls M3.1 und bietet eine Übersicht über die verwendeten Lehrformate (siehe Abbildung 1).

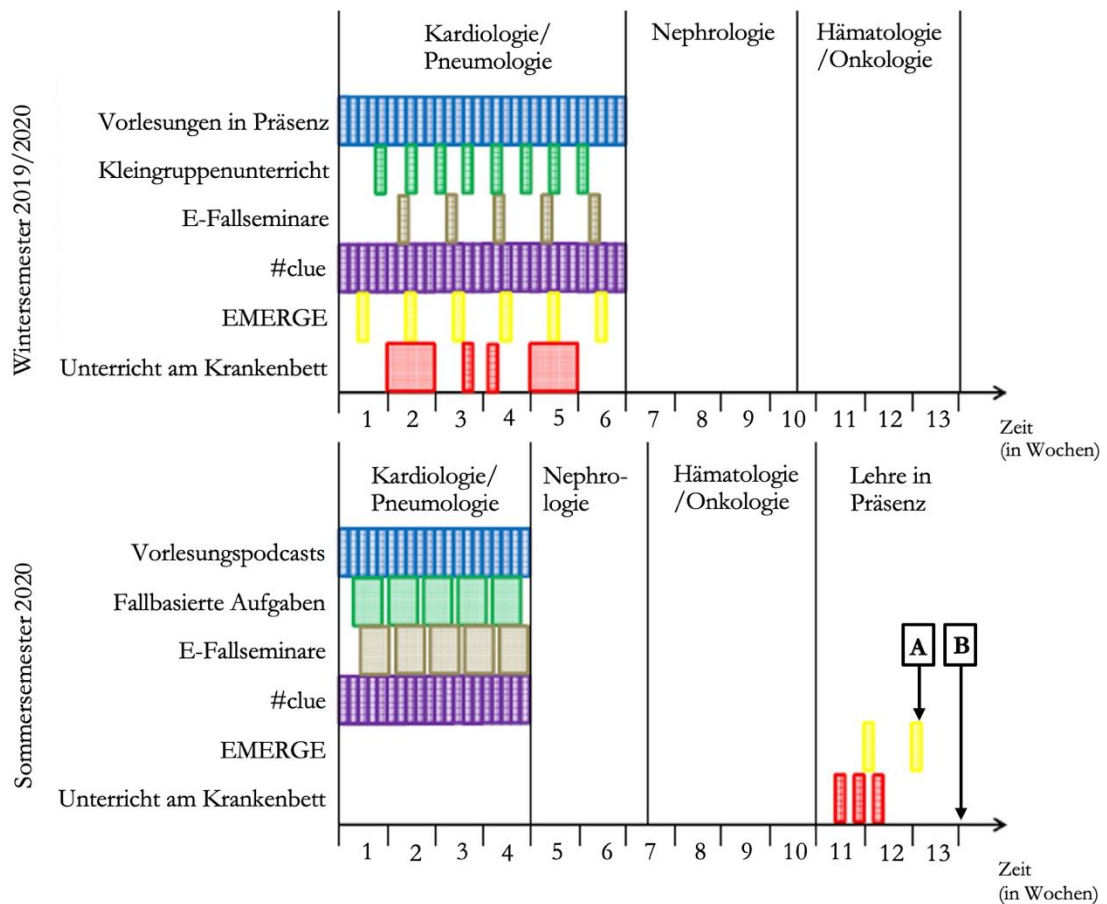


Abbildung 1: Zeitlicher Ablauf des Moduls M3.1 vor Beginn der Pandemie und im ersten Pandemie-Semester.

Im oberen Bildabschnitt ist der zeitliche Ablauf des Moduls M3.1 vor dem Beginn der Pandemie im Wintersemester 2019/2020 abgebildet. Im darunterliegenden Bildabschnitt sind die zeitlichen Veränderungen im Sommersemester 2020 dargestellt. Die jeweils oben genannten Fachrichtungen bilden das zugehörige Modul ab. A kennzeichnet die zweite EMERGE-Sitzung, in der die Studiendaten erhoben wurden; B kennzeichnet den Zeitpunkt der Modulabschlussprüfung.

Vorlesungen und Tutorials waren auch schon vor der Corona-Pandemie ein eingesetztes Lehrformat im Modul 3.1. Mittels Camtasia Studio 8 (TechSmith) wurden die Vorlesungen als Podcasts aufgezeichnet und über das Content-Management-System StudIP der Universität den Studierenden zur Verfügung gestellt. Die Anzahl und Themen der Vorlesungen veränderte sich dabei nicht. Da Vorlesungen als freiwilliges Angebot galten, fand in diesem Format keine Erhebung von Daten bezüglich der Ressourcennutzung statt.

Zu Zeiten der Präsenzlehre diente der Kleingruppenunterricht, die sogenannten Tutorials, zur Vertiefung der Vorlesungsinhalte. Die acht Tutorials bestanden aus einer 90-minütigen Lehrintervention, in denen, passend zu den Themen der Vorlesung, Patientenfälle vorgestellt und gemeinsam diskutiert wurden. Die Patientenfälle, die zur gemeinsamen Diskussion in den Lehrinterventionen anregten, bezogen sich auf folgende Erkrankungen: Myokardinfarkt, Herzinsuffizienz mit Vorhofflimmern, Vorhofseptumdefekt, chronisch obstruktive Lungenerkrankung und Lungenarterienembolie.

Im Zuge der Corona-Pandemie wurde dieser Unterricht unter Zuhilfenahme des *flipped-classroom-model's* in ein alternatives Lehrformat umgewandelt: Die als Vorlesungsersatz erstellten Podcasts gaben den Studierenden die Möglichkeit, sich Fachwissen anzueignen. Aus der zu jeder Zeit möglichen Nutzung der Podcasts entstand eine hohe Flexibilität, mit der die Studierenden sich eine eigene Tagesstruktur schaffen konnten. Um das erarbeitete Wissen weiter zu vertiefen, erstellten Dozierende fünf fallbasierte, patientenbezogene Aufgaben, mit den gleichen Schwerpunktthemen bzw. Erkrankungen wie zuvor im Kleingruppenunterricht in Präsenz. In der kontaktlosen, digitalen Lehre war es vorgesehen, die Aufgaben, die sich größtenteils auf das Patientenmanagement bezogen, schriftlich zu lösen.

Die Bearbeitung der einzelnen Fallvignetten erfolgte individuell innerhalb von fünf Tagen. Die Freischaltung der Fallvignetten erfolgte jeweils montags. Spätestens am Freitag derselben Woche musste die schriftliche Ausarbeitung abgegeben werden. Im Vorfeld erfolgte die Zuteilung der Studierenden zu einer Tutorial-Gruppe und ihrem Dozierenden. Nach der Bearbeitung der Patientenfälle bekam der jeweilige Tutorial-Dozierende die schriftlichen Ausarbeitungen zugeschickt. Auf Grundlage eines vordefinierten Bewertungsschlüssels erhielt jeder Studierende ein individuelles Feedback. Das Feedback wurde dabei ebenfalls in schriftlicher Form zugesendet. Des Weiteren standen die Tutorial-Dozierenden für Rückfragen zur Verfügung, sodass trotz der digitalen Lehre eine direkte Kontaktaufnahme möglich war und Fragen geklärt werden konnten. Damit verfolgten die modifizierten Tutorials das Bildungsziel (BZ) Wissensanwendung und Stärkung des *clinical reasoning*.

Eine Bewertung der Leistung fand nicht statt. Es konnte jedoch von jedem Studierenden eingesehen werden, zu welchem Zeitpunkt er die Aufgaben abgegeben hat. Anhand der Daten konnten die Zeiträume zwischen dem Bereitstellen der Aufgaben und dem Hochladen der Ausarbeitungen berechnet werden. Studierende, die drei oder mehr fallbasierte Aufgaben eher als 50 % der Kohorte abgaben, wurden als *early-completer* eingestuft.

Als weiteres Lehrangebot bot das Modul M3.1 elektronische Fallseminare an, deren Implementierung bereits in vorherigen Semestern erfolgreich war. Diese Seminare wurden bislang im Digitalen Prüfungs- und Schulungszentrum (DIPS) der Universitätsmedizin Göttingen in Präsenz abgehalten. Während des Präsenz-Seminars erhielten die Studierenden fallbasierte Fragen zum diagnostischen und therapeutischen Vorgehen. Insgesamt standen pro Sitzung vier Patientenfälle mit jeweils fünf Fragen zur Bearbeitung bereit. Die Fragen, die als *key-feature*-Fragen konzipiert waren, waren auf die Vorlesungsinhalte der vorangegangenen Woche abgestimmt. Dabei hatten die Studierenden nicht nur eine aus fünf möglichen Antworten auszuwählen, sondern wählten aus einem *long menu* eine Antwort aus. Dazu wurde von den Studierenden gefordert, die ersten Buchstaben der richtigen Antwort in ein Dialogfeld einzufügen, woraufhin eine definierte Liste mit Begriffen erschien. Dementsprechend führt das gewählte Fragen-Format zum Abrufen des erlernten Wissens aus dem Gedächtnis. Zusätzlich eignen sich *key-feature*-Fragen dazu, das klinisch orientierte Denken der Studierenden zu prüfen (Hrynchak et al. 2014; Kopp et al. 2006). Neben dem Format der Fragen, tragen auch die wiederholten Testungen durch die E-Fallseminare zur langfristigen Speicherung des Erlernten bei. Durch ein wiederholtes Testen des Erlernten werden die studentischen Fähigkeiten nicht nur bewertet, sondern Fakten bewusst wieder ins Gedächtnis gerufen und das Lernen effektiv verbessert (Roediger und Karpicke 2006). Dabei handelt es sich um das Konzept des direkten *testing effect's*.

Im Rahmen der notwendigen Digitalisierung aufgrund der Pandemie, mussten die Seminare, wie auch alle anderen Lehrangebote, digitalisiert werden. Dazu stellten die Modulkoordinatoren die bereits in den vorherigen Semestern entworfenen, fallbasierten Aufgaben auf der Online-Plattform ILIAS (Integriertes Lern-, Informations- und Arbeitskooperations-System) zu Verfügung. Von dort aus konnten die Aufgaben bearbeitet werden. Auch hier bekam jeder Studierende zu jedem bearbeiteten E-Fallseminar ein individuelles Feedback per E-Mail zugeschickt. Korrigierte Fehler konnten dadurch auf eventuelle Wissenslücken hinweisen und zu einem vertiefenden Lernen anregen. Ein wesentlicher Effekt dieses Lehrformats ist die Stärkung des *clinical reasoning* (Hrynchak et al. 2014; Wahlgren et al. 2006). Durch die wiederholten Testungen mit Ausnutzen des *testing*

effect's sowie dem individuellen Feedback wird zudem eine Steigerung der Lerneffektivität erreicht (Larsen et al. 2009; Yeh und Park 2015).

Insgesamt gab es fünf elektronische Seminare. Jede richtig beantwortete Frage ergab einen Punkt. Dementsprechend konnte jeder Studierende in jedem E-Fallseminar maximal 20 Punkte erreichen. Durch die fünf angebotenen E-Fallseminare bestand daher die Möglichkeit, insgesamt 100 Punkte zu sammeln. Die E-Seminare konnten jeweils bis zu acht Tage nach Freischaltung bearbeitet werden. Die Bearbeitung bestand auf freiwilliger Basis, wurde jedoch sehr empfohlen. In der Auswertung des Seminars konnte erhoben werden, wie viele Punkte jeder Studierende erreicht hat, Datum und Uhrzeit der Bearbeitung, wie viel Zeit er pro Frage gebraucht hat und mit welcher Verzögerung (von Bereitstellung bis Bearbeitung der Fragen) er die Fragen beantwortet hat. Dabei wurden Studierenden als *early-completer* eingeteilt, wenn sie drei oder mehr E-Fallseminare eher als 50% der Kohorte bearbeiteten.

Wie auch bei allen anderen Modulen im klinischen Studienabschnitt liegt das Ziel nicht nur darin, möglichst viel Wissen zu vermitteln und die Studierenden bestmöglich auf ihren späteren Alltag vorzubereiten, sondern sie auch für das kommende Examen zu rüsten. Um entsprechende Übungen zu ermöglichen, wird seit dem Wintersemester 2019/2020 das Lehrformat #clue angeboten. Dabei können Studierende mit Hilfe einer App täglich zwei originale IMPP-Fragen beantworten. Die MC-Fragen sind dabei inhaltlich auf den jeweiligen Modul-Tag abgestimmt, sodass das angeeignete Wissen kontinuierlich überprüft werden kann. Zahlreiche Studien belegen, dass *Gamification*-Bausteine potenziell die Motivation erhöhen und zu einer höheren Prüfungsleistung führen (McCoy et al. 2016; Rutledge et al. 2018).

Das Angebot, #clue in Anspruch nehmen zu können, bestand auch im Sommersemester 2020. In den vier Wochen des Moduls (mit jeweils fünf Werktagen) konnten die Studierenden 40 Fragen (zwei Fragen pro Tag) auf freiwilliger Basis beantworten. Dabei hatten sie pro Frage eine Bearbeitungszeit von zehn Tagen. Im Anschluss an die Beantwortung, wurden die Fragen in einer Liste gespeichert. Nach zehn Tagen waren die Fragen nicht mehr einsehbar. Der Bearbeitung einer Frage folgte ein vertiefendes Feedback. Zudem hatten die Studierenden die Möglichkeit, über die App ihren Leistungsstand und -verlauf zu sehen, sowie diesen Stand mit den Leistungen der anderen Studierenden in der Kohorte zu vergleichen. Die Fragen dienten dem Bildungsziel Wissenserwerb und Wissensbewahrung.

In der Auswertung konnten folgende Daten erhoben werden: Datum und Uhrzeit der Ausarbeitung, Verzögerung der Beantwortung der Fragen, die Anzahl beantworteter und die Anzahl richtig beantworteter Fragen. Erledigten die Studierenden 50 % der Fragen am Tag der Freischaltung oder am darauffolgenden Tag, galten sie als *early-completer*.

Seit dem Sommersemester 2017 gehört EMERGE zum regulären Lehrangebot im Modul M3.1 an der Universitätsmedizin Göttingen. Dabei handelt es sich um eine virtuelle Notaufnahme, in der virtuelle Patienten*innen mit realistischen Erkrankungen, wie sie auch im späteren Klinikalltag zu finden sind, behandelt werden müssen. Bei der Auswahl der Fälle wurde weiterhin darauf geachtet, dass es einen engen Bezug zu den bereits gelernten Inhalten aus dem Modul M3.1 gibt.

Da das Programm für EMERGE ausschließlich auf Computern im DIPS der Universitätsmedizin Göttingen installiert ist, konnte dieses Format lediglich in Präsenz stattfinden. Jeder Studierende bekam zwei Termine zugeteilt, an denen er im DIPS die virtuellen Patienten*innen behandeln konnte. Der erste 90-minütige Termin galt einer Übungsphase zum Kennenlernen des Programms. In der zweiten 90-minütigen Sitzung konnte der Student bis zu acht virtuelle Patientenfälle bearbeiten. Die Auswertung der Daten bezieht sich hierbei allein auf die zweite Sitzung. Dabei war die zweite Sitzung so konzipiert, dass die Studierenden die virtuellen Patientenfälle in zufälliger Reihenfolge erhielten. Jedoch bearbeitete jeder Studierende den Fall „ambulant erworbene Pneumonie“ sowie den Fall „infektiöse Endokarditis“. Die Antworten bzw. Handlungen der Studierenden zu diesen beiden Patientenfällen waren die Ausgangswerte für die anschließende Analyse.

In dem *serious game* EMERGE erhielten die zukünftigen Ärzte*innen den Auftrag, eine Anamnese zu erheben und auf Grundlage dessen die ersten diagnostischen Maßnahmen einzuleiten (Abbildung 2). Nach den Auswertungen der virtuellen diagnostischen Tests wurde der Studierende dazu angehalten eine Diagnose zu stellen, die geeignete Therapie für diese Diagnose einzuleiten und den Patienten auf eine periphere Station zu verlegen. Im Anschluss an die virtuelle Behandlung des*der Patienten*in war es zudem notwendig, eine schriftliche Übergabe zu verfassen.



Abbildung 2: Screenshots einer EMERGE-Spielsitzung (Middeke et al. 2018).

Oben links ist die Anmeldung und Ankunft eines virtuellen Patienten*in zu sehen; oben rechts wird die Übergabe des Patienten*in durch den Rettungssanitäter*in an den Arzt*in dargestellt; unten links sind die vom Spieler*in auszuwählenden Laborparameter zu sehen; unten rechts wird das Erheben einer Anamnese simuliert (Middeke et al. 2018).

Wie in der prospektiven Studie von Middeke und Kollegen gezeigt, kann durch EMERGE das klinisch orientierte Denken effektiv geübt und verbessert werden (Middeke et al. 2018). Sie berichtete zudem, dass EMERGE eine vielversprechende Alternative zum intensivierten Kleingruppenunterricht darstellt. Die Verbesserung des klinischen Denkens durch Computersimulationsprogramme belegten auch Jawaid und Kollegen in ihrer Studie (Jawaid et al. 2019).

Im Sommersemester 2020 behielten die Modulverantwortlichen teils alt bewährte Lehrformate bei, teils entwickelten sie neue alternative Lehrformate. Dabei galt es, sowohl den Wissenserwerb, als auch die Wissensanwendung im Sinne des *clinical reasoning* zu stärken. Den angewandten Lehr- und Prüfungsformaten wurden dabei jeweils ein Bildungsziel zugeordnet (siehe Abbildung 3) (Biggs 1996). Während die Vorlesungen und teils auch #clue (durch das detaillierte Feedback) als Ziel die Wissensvermittlung/-erwerb verfolgten, überprüften die fallbasierten Aufgaben und die E-Fallseminare eine Anwendung des erlernten Wissens.

2.3 Messung der Lernergebnisse

Wissenserwerb und Wissensanwendung stellten zwei Kernziele des Modus M3.1 „Erkrankungen des Herz-Kreislauf-Systems und der Lunge“ dar. Jedoch bilden der Erwerb und die letztliche Anwendung des Wissens zwei vollkommen unterschiedliche Kompetenzen ab und werden daher mit unterschiedlichen Instrumenten geprüft. Dazu wurden zwei divergente Prüfungsformate am Ende des Semesters gewählt. Während die Prüfung des Wissenserwerbs durch eine summative Abschlussprüfung mittels MC-Fragen erfolgte, konnte die Wissensanwendung durch die richtigen Handlungsschritte in EMERGE beurteilt werden (siehe Abbildung 3).

Die summative Abschlussprüfung im Sommersemester 2020 fand am Ende des Semesters nach der dreiwöchigen praktischen Phase statt. In der Prüfung beantworteten die Studierenden 45 MC-Fragen, die Themen zu den Bereichen Kardiologie inklusive pädiatrischer Teilaspekte, Herz-Thorax-Gefäßchirurgie sowie Pneumologie abbildeten. Erfahrene Dozierende in den oben genannten Bereichen verfassten die Fragen und stellten die Abschlussprüfung zusammen. Ein Gremium nahm zudem eine nochmalige Prüfung der Fragen vor. Die Studierenden erhielten die Klausur in Papierform und mussten unter Zuhilfenahme eines Stifts die richtige Antwort auf einem Antwortbogen markieren. Im Anschluss wertete eine zentrale Bewertungsstelle den Antwortbogen aus und lieferte die Roh- sowie Prozentwerte.

Um das Leistungsniveau im neuen digitalen Modul mit vorherigen Leistungen vergleichen zu können, wurden die Prüfungsergebnisse aus dem vorherigen Semester herangezogen. Die Ergebnisse der Klausuren aus den Modulen M2.1 „Grundlagen der Pharmakologie und Toxikologie“, M2.2 „Grundlagen von Infektion und Abwehr“ und M2.4 „Methodische Grundlagen der Evidenzbasierten Medizin“ wurden für jeden Studierenden entnommen und die in der Prüfung erreichten Prozentwerte in einen Perzentilränge umgerechnet. Für den direkten Vergleich wurden die erreichten Prozentwerte aus der summativen Abschlussprüfung des Moduls M3.1 ebenfalls in Perzentilränge umgerechnet.

Die Anwendung des erlernten Wissens prüfte EMERGE durch die Behandlung der virtuellen Patienten. In dem *serious game* erhielten die Studierenden für ihre richtigen Handlungsschritte Punkte, die anschließend in Prozentwerte überführt wurden. Insgesamt konnten so maximal 31 Punkte erreicht werden. Die erreichten Prozentwerte flossen nicht in die Abschlussnote des Moduls mit ein, das heißt es handelt sich um eine formative Prüfung.

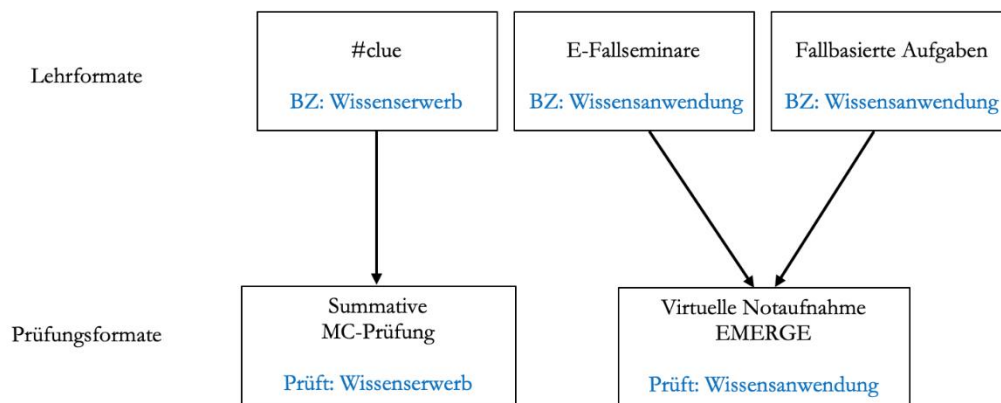


Abbildung 3: Lehr- und Prüfungsformate nach Bildungszielen kategorisiert.

Abgebildet sind die verschiedenen Lehrformate, die im Modul M3.1 den Studierenden angeboten wurden mit ihren jeweiligen Bildungszielen (BZ). Die Bildungsziele (Wissenserwerb und Wissensanwendung) werden durch unterschiedliche Prüfungsformate abgefragt. Die Hypothese war, dass eine intensivere Nutzung der digitalen Ressourcen sich auf das Prüfungsformat mit gleichem BZ auswirkt (verdeutlicht durch die Pfeile). Nicht mit in die Abbildung aufgenommen wurden der Unterricht am Krankenbett und die Vorlesungen, da während dieser Lehrformate keine Daten zur Auswertung erhoben wurden.

2.4 Datensammlung und statistische Analyse

Mit dem positiven Votum der Ethikkommission der Universitätsmedizin Göttingen vom 16.07.2020 (Antragsnummer 38/7/20) begann die Erhebung der Daten im Sommersemester 2020. Per E-Mail erhielten die Studierenden zum Anfang des Semesters über das Vorhaben der Studie eine Benachrichtigung. Die möglichen Teilnehmer der Studie erhielten unter anderem Informationen über Gründe sowie Ziele der Studie. Des Weiteren gab es Raum für Rückfragen. Da die EMERGE-Sitzungen in Präsenz stattfanden, holten die Studienverantwortlichen zu diesem Zeitpunkt die schriftliche Einwilligung ein. Auch während der EMERGE-Sitzungen standen die Studienverantwortlichen für Rückfragen zur Verfügung. Die Entscheidung für oder gegen eine Teilnahme an der Studie hatte keine Auswirkungen auf die Teilnahme an den Lehrangeboten oder an der Abschlussprüfung.

Für eine reibungslose Kommunikation übermittelte das Studiendekanat die Namen sowie E-Mail-Adressen der an der Studie teilnehmenden Personen. Dieses stellte auch die Prüfungsergebnisse aus dem Vorsemester bereit, die in den Vergleich der Leistungen mit einfließen. Für die jeweilige Zuordnung der Studierenden zu ihren Leistungen aus dem vorherigen Semester sowie Sommersemester 2020 war die Angabe der Matrikelnummer

notwendig. Diese gaben die Studierenden selbst an. Alle personenbezogenen Daten wurden nach den Vorgaben des Datenschutzes auf einem Rechner mit Standort im UBFT (Untersuchung, Behandlung, Forschung, Theorie)-Gebäude der Universitätsmedizin Göttingen gespeichert, sodass nur der Projektleiter zu den Daten Zugang hatte.

Die Nutzungsdaten der digitalen Lehrformate wurden elektronisch erhoben. Dabei erfolgte die Zuordnung zu jedem einzelnen Studierenden über die Login-Namen. Somit konnten Daten wie Datum, Uhrzeit und Dauer der Bearbeitung präzise gemessen werden. Die App #clue verwaltete die Daten in pseudonymisierter Form. Nur von Probanden, die in die Teilnahme der Studie einwilligten, wurden die Daten gespeichert. Sie konnten im Nachgang über einen individuellen Code mit den Stammdaten der Studierende zusammengeführt werden.

Nach Abschluss der Datensammlung erfolgte die Zusammenführung der unterschiedlichen Messungsergebnisse. Nicht benötigte personenbezogene Daten, wie Name, Matrikelnummer etc., wurden gelöscht. Im Anschluss begann die Analyse der Daten.

Alle in dieser Studie erhobenen Daten (Daten aus dem Modul M3.1 sowie Leistungsergebnisse aus dem vorherigen Semester) wurden in IBM SPSS Statistics 25.0 (IBM Corp, Armonk, NY, USA) überführt. Zunächst erfolgte die Durchführung der deskriptiven Analyse der Kohorte. Die daran anschließenden Analysen verfolgten diese Ziele:

Zur (I) Beurteilung der Leistungsverläufe der Studierenden wurden die Ergebnisse der Prüfungen aus dem zweiten klinischen Semester eingeholt und mit den Prüfungsergebnissen der Abschlussklausur des Moduls M3.1 verglichen. Die erhaltenen Prozentwerte aus den Leistungsüberprüfungen der Module M2.1, M2.2 und M2.4 wurden dabei in Perzentilränge überführt. Für den Vergleich wurden auch die Prozentwerte aus der Abschlussklausur des Moduls M3.1 in Perzentilränge überführt. Da die Studierenden die Module aus dem dritten klinischen Semester (M3.1, M3.2, M3.3) nicht alle im Anschluss an das zweite klinische Semester absolvierten, wurden die Studierenden nach Kohorten betrachtet: Kohorte (1) nahm im Wintersemester 2019/2020 an den Modulen des zweiten klinischen Semesters teil; Kohorte (2) belegte diese Module bereits im Sommersemester 2019. Indem die Prozentwerte in Perzentilen umgerechnet wurden, war es möglich, die Kohorten zusammenzulegen. Daher wird im weiteren Verlauf auf die Einteilung der Kohorten nicht weiter eingegangen.

Durch die Überführung der Prozentwerte in Perzentilen war ein direkter Vergleich der Leistung vor der Pandemie im Präsenzsemester mit den Leistungen im ersten pandemiebedingten Online-Semester möglich. Nach der Berechnung der Perzentilränge

wurden die Leistungsverläufe anhand des Geschlechts der Studierenden gruppiert betrachtet, um mögliche Unterschiede aufzudecken. Dazu wurden die Perzentilränge in einem Streudiagramm dargestellt.

Zur (II) Untersuchung des Zusammenhangs zwischen Leistungen und Ressourcennutzung, wurden univariate und multivariate lineare Regressionen durchgeführt. Als abhängige Variable galten dabei die Leistungen in der Modulabschlussklausur sowie in EMERGE. Während die Leistungen der Abschlussklausur für den Wissenserwerb der Studierenden standen, spiegelten die Leistungen in EMERGE die Wissensanwendung wider. Es fungierten die unterschiedlichen Nutzungscharakteristika der Lehrformate fallbasierte Aufgaben, E-Fallseminare und #clue (Anzahl der bearbeiteten Fragen/Seminare, Anzahl richtiger Fragen, erreichte Prozentwerte, Verzögerung bei der Bearbeitung, Zeit pro Frage und benötigte Gesamtzeit) als unabhängige Variablen. Ebenso galten die Studierendencharakteristika Geschlecht, Alter sowie Perzentilränge in den Prüfungen des zweiten klinischen Semesters als unabhängige Variablen. Die Analyse verfolgte das Ziel, die zuvor aufgestellte Hypothese, dass eine intensive Nutzung digitaler Lehrformate mit dem Ziel Wissenserwerb und -bewahrung zu besseren Prozentwerten in der MC-Prüfung führe, während Formate mit dem Ziel Wissensanwendung zu höheren Prozentwerten im *serious game* EMERGE beitragen, zu bestätigen (siehe Abbildung 3).

Um (III) Prädiktoren für eine hohe oder niedrige Prüfungsleistung zu erkennen, wurden ebenfalls univariate und multivariate lineare Regressionen durchgeführt. Dabei galt als abhängige Variable die Prüfungsleistung (Prozentpunkte in der MC-Prüfung). Als unabhängige Variable galt das studentische Verhalten in Bezug auf die Ressourcennutzung in der ersten Hausarbeit, dem ersten E-Fallseminar und den ersten zehn #clue Fragen. Zu den gewählten Nutzungscharakteristika zählten die Bearbeitung des ersten E-Fallseminars/der ersten fallbasierten Aufgabe, die Anzahl der gelösten Fragen, die erreichten Prozentwerte, die Verzögerung bei der Bearbeitung und die benötigte Gesamtzeit. Ziel der Auswertung war es, die Hypothese zu prüfen, dass Studierende, die mit einer geringen Verzögerung die erste fallbasierte Aufgabe abgaben sowie das erste E-Fallseminar und die ersten zehn #clue-Fragen bearbeiteten, in der Modulabschlussklausur eine höhere Leistung zeigen. Im Gegensatz dazu zeige sich eine niedrigere Leistung in der Prüfung bei denjenigen Studierenden, die die Aufgabe und Fragen erst zu einem späteren Zeitpunkt bearbeiten würden. Ebenso sollte geprüft werden, ob Studierende, die in dem ersten E-Fallseminar sowie den ersten zehn #clue-Fragen höhere Prozentwerte erzielten, auch ein besseres Ergebnis in der Abschlussprüfung erreichten.

Die Ergebnisse werden in Mittelwert \pm Standardabweichung, Anteilen und Rängen angegeben. Ebenso wurden Effektstärken nach Cohens d berechnet. Zur Datenanalyse wurden T-Tests zur Darstellung von Gruppenunterschieden, Korrelationen sowie univariate und multivariate lineare Regressionsanalysen durchgeführt. Für die Regressionen wurden die Betakoeffizienten mit dem 95%-Konfidenzintervall (KI) angegeben. Das Signifikanzniveau wurde auf 5% festgelegt.

3 Ergebnisse

3.1 Charakteristika der Studienkohorte

Insgesamt nahmen 169 Studierende an den Lehrveranstaltungen des Moduls M3.1 teil. Von ihnen willigten 162 in die Teilnahme der Studie ein, was einem Rücklauf von 95,6 % entspricht. Die Stichprobengröße variiert in den einzelnen Analysen aufgrund fehlender Daten.

In der Stichprobe waren Frauen in der Überzahl mit einem Anteil von 67,9 % (110 Frauen, 52 Männer). Das durchschnittliche Alter betrug $24,8 \pm 3,5$ Jahre. Dabei waren die Männer signifikant älter als die Frauen (Männer: $25,7 \pm 3,5$ Jahre, Frauen $24,4 \pm 3,5$ Jahre; $p = 0,027$; $d = 0,38$). Mit einer Mehrheit von 75 % befanden sich die meisten Studierenden im dritten klinischen Semester, gefolgt von 19 %, die sich im vierten klinischen Semester befanden.

3.2 Analyse der Nutzungscharakteristika

3.2.1 Tutorials und fallbasierte Aufgaben

In den Tutorials bekamen die Studierenden fünf fallbasierte Aufgaben, die sie anhand des Wissens aus den Vorlesungen bearbeiten sollten. Gezeigt werden konnte, dass die Zeit zwischen der Bereitstellung der Tutorial-Aufgaben und der Einreichung der entsprechenden Lösungen durch die Studierenden mit dem Fortschreiten des Moduls abnahm. So wurde der erste Fall im Durchschnitt nach $5,6 \pm 5,8$ Tagen abgegeben, der zweite Fall nach $5,8 \pm 4,1$ Tagen und der Letzte nach $2,9 \pm 3,4$ Tagen (Fall drei: $4,2 \pm 3,0$ Tage, Fall vier: $3,5 \pm 2,3$ Tage). Die mittlere Verzögerung der Fälle ($4,4 \pm 2,9$ Tage) korrelierte schwach negativ, aber signifikant mit den Perzentilrängen der Prüfungsergebnisse aus dem vorherigen Semester ($r = -0,190$; $p = 0,018$).

Insgesamt zählten 53,4 % der Studierenden zu den *early-completern*. In der Gruppe der Männer war 50 % der Studenten *early-completer*, 50 % waren es nicht. Bei den Frauen überwog leicht die Anteil der *early completer*, 55 % der Studentinnen waren *early-completer*. Der Unterschied zwischen Männer und Frauen war dabei nicht signifikant ($p = 0,551$). Das Alter der *early-completer* betrug im Durchschnitt $24,7 \pm 3,0$ Jahre. Damit waren sie nicht signifikant jünger als die Gruppe der *non early-completer*, die im Mittel $24,9 \pm 4,1$ Jahre alt waren ($p = 0,659$).

Hinsichtlich der Perzentilränge in den Prüfungen des vorherigen Semesters ließen sich keine signifikanten Unterschiede zwischen *early-completern* und *non early-completern* feststellen (*early-completer*: $54,3 \pm 27,8$; *non early-completer*: $50,4 \pm 28,7$; $p = 0,390$; $d = 0,14$). Dieser nicht-signifikante Trend konnte auch bei den *early-completern*/ *non early-completern* in der Prüfung des Sommersemesters 2020 im Modul M3.1 gesehen werden. Studierende, die eher als 50 % der Kohorte drei oder mehr Aufgaben abgaben, erzielten in der Prüfung des Moduls M3.1 im Durchschnitt einen Perzentilrang von $51,8 \pm 29,6$. Diejenigen Studierenden, die drei oder mehr Aufgaben später als 50 % der Kohorte abgaben, erreichten hingegen im Mittel einen Perzentilrang von $49,4 \pm 28,5$. Wie bereits oben genannt, ist dieser Trend nicht signifikant ($p = 0,608$; $d = 0,08$).

3.2.2 Elektronische Fallseminare (E-Fallseminare)

Ungefähr die Hälfte aller Studierenden, 51,9 %, bearbeitete alle fünf E-Seminare; 72,3 % der Studierenden bearbeiteten mindestens vier Seminare (29 % erledigten drei oder vier Seminare). Nur 16 Studierende (9,9 %) nahmen das Lehrangebot nicht an.

Im Mittel erreichten die Studierenden, die mindestens ein E-Fallseminar bearbeitet hatten, $67,5 \pm 11,0$ % der erhältlichen Punkte. Es wurde ersichtlich, dass Studentinnen signifikant weniger Punkte (in Prozentwerten) in den E-Fallseminare erzielten als ihre männlichen Kollegen (Studentinnen: $65,6 \pm 10,8$ %; Studenten: $72,2 \pm 10,1$ %; $p = 0,001$; $d = 0,62$), obwohl sie signifikant mehr E-Fallseminare bearbeiteten (Studentinnen: $4,0 \pm 1,4$ E-Fallseminare; Studenten: $3,4 \pm 2,0$ E-Fallseminare; $p = 0,022$; $d = 0,37$). Unabhängig vom Geschlecht gab es eine positive Korrelation zwischen der Anzahl der bearbeiteten Aufgaben und den erreichten Punkten in den E-Fallseminaren ($r = 0,872$; $p < 0,001$). Zwischen dem Alter der Studierenden und den in den E-Fallseminaren erreichten Prozentwerten zeigte sich keine signifikante Korrelation ($r = 0,157$; $p = 0,059$). Allerdings konnten Korrelationen mit den Prozentwerten der Prüfungen nachgewiesen werden. Die erreichten Prozentwerte in den E-Fallseminaren korrelierten signifikant mit den Prozentwerten der Prüfungen aus dem vorherigen Semester ($r = 0,260$; $p = 0,002$) sowie mit denen der Prüfung des Moduls M3.1 ($r = 0,281$; $p = 0,001$).

Die mittlere Zeitspanne, die zwischen der Veröffentlichung und der Bearbeitung der E-Fallseminare lag, betrug $1,6 \pm 1,5$ Tage. In der Analyse der einzelnen Seminare konnte festgestellt werden, dass die Studierenden für die Bearbeitung des ersten Seminars mehr Zeit in Anspruch nahmen ($3,1 \pm 2,3$ Tage) als für die Bearbeitung der restlichen Seminare. Diese Seminare wurden im Durchschnitt am Tag der Bereitstellung oder am Folgetag bearbeitet

(Seminar zwei: $0,5 \pm 0,7$ Tage, Seminar drei: $1,4 \pm 1,8$ Tage; Seminar vier: $0,7 \pm 1,0$ Tage, Seminar fünf: $1,1 \pm 1,4$ Tage). Die weiblichen Studierenden bearbeiteten die E-Fallseminare im Mittel nicht signifikant später als die männlichen Studierenden (Männer: $1,5 \pm 1,6$ Tage; Frauen: $1,7 \pm 1,4$ Tage; $p = 0,498$; $d = 0,12$). Weiterhin wurde keine signifikante Korrelation zwischen der Verzögerung und den erreichten Prozentwerten in den E-Fallseminaren nachgewiesen ($r = -0,121$; $p = 0,145$). Jedoch zeigte sich eine signifikante negative Korrelation zwischen der Verzögerung in den E-Fallseminaren und den Prozentwerten, die die Studierenden in der Abschlussprüfung des Moduls M3.1 erzielten ($r = -0,187$; $p = 0,024$). Auch zwischen den Prozentwerten der vorherigen Prüfungen und der Verzögerung zeigte sich eine signifikante negative Korrelation ($r = -0,185$; $p = 0,028$). Zum Alter der Studierenden konnte keine Korrelation bezogen auf die Verzögerung, mit der die Studierenden die E-Fallseminare bearbeiteten, festgestellt werden ($r = 0,117$; $p = 0,158$).

Fast die Hälfte der Studierenden (45,9 %) in dieser Stichprobe waren *early-completer*. Unter den Frauen, sowie unter den Männern überwog in geringem Maße der Anteil der *non early-completer* (*non early-completer* Frauen: 54,8%; *non early-completer* Männer: 52,4%). Der Unterschied dabei war nicht signifikant ($p = 0,792$). Ersichtlich wurde in der Analyse, dass die *early-completer* im Gegensatz zu den *non early-completern* eine signifikant höhere Prozentzahl richtiger Antworten in den Fragen der E-Fallseminare erzielten (*early-completer*: $70,8 \pm 10,2$ %; *non early-completer*: $64,7 \pm 10,9$ %; $p = 0,001$; $d = 0,57$). Nicht nur in den E-Fallseminaren erzielten die *early-completer* einen signifikant höheren Prozentwert. Weiterhin wurde erkannt, dass diejenigen Studierenden, die zu den *early-completern* zählten, signifikant höhere Prozentwerte in den Prüfungen des vorherigen Semesters erzielten (*early-completer*: $83,4 \pm 6,8$ %; *non early-completer*: $80,1 \pm 7,9$ %; $p = 0,009$; $d = 0,45$). Auch in den Prüfungen des Moduls M3.1 hatten die *early-completer* signifikant höhere Prozentwerte als die *non early-completer* (*early-completer*: $88,6 \pm 6,9$ %; *non early-completer*: $84,8 \pm 9,0$ %; $p = 0,006$; $d = 0,47$).

Aus den Daten ging ebenso hervor, wie lange die Studierenden für die Beantwortung der Fragen gebraucht haben. Pro Frage brauchten sie im Durchschnitt $103,2 \pm 35,7$ Sekunden, für die Bearbeitung aller Fälle im Mittel $8534,0 \pm 3543,9$ Sekunden. Die benötigte Zeit pro Frage ($r = 0,183$; $p = 0,027$) sowie die aufgewandte Gesamtzeit ($r = 0,228$; $p = 0,006$) korrelierte dabei jeweils signifikant mit den erreichten Prozentwerten in den E-Fallseminaren. Die mittlere benötigte Zeit pro Frage unterschied sich zwischen Männern und Frauen nicht signifikant voneinander (Männer: $104,3 \pm 34,5$ Sekunden; Frauen: $102,7 \pm 36,4$ Sekunden; $p = 0,795$; $d = 0,05$). Ebenso gab es keinen Unterschied zwischen den Geschlechtern in Bezug auf die verbrauchte mittlere Gesamtzeit (Männer: $8257 \pm 2797,1$

Sekunden; Frauen: $8645,9 \pm 3810,8$ Sekunden; $p = 0,310$; $d = 0,12$). Das Alter der Studierenden korrelierte in dieser Stichprobe ebenfalls nicht signifikant mit der benötigten Zeit pro Frage ($r = 0,141$; $p = 0,090$) oder der benötigten Gesamtzeit ($r = 0,114$; $p = 0,169$). Weiterhin konnte keine signifikante Korrelation zwischen der Zeit pro Frage und den Prozentwerten in den Prüfungen aus dem vorherigen ($r = -0,094$; $p = 0,267$) sowie jetzigen Semester im Modul M3.1 ($r = -0,049$; $p = 0,559$) nachgewiesen werden. Die Gesamtzeit korrelierte ebenso nicht signifikant mit den Prozentwerten der Prüfungen des vorherigen Semesters ($r = -0,007$; $p = 0,931$) sowie der Prüfung aus dem Modul M3.1 ($r = 0,096$; $p = 0,252$).

Zudem wurde ersichtlich, dass über die Hälfte der Studierenden (73,9 %) die Fälle zwischen 12:00 Uhr und 18:00 Uhr bearbeiteten.

3.2.3 #clue

In dem Lehrformat #clue wurden den Studierenden jeden Tag zwei originale IMPP-Prüfungsfragen gestellt, die die Studierenden mit Hilfe einer App beantworten können. 79,6 % der Studierenden nutzen dieses Format mindestens ein Mal. In der Gruppe der Studierenden, die mindestens eine Frage bearbeiteten, wurden durchschnittlich $29,0 \pm 13,3$ Fragen beantwortet, von denen durchschnittlich $70,5 \pm 13,4\%$ richtige Antworten waren. 54 Studierende (33,3 %) beantworteten alle Fragen, wohingegen 33 Studienteilnehmer (20,4 %) keine Frage beantworteten.

Aus den Daten wurde ersichtlich, dass die Frauen mehr Fragen beantworteten als ihre männlichen Kollegen (Frauen: $30,7 \pm 11,9$; Männer: $25,5 \pm 15,5$; $p = 0,037$; $d = 0,40$). Die männlichen Studierenden jedoch erreichten eine signifikant höhere Prozentzahl richtig beantworteter Fragen (männliche Studierende: $75,3 \pm 13,5$; weibliche Studierende: $68,3 \pm 12,9$; $p = 0,005$; $d = 0,54$). Zwischen den in #clue erreichten Prozentwerten und dem Alter konnte keine signifikante Korrelation nachgewiesen werden ($r = 0,026$; $p = 0,772$). In den folgenden Analysen zeigte sich weiterhin, dass die erreichten Prozentwerte in #clue signifikant mit den in vorherigen Klausuren erreichten Prozentwerten ($r = 0,294$; $p = 0,001$) und den Prozentwerten aus der Modulabschlussklausur ($r = 0,252$; $p = 0,004$) korrelierten. Zudem konnte eine signifikante Korrelation zwischen den erreichten Prozentwerten in #clue und den E-Fallseminaren nachgewiesen werden ($r = 0,441$; $p < 0,001$).

Durch das zeitlich vorher definierte Freischalten der #clue-Fragen sowie die erhobenen Daten der Bearbeitung, konnte die mittlere Verzögerung, mit der die Studierenden die #clue-Fragen lösten, bestimmt werden: Sie betrug $2,2 \pm 1,9$ Tage. Mit Blick auf die einzelnen

Wochen lassen sich in der mittleren Verzögerung keine starken Unterschiede feststellen (Woche 1: $2,1 \pm 2,5$ Tage; Woche 2: $2,3 \pm 2,2$ Tage; Woche 3: $2,1 \pm 2,2$ Tage; Woche 4: $1,9 \pm 2,2$ Tage).

Zudem konnte analysiert werden, dass 61,2 % der Studierenden mindestens 50 % der #clue-Fragen am Tag der Veröffentlichung oder einen Tag später bearbeitet haben. Sie zählten damit als *early-completer*. Unter den Männern, sowie auch unter den Frauen, überwog jeweils der Anteil der Studierenden, die zu den *early-completern* gehörten (in der Gruppe der Männer waren 68,3 % *early-completer*; in der Gruppe der Frauen waren 58 % *early-completer*). Der Unterschied war jedoch nicht signifikant ($p = 0,262$). Die *early-completer* lieferten weiterhin signifikant mehr richtige Antworten (in Prozent) auf die #clue-Fragen als diejenigen Studierenden, die 50 % der #clue-Fragen erst nach dem zweiten Tag nach Freischaltung bearbeiteten (*early-completer*: $73,6 \pm 14,6$ %; *non early-completer*: $65,6 \pm 9,6$ %; $p = 0,001$; $d = 0,63$). Im Mittel waren die *early-completer* nicht signifikant jünger als die *non early-completer* (*early-completer*: $24,7 \pm 3,2$; *non early-completer*: $24,8 \pm 3,0$; $p = 0,749$; $d = 0,06$). In Bezug auf die Abschlussprüfungen der einzelnen Module, ergaben sich folgende Ergebnisse: *early-completer* erreichten in den Prüfungen des vorherigen Semesters nicht signifikant mehr Prozentwerte als die *non early-completer* (*early-completer*: $82,6 \pm 7,5$ %; *non early-completer*: $82,1 \pm 7,2$ %; $p = 0,703$; $d = 0,07$). Auch im Sommersemester 2020 erreichten die *early-completer* nicht signifikant mehr Prozentwerte als die *non early-completer* (*early-completer*: $86,8 \pm 8,3$ %; *non early-completer*: $86,4 \pm 8,6$ %; $p = 0,801$; $d = 0,05$).

Die favorisierte Zeit, um die #clue Fragen zu lösen, lag zwischen 12:00 Uhr und 18:00 Uhr. Zu dieser Zeit bearbeiteten 48,5 % der Studierenden die Fragen.

3.3 Überprüfung des Wissenserwerbs und der Wissensanwendung

3.3.1 EMERGE

Das *serious game* EMERGE wurde im Sommersemester 2020 in den letzten drei Wochen des Semesters in einer direkten Kontaktlehre durchgeführt. Dabei nahmen 149 Personen (92 %) an diesem Lehrformat teil. Die interne Konsistenz dieses Lehrformats war akzeptabel mit einem Cronbachs alpha von 0,717.

Im Mittel erreichten die Teilnehmer einen Prozentwert von $47,6 \pm 14,0$ %. In dem *serious game* erreichten die weiblichen Studierenden einen etwas höheren Prozentwert als ihre männlichen Kollegen (Frauen: $48,2 \pm 13,0$ %; Männer: $46,5 \pm 16,0$ %). Es lag jedoch kein

signifikanter Unterschied vor ($p = 0,503$; $d = 0,11$). Des Weiteren konnte keine signifikante Korrelation zwischen dem Alter und den in EMERGE erreichten Prozentwerten festgestellt werden ($r = 0,063$; $p = 0,444$). Außerdem ergab sich keine signifikante Korrelation zwischen den in EMERGE erreichten Prozentwerten und den Prozentwerten aus den vorherigen Klausuren ($r = 0,017$; $p = 0,840$) sowie den Prozentwerten aus der Modulabschlussklausur ($r = -0,049$; $p = 0,557$). Auch zwischen den in #clue erreichten Prozentwerten und den Prozentwerten, die in EMERGE erzielt wurden, gab es keine signifikante Korrelation ($r = 0,172$; $p = 0,063$).

3.3.2 Studentische Leistungen der Abschlussprüfungen

Da die Prozentwerte der Prüfungen der Module M2.1, M2.2 und M2.4 aus dem vorherigen Semester (Wintersemester 2019/2020 oder Sommersemester 2019) vorlagen, konnte ein Mittelwert über die Prozentwerte der vorherigen Klausuren ermittelt werden. Dieser Wert lag bei $81,7 \pm 7,7 \%$.

In den Prüfungen des vorherigen Semesters erlangten die Studenten höhere Prozentwerte als ihre weiblichen Kommilitoninnen (Studenten: $82,9 \pm 7,7 \%$; Studentinnen: $81,2 \pm 7,6 \%$). Dieser Unterschied war jedoch nicht signifikant ($p = 0,183$; $d = 0,23$). Die Analyse der weiteren Daten zeigte, dass es eine signifikante negative Korrelation zwischen dem Alter und den Perzentilrängen der Prüfungen aus dem zweiten klinischen Semester gab ($r = -0,189$; $p = 0,018$).

Zudem wurde ersichtlich, dass die Perzentilränge der Prüfungen aus dem Wintersemester 2019/2020 (bzw. Sommersemester 2019) sowohl mit den Prozentwerten, die in #clue ($r = 0,278$; $p = 0,002$) als auch mit denjenigen, die in den E-Fallseminaren ($r = 0,238$; $p = 0,004$) erreicht wurden, positiv und signifikant korrelierten. Nicht nur die in den E-Fallseminaren erreichten Prozentwerte wiesen eine signifikante Korrelation zu den Perzentilrängen der vorherigen Prüfungsergebnisse auf. Auch die Verzögerung, mit der die Studierenden die E-Fallseminare bearbeiteten, zeigte eine signifikante, negative Korrelation zu den Perzentilrängen der vorherigen Prüfungen ($r = -0,174$; $p = 0,039$).

Der Lernerfolg der Studierenden im Sommersemester 2020 wurde mit Hilfe einer summativen Abschlussprüfung aus MC-Fragen erhoben. Die interne Konsistenz der Abschlussprüfung war akzeptabel mit einem Cronbachs alpha von 0,698. Im Mittel erreichten die Studierenden bei der Prüfung im Modul M3.1 einen Prozentwert von $86,4 \pm 8,4 \%$. Die männlichen Studierenden erzielten in dieser Abschlussprüfung signifikant höhere Prozentwerte als ihre weiblichen Kommilitoninnen. Während die Studenten im

Durchschnitt $89,2 \pm 7,3$ % erreichten, erlangten die Studentinnen im Mittel nur $85,1 \pm 8,6$ % der möglichen Punkte ($p = 0,004$; $d = 0,46$). Zwischen dem Alter und dem Perzentilrang der jeweiligen Studierenden ließ sich keine signifikante Korrelation finden ($r = -0,120$; $p = 0,129$).

3.4 Analyse der Leistungsverläufe

Zur Überprüfung der Leistungsverläufe der Studierenden wurden die Ergebnisse der Prüfungen aus dem zweiten klinischen Semester mit den Prüfungsergebnissen der Abschlussklausur des Moduls M3.1 verglichen (siehe Kapitel 2.3 und Kapitel 2.4). In der anschließenden Analyse konnte eine signifikante Korrelation zwischen den Leistungen in den Prüfungen des vorherigen Semesters und den Leistungen in der Prüfung des Moduls M3.1 nachgewiesen werden ($r = 0,524$, $p < 0,001$).

Im Hinblick auf die beiden Geschlechter ließ sich erkennen, dass Studentinnen in den Prüfungen des zweiten klinischen Semesters einen nicht-signifikant niedrigeren Perzentilrang erzielten als ihre männlichen Kommilitonen (Studentinnen: $50,3 \pm 27,8$; Studenten: $57,3 \pm 28,5$; $p = 0,148$; $d = 0,25$). Diese Tendenz ergab sich auch in der Prüfung des Sommersemesters 2020 im Modul M3.1. Jedoch zeigte sich im Sommersemester 2020 ein signifikanter und deutlicherer Unterschied. Die Frauen erzielten im Mittel einen Perzentilrang von $46,0 \pm 28,3$, während die Männer einen Perzentilrang von $60,7 \pm 28,2$ erreichten ($p = 0,003$; $d = 0,52$).

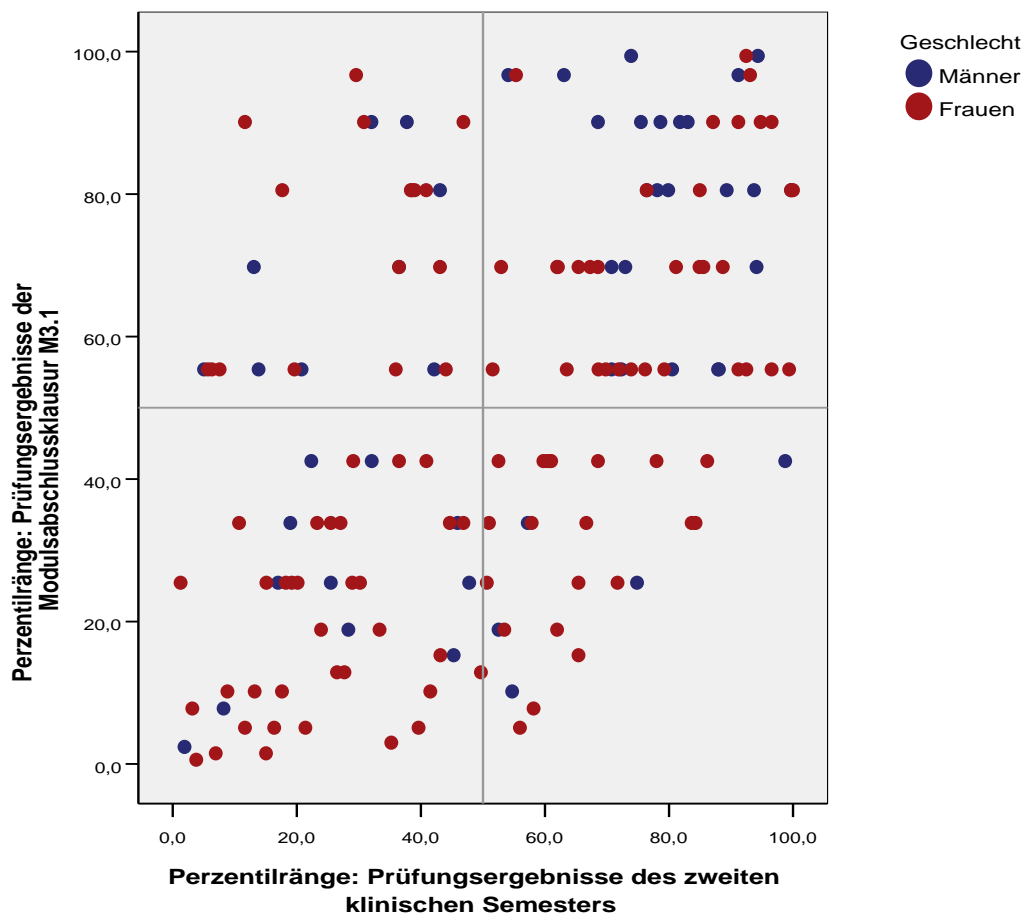


Abbildung 4: Streudiagramm Perzentilränge der Prüfungsergebnisse.

Abgebildet sind die Perzentilränge der Prüfungsergebnisse der Modulabschlussklausur M3.1 im Sommersemester 2020 auf der Y-Achse; auf der X-Achse sind die Perzentilränge der Prüfungsergebnisse des zweiten klinischen Semesters (im Wintersemester 2019/2020) abgebildet. Männer sind dabei als blaue Punkte, Frauen als rote Punkte dargestellt. Gruppe I befindet sich im Quadranten unten links, Gruppe II unten rechts, Gruppe III oben links und Gruppe IV oben rechts.

Durch die Umwandlung der Prozentwerte (der Abschlussklausur vom Modul M3.1 und der Klausuren aus dem vorherigen Semester) in Perzentilen konnten die Leistungen der Studierenden aus dem zweiten und dritten Semester miteinander verglichen werden (siehe Abbildung 4). Dazu wurden die Studierenden in vier Gruppen eingeteilt: I. Leistungsschwach, keine Veränderung; II. Verschlechterung; III. Verbesserung; IV. Leistungsstark, keine Veränderung. In der Gruppe I befanden sich 46 Studierende (29,7 %), in der Gruppe IV waren 57 Studierende (36,8 %). Die Leistungen der Studierenden aus der Gruppe I und IV veränderten sich nicht; sie waren gleich schwach oder gleich stark. Leistungsveränderungen zeigten sich vor allem in den Gruppen II ($n = 26$; 18,6 %) und III ($n = 26$; 18,6 %). Während die Leistungen der Studierenden aus der Gruppe II abnahm, stieg diese in Gruppe III.

Allerdings konnte kein signifikanter Unterschied zwischen den Gruppe II und III in Bezug auf den Anteil der Studentinnen (in Prozent) nachgewiesen werden (Gruppe II: 80,8 %; Gruppe III: 65,4 %; $p = 0,221$).

3.5 Analyse der Nutzungscharakteristika als Einflussfaktoren auf den Lernerfolg

In der folgenden Analyse wurden univariate und multivariate Regressionsanalysen durchgeführt, um folgende Hypothesen zu testen: (1) Eine intensive Nutzung der Lehrformate, die auf eine Förderung des Wissenserwerbs und -bewahrung abzielen, sind mit höheren Leistungen in der summativen MC-Prüfung assoziiert. Ebenso (2) trägt die Förderung der Wissensanwendung zu höheren Prozentwerten in dem Spiel EMERGE bei. Ziel war es, Einflussfaktoren auf den Lernerfolg zu identifizieren (siehe Kapitel 2.2.3 und Kapitel 2.4).

3.5.1 Regressionsanalyse mit den Prüfungsleistungen als abhängige Variable

Die Auswertung der Daten in der univariaten linearen Regressionsanalyse zeigte folgende Ergebnisse (siehe Tabelle 1):

Mit der Prüfungsleistung der Studierenden als abhängige Variable konnte eine signifikante Assoziation mit den Perzentilrängen der vorherigen Prüfungsleistungen des zweiten klinischen Semesters nachgewiesen werden. Ebenso waren in der univariaten Regressionsanalyse das männliche Geschlecht, hohe Prozentwerte in #clue und in den E-Fallseminare positive Einflussfaktoren auf die Prüfungsleistung der Studierenden am Ende des dritten klinischen Semesters. Neben diesen Einflussfaktoren konnte auch eine Assoziation zwischen der Prüfungsleistung und der Verzögerung der E-Fallseminare erkannt werden. Eine geringe Verzögerung war ein Einflussfaktor für eine stärkere Leistung.

Nach Durchführung der univariaten linearen Regressionsanalyse und Erkennen der signifikanten Assoziationen, wurde eine multivariate lineare Regressionsanalyse durchgeführt. In der multivariaten Regressionsanalyse ließ sich feststellen, dass nur die vorherige Prüfungsleistung (Perzentilrang der Prüfungen des zweiten klinischen Semesters) sowie das männliche Geschlecht weiterhin signifikante Einflussfaktoren auf die Leistung in der Modulabschlussklausur waren (siehe Tabelle 2).

Tabelle 1: Einflussfaktoren auf die Prüfungsleistung der Studierenden im Sommersemester 2020 in einer univariaten linearen Regressionsanalyse.

Abhängige Variable: Prüfungsleistungen im Sommersemester 2020 im Modul M3.1 (in Prozent).

Unabhängige Variabel	Beta (Koeffizient)	95% KI - untere Grenze	95% KI - obere Grenze	p-Wert
Weibliches Geschlecht	-4,05	-6,81	-1,29	0,004
Alter	-0,30	-0,68	0,07	0,112
Prüfungen zweites KS (Perzentilrang)	0,15	0,11	0,20	< 0,001
Fallbasierte Aufgaben: mittlere Verzögerung	-0,16	-0,63	0,31	0,490
E-Fallseminar: Anzahl bearbeiteter Seminare	0,77	-0,02	1,56	0,057
E-Fallseminar: erreichte Prozentpunkte	0,21	0,09	0,33	0,001
E-Fallseminar: mittlere Verzögerung	-1,05	-1,95	-0,14	0,024
E-Fallseminar: Zeit pro Frage	-0,01	-0,05	0,03	0,559
E-Fallseminar: Gesamtzeit	0,00	0,00	0,00	0,252
#clue: Anzahl bearbeiteter Fragen	0,03	-0,05	0,11	0,468
#clue: Anzahl richtiger Fragen	0,08	-0,03	0,18	0,141
#clue: erreichte Prozentpunkte	0,16	0,05	0,27	0,004
#clue: mittlere Verzögerung	0,08	-0,70	0,86	0,849

(KI = Konfidenzintervall; KS = Klinisches Semester)

Tabelle 2: Multivariate lineare Regressionsanalyse der Einflussfaktoren auf die studentische Prüfungsleistung im Sommersemester 2020.

Abhängige Variable: Prüfungsleistungen im Sommersemester 2020 im Modul M3.1 (in Prozent).

Unabhängige Variabel	Beta (Koeffizient)	95% KI - untere Grenze	95% KI – obere Grenze	p-Wert
Weibliches Geschlecht	-3,10	-6,14	-0,05	0,037
Prüfungen zweites KS (Perzentilrang)	0,13	0,08	0,18	< 0,001
E-Fallseminare: erreichte Prozentpunkte	0,07	-0,06	0,22	0,456
E-Fallseminare: mittlere Verzögerung	-0,55	-1,46	0,37	0,900
#clue: erreichte Prozentpunkte	-0,02	-0,13	0,10	0,637

(KI = Konfidenzintervall; KS = Klinisches Semester)

3.5.2 Regressionsanalyse mit den EMERGE-Leistungen als abhängige Variable

In der anschließenden univariaten Regressionsanalyse zur Identifikation von Einflussfaktoren auf die Leistung in der Spielsitzung, wurde als abhängige Variable die Leistung der Studierenden in EMERGE (in Prozent) festgelegt. Die Analyse zeigte, dass nur ein hoher Prozentwert in den E-Fallseminaren als signifikanter, prädiktiver Faktor für eine starke Leistung in der Spielsitzung EMERGE (siehe Tabelle 3) gilt.

Tabelle 3: Einflussfaktoren auf die Leistung der Studierenden im computergestützten Spiel EMERGE in einer univariaten linearen Regressionsanalyse.

Abhängige Variable: Leistungen der Studierenden in EMERGE (in Prozent).

Unabhängige Variabel	Beta (Koeffizient)	95% KI- untere Grenze	95% KI – obere Grenze	p-Wert
Weibliches Geschlecht	1,65	-3,20	6,50	0,503
Alter	0,25	-0,40	0,91	0,444
Prüfungen zweites KS (Perzentil)	0,01	-0,08	0,09	0,894
Fallbasierte Aufgaben: mittlere Verzögerung	-0,58	-1,39	0,24	0,165
E-Fallseminar: Anzahl bearbeiteter Seminare	-0,60	-1,99	0,81	0,399
E-Fallseminar: erreichte Prozentpunkte	0,30	0,07	0,53	0,010
E-Fallseminar: mittlere Verzögerung	0,90	-0,74	2,55	0,279
E-Fallseminar: Zeit pro Frage	0,05	-0,02	0,12	0,130
E-Fallseminar: Gesamtzeit	0,00	-0,00	0,00	0,674
#clue: Anzahl bearbeiteter Fragen	0,01	-0,13	0,15	0,932
#clue: Anzahl richtiger Fragen	0,04	-0,14	0,22	0,665
#clue: erreichte Prozentpunkte	0,18	-0,01	0,37	0,063
#clue: mittlere Verzögerung	-0,83	-2,23	0,58	0,246

(KI = Konfidenzintervall; KS = Klinisches Semester)

3.6 Prädiktoren für negative Leistungsveränderungen

Ein weiteres Ziel der Studie war es, prädiktive Einflussfaktoren zu identifizieren, die mit einer ungünstigen Prüfungsleistung einhergehen. Dazu wurden univariate und multivariate Regressionsanalysen durchgeführt mit der Prüfungsleistung als abhängige Variable. Wie in Kapitel 1.6 und Kapitel 2.4 beschrieben, wurde erwartet, aus den Nutzungscharakteristika der ersten fallbasierten Aufgabe, der ersten zehn #clue-Fragen sowie des ersten E-Fallseminars Prädiktoren erkennen zu können. In der Analyse galten Nutzungscharakteristika daher als unabhängige Variablen.

Die Ergebnisse der univariaten Regressionsanalyse (siehe Tabelle 4) zeigten, dass es eine signifikante Assoziation zwischen den Prozentwerten der Abschlussprüfung und den Prozentwerten in den E-Fallseminaren sowie #clue gab. Je mehr Prozente in den ersten zehn #clue-Fragen sowie dem ersten E-Fallseminar erreicht wurden, desto höher war die Leistung in der summativen Abschlussprüfung. Ebenso konnte erkannt werden, dass eine geringere Verzögerung in der Bearbeitung des ersten E-Fallseminars assoziiert ist mit besseren Prüfungsleistungen. Die Verzögerungen der ersten fallbasierten Aufgabe sowie der ersten zehn #clue-Fragen standen jedoch in keinem Zusammenhang zu den Prozentwerten der Abschlussprüfung. Weiterhin konnte die Bearbeitung des ersten E-Fallseminars, die Zeit, die die Studierenden für das erste E-Fallseminar aufwandten, sowie die Anzahl der beantworteten #clue-Fragen in der ersten Woche nicht als Prädiktoren nachgewiesen werden.

In der anschließenden multivariaten linearen Regressionsanalyse blieb nur eine signifikante Assoziation zwischen den Prozentwerten des ersten E-Fallseminars sowie der Prüfungsleistung bestehen. Höhere Prozentwerte in dem ersten E-Fallseminar können eine günstigere Prüfungsleistung prognostizieren. Zu den Prozentwerten der #clue-Fragen sowie der Verzögerung des ersten E-Fallseminars ließ sich weiterhin kein Zusammenhang feststellen (siehe Tabelle 5).

Tabelle 4: Prädiktoren für die studentische Prüfungsleistung in der ersten Modulwoche: eine univariate lineare Regressionsanalyse.

Abhängige Variable: Prüfungsleistungen im Sommersemester 2020 im Modul M3.1 (in Prozent).

Unabhängige Variable	Beta (Koeffizient)	95% KI- untere Grenze	95% KI- obere Grenze	p-Wert
Erste fallbasierte Aufgabe: mittlere Verzögerung	-0,04	-0,27	0,19	0,751
Erstes E-Fallseminar: Bearbeitung	-0,17	-3,51	3,18	0,922
Erstes E-Fallseminar: erreichte Prozentpunkte	0,22	0,11	0,33	< 0,001
Erstes E-Fallseminar: mittlere Verzögerung	-0,91	-1,54	0,29	0,005
Erstes E-Fallseminar: Gesamtzeit für alle 4 Fälle	0,00	-0,00	0,00	0,635
#clue: Anzahl bearbeiteter Fragen 1-10	-0,01	-0,31	0,31	0,968
#clue: erreichte Pro- zentpunkte Fragen 1-10	0,10	0,01	0,18	0,022
#clue: mittlere Verzögerung Fragen 1-10	0,21	-0,39	0,80	0,492

(KI = Konfidenzintervall)

Tabelle 5: Multivariate lineare Regressionsanalyse der Prädiktoren der studentischen Prüfungsleistung in der ersten Modulwoche.

Abhängige Variable: Prüfungsleistungen im Sommersemester 2020 im Modul M3.1 (in Prozent).

Unabhängige Variable	Beta (Koeffizient)	95% KI- untere Grenze	95% KI- obere Grenze	p-Wert
Erstes E-Fallseminare: erreichte Prozentpunkte	0,159	0,017	0,295	0,028
Erstes E-Fallseminar: mittlere Verzögerung	-0,670	-1,371	0,031	0,061
#clue: erreichte Prozentpunkte Fragen 1-10	0,062	-0,030	0,153	0,183

(KI = Konfidenzintervall)

4 Diskussion

4.1 Darlegung der Ergebnisse und Studienlage

In der vorliegenden prospektiven Studie aus dem Sommersemester 2020 konnte eine unterschiedlich ausgeprägte Nutzung der einzelnen digitalen Ressourcen beobachtet werden. Dabei konnte ein Zusammenhang zwischen den fragenbasierten *tools*, #clue und E-Fallseminaren, und der Abschlussprüfung des Moduls M3.1 erfasst werden. Es scheint, als seien Studierende, die in #clue sowie den E-Fallseminaren gut abschneiden auch diejenigen, die in der Abschlussprüfung eine gute oder gar sehr gute Leistung zeigen. Die Bearbeitung der fallbasierten Aufgaben hingegen hatte keinen Einfluss auf die Leistungen der Studierenden in der Modulabschlussklausur. Darüber hinaus konnte erkannt werden, dass EMERGE in keiner Assoziation zu den fragenbasierten *tools* oder der Abschlussprüfung stand. Diese Tatsache legt nahe, dass hier unterschiedliche Konstrukte gemessen wurden.

In der weiteren Analyse des adjustierten Modells wurde erkannt, dass einzig das männliche Geschlecht sowie die Prüfungsleistungen aus dem vorherigen Semester als Prädiktoren für eine starke Leistung in der Abschlussprüfung übrigbleiben. Es ergaben sich zwei wesentliche Ergebnisse aus dieser Studie:

Zum einen wurde gezeigt, dass Studentinnen der Humanmedizin durch die Umstellung von der Präsenzlehre auf eine rein digitale Lehre gegenüber ihren männlichen Kommilitonen benachteiligt sind. Zum anderen konnten Prädiktoren erkannt werden, die eine Vorhersage über die Leistung der Studierenden treffen. Dadurch ist es möglich im weiteren Verlauf Studierende zu identifizieren, die womöglich eine weniger gute Prüfungsleistung erbringen. Solchen Studierenden könnte eine gezielte Unterstützung und Förderung angeboten werden.

Durch die Corona- Pandemie musste die universitäre Lehre umstrukturiert werden. Der seit Jahrzehnten etablierte Präsenzunterricht musste aufgrund der strikten Kontaktbeschränkungen in ein digitales Format überführt werden. Auch andere universitären Kliniken boten erstmals ein reines Online-Curriculum an (Bączek et al. 2021; Dost et al. 2020; Redinger et al. 2020). Im Zuge dessen wurden neue Unterrichtsformate entwickelt und erstmals an den Studierenden getestet. Wesentliche Formate sind nach dem *flipped-classroom-model* umstrukturiert (Röhle et al. 2021; van der Keylen et al. 2020) und *tools*, wie Quiz basierend auf Patientenfällen (Huber et al. 2021), computergestützte Simulationen im klinischen Setting (De Ponti et al. 2020) sowie Tests mit Multiple-Choice-Fragen (Darici et al. 2021) angeboten worden. Bestandteile der Lehre zur Zeit der Pandemie waren zudem

digitale Gruppendiskussionen (Khalil et al. 2020; Knie et al. 2020) und synchrone Online-Konferenzen (Huber et al. 2021).

Viele Studien befragten die Studierenden während der Pandemie nach ihrer Zufriedenheit und ihrem subjektiven Lernerfolg (Khalil et al. 2020; Puljak et al. 2020; Ruiz-Grao et al. 2022). Eine Gruppe in Kalifornien untersuchte die Vor- und Nachteile des digitalen Lehrplans (Shahrivini et al. 2021). Neben der Flexibilität gaben die Studierenden als weiteren Vorteil die selbständige Anpassung des Lernpensums an den individuellen Lernfortschritt an. Allerdings teilten sie auch mit, dass die Unterrichtsqualität erheblich gesunken sei (Shahrivini et al. 2021). Studierende aus Mexiko gaben an, dass ihnen besonders die klinische Erfahrung fehlt (Servin-Rojas et al. 2022). Wie Olmes und Kollegen erklärten, seien digitale Möglichkeiten im Studium ein Zugewinn, ersetzen jedoch keine *bedside*-Lehre (Olmes et al. 2021).

Die vorhandenen Studien sind momentan jedoch qualitativ ausgelegt und bilden meist nur die subjektive Zufriedenheit und die eigene Wahrnehmung der Studierenden ab. In der aktuellen Literatur findet sich keine Studie, die die Nutzung der digitalen Ressourcen beschreibt und ihre Auswirkungen auf den objektiv messbaren Lernerfolg untersucht. Ebenso ist es durch die Pandemie und die sich daraus ergebenden Kontaktbeschränkungen wesentlich schwieriger geworden, gefährdete Studierende zu erkennen. Eine Identifizierung Derselben könnte helfen, durch ein vorzeitiges Unterstützen, Unzufriedenheit und Studienabbrüche zu reduzieren.

4.2 Beurteilung der Nutzungscharakteristika der digitalen Ressourcen

4.2.1 Hohe Leistungen in den fragenbasierten *tools*

In der Analyse der Studiendaten wurde ersichtlich, dass es in #clue eine signifikante Assoziation zwischen den erreichten Prozentwerten in #clue und der Abschlussprüfung des Moduls M3.1 gab. Hieraus wird geschlossen, dass Studierende, die sich mit #clue beschäftigten, auch in der Abschlussklausur besser abschnitten. Diese Beobachtung entspricht den vorherigen Erwartungen. Zum einen wurde angenommen, dass #clue und die Abschlussprüfung des Moduls M3.1 jeweils auf das Bildungsziel Wissenserwerb ausgerichtet sind. Während #clue zum Wissenserwerb beitragen sollte, prüfte die summative MC-Prüfung den erhaltenen Wissenserwerb ab.

Zum anderen ließ sich in der vorliegenden Untersuchung das Phänomen des *testing effect's* beobachten. Nach dem *testing effect*, der bereits in diversen Studien untersucht wurde (Larsen et al. 2013a; McConnell et al. 2015; Roediger und Karpicke 2006), führt wiederholtes Testen zu einer höheren Merkfähigkeit. Dabei ist das erlernte Wissen meist nicht nur auf einen Kontext beschränkt, sondern auf eine Vielzahl anderer Situationen, in diesem Fall auf verschiedene Krankheitsbilder, anwendbar (Butler 2010; Carpenter 2009; Roediger und Karpicke 2006). In den Untersuchungen zur vorliegenden Studie zeigte sich, dass vorheriges und wiederholtes Testen durch #clue zu bessern Leistungen sowohl in #clue als auch in der summativen MC-Prüfung führte, die auf eine höhere Merkfähigkeit zurückzuführen sind. Mit der Verbesserung der Merkfähigkeit wird eine höherer Lernqualität und -effektivität erreicht. Dazu bestätigten die Untersuchungen von Augustin sowie Roediger und Karpicke zusätzlich, dass Studierende, die eine Lehrintervention erhalten haben, besser in Überprüfungstests abschneiden als die Kontrollgruppe, die keine Intervention erhielt (Augustin 2014; Roediger und Karpicke 2006).

Ein derartiger Zusammenhang zwischen Lernquantität und Lerneffektivität, wie er in #clue zu sehen war, konnte ebenso bei den E-Fallseminaren erkannt werden. Studierende, die viele elektronische Fälle bearbeiteten, erreichten hierin mehr Punkte. Höhere Prozentwerte sowie eine geringe Verzögerung bei der Bearbeitung der E-Fallseminare waren assoziiert mit höheren Prozentwerten in der Modulabschlussklausur. Da es sich bei den *key-feature*-Fragen, die in den E-Fallseminaren zu beantworten waren, und den MC-Fragen aus der Prüfung um unterschiedliche Frage-Typen handelt, besteht die Annahme, dass leistungsstarke Studierende mit beiden Frage-Typen problemlos zurechtkommen.

Bereits 2016 führten Raupach und Kollegen eine Studie zur Untersuchung des Lerneffekts durch computergestützte Aufgaben durch (Raupach et al. 2016). Die Studierenden wurden dazu während der Seminare in zwei Gruppen aufgeteilt. In der Interventionsgruppe wurden computergestützt *key-feature*-Fragen zu einzelnen Patientenfällen gestellt. Die Kontrollgruppe lernte ohne diese Fragen. Die Studierenden erzielten 2016 in den *key-feature*-Fragen im Mittel $56,0 \pm 25,8$ % und damit geringfügig weniger Prozentpunkte als die Studierenden in dieser Studie 2020 ($67,5 \pm 11,0$ %). In der Abschlussprüfung, 13 Wochen nach der Intervention, konnte gezeigt werden, dass Studierende in der Interventionsgruppe signifikant besser in der Abschlussprüfung abschnitten als die Studierenden in der Kontrollgruppe (Interventionsgruppe: $59,3 \pm 25,8$ %; Kontrollgruppe: $46,7 \pm 24,8$ %). Dieses Ergebnis ähnelt den Beobachtungen der vorliegenden Studie. Die wiederholte Teilnahme an den E-

Fallseminaren (mit *key-feature*-Fragen) und das Erreichen höherer Prozentwerte waren assoziiert mit höheren Leistungen in der Abschlussprüfung.

Ähnliche Resultate erzielten auch Fransen und Kollegen (Fransen et al. 2018). In der randomisierten Studie an der Universität Maastricht erhielten 32 Medizinstudenten im vierten Jahr ihrer Ausbildung neben einer konventionellen Lehre im Fachbereich Dermatologie zusätzlich eine *E-Learning* Intervention. Bestandteil der Intervention waren patientenzentrierte Fälle und zugehörige Fragen. Die Kontrollgruppe aus 30 Studierenden erhielt nur eine konventionelle Lehre. In abschließenden Tests konnte gezeigt werden, dass Studierende von dem zusätzlichen *E-Learning* Angebot profitierten und in den Tests signifikant besser abschnitten (*E-Learning*-Gruppe: $73,1 \pm 5,1$; Kontrollgruppe: $51,4 \pm 6,4$ %). Der Einbezug dieses Formats in die klinische Lehre habe den Wissenserwerb signifikant verbessert und Wissensdefizite reduziert. Auch in der vorliegenden Studie aus dem Sommersemester 2020 haben Studierende von dem Einsatz der digitalen Lehrformate (E-Fallseminare und #clue) profitiert, wie an der besseren Prüfungsleistung der Studierenden zu erkennen ist, die die Formate regelmäßig nutzten.

Larsen und Kollegen zeigten gleichartige Ergebnisse in ihrer randomisierten Crossover-Studie an Assistenzärzten*innen der Neurologie (Larsen et al. 2009). In der Studie erhielten die Ärzte*innen eine Schulung zu den Themen *Myasthenia gravis* und *Status epilepticus*. Anschließend nahmen sie an drei Lernsitzungen oder Testungen (mittels *short-answer-questions*) teil. In einem Abschlusstest sechs Monate später wurden die Ärzte*innen zu den Themen aus den Lehrinterventionen befragt. Dabei konnte gezeigt werden, dass wiederholtes Testen zu besseren Lernergebnissen führt als wiederholtes Studieren (wiederholtes Testen: 39 %; wiederholtes Studieren: 26 %). Zahlreiche folgende Studien belegten im Laufe der Jahre, dass der Lernerfolg durch wiederholtes Testen höher ist als durch wiederholtes Studieren (Larsen et al. 2013a; Larsen et al. 2013b). In den Ergebnissen der vorliegenden Untersuchung zeigt sich ein gleichartiger Trend. Studierende, die mehr E-Fallseminare lösten, erzielten mehr Punkte und zeigten auch in der Abschlussprüfung eine höhere Leistung. Obwohl es im Vergleich zu Studienkohorten vorheriger Jahrgänge keine Kontrollgruppe gab, aufgrund der besonderen Situation, stimmen die Ergebnisse mit Resultaten vorheriger Studien überein. Dieser Trend spiegelte sich ebenso in der Auswertung der #clue bezogenen Daten wider. Häufigeres Testen führte zu besseren Ergebnissen und geht mit einem höheren Lernerfolg einher.

4.2.2 Geschlechterunterschiede

Mit spezifischem Blick auf die Geschlechter zeigte die Analyse der Studiendaten wesentliche Unterschiede. In den E-Fallseminaren beantworteten die Frauen signifikant mehr Fragen. Trotz der Tatsache, dass die Männer weniger Fragen beantworteten als die Frauen, erreichten sie signifikant mehr Punkte in den *key-feature*-Fragen. Dieselbe Konstellation ließ sich bei den *#clue*-Fragen finden. Auch hier beantworteten die Frauen signifikant mehr Fragen, während die Männer signifikant mehr Punkte erreichten. Zudem konnte erfasst werden, dass Männer in der Modulabschlussprüfung im Sommersemester 2020 mit höheren Prozentwerten abschlossen als die Frauen. Aufgrund dieser Ergebnisse scheint es, als seien Frauen trotz eines erhöhten Zeitaufwands im Sommersemester 2020 benachteiligt gewesen.

Bereits in früheren Forschungen gab es Hinweise darauf, dass Geschlechtsunterschiede bezüglich des Lernverhaltens im Studium vorherrschen. In dem systematischen Review von Ali und Kollegen wurden Studien untersucht, die die klinisch chirurgischen Fähigkeiten sowie deren Erwerb von Medizinstudierenden und Assistenzärzten*innen beobachteten (Ali et al. 2015). Die Ergebnisse zeigten, dass männliche Medizinstudenten tendenziell leistungstärker sind als ihre weiblichen Kommilitoninnen. Weibliche Studierende sind anfänglich auf einem niedrigeren Niveau anzusiedeln als ihre männlichen Kollegen, dennoch haben sie sich durch Einzeltraining und Feedback maßgeblich verbessert (Ali et al. 2015). Auch Oussi und Kollegen wiesen darauf hin, dass eine technische Überlegenheit bei Männern zu besseren Ergebnissen führen kann (Oussi et al. 2021). Durch repetitives Training ist es jedoch möglich, diesen Effekt zu kompensieren. Wie die erläuterten Studien zeigen, konnten bereits in der Vergangenheit Unterschiede in den Fähigkeiten erkannt werden. Die Tatsache, dass Studentinnen in der vorliegenden Studie trotz eines erhöhten Aufwandes schlechter abschnitten als die männlichen Kommilitonen, könnte durch eine technische Überlegenheit der Männer erklärt werden. Durch gezieltes Training könnten solche Effekte möglicherweise hingegen vermieden werden.

Ansätze für Erklärungen auf biologischer Ebene lieferte die Arbeitsgruppe um Kromann und Kollegen. In ihrer Forschung wurde der Cortisolspiegel bei Männern und Frauen während eines Seminars zur Herz-Lungen-Wiederbelebung untersucht (Kromann et al. 2011). Ein Resultat der Studie war, dass der Lernerfolg bei Männern positiv mit dem Cortisolspiegel korrelierte, wohingegen dies bei Frauen nicht der Fall war. Glukokortikoide wie Cortisol haben eine modulierende Wirkung auf das Gehirn und sind maßgeblich an Lern- und Gedächtnisprozessen beteiligt (Lupien et al. 2007). Auch dieser biologische Ansatz

könnte die Unterschiede zwischen Männern und Frauen erklären, die in der vorliegenden Untersuchung gefunden wurden (weitere Erklärungsansätze siehe in Kapitel 4.3).

4.2.3 Altersunterschiede

Zu Anfang der Studie wurde vermutet, dass sich die Altersunterschiede in den Leistungen widerspiegeln. Im Zuge dessen wurde angenommen, dass ältere Studierende höhere Leistungen in den fragenbasierten *tools* sowie der Abschlussprüfung erbringen. Wie bereits Simmenroth-Nayda und Kollegen darauf hinwiesen, schneiden Studienplatzanwärter mit Berufserfahrung besser in den Prüfungen ab (Simmenroth-Nayda et al. 2014). Aufgrund einer vorherigen Ausbildung bzw. der vorherigen Ausübung eines Berufs liegt es nahe, dass Studierende mit Berufserfahrung älter sind, als Studierende, die direkt nach dem Abitur mit dem Studium beginnen. Ebenso scheinen sich Studierende unterschiedlicher Altersgruppen in unterschiedlichen Lebensumständen zu befinden. Während jüngere Studierende in den Anfängen ihres beruflichen Werdegangs stehen, befinden sich ältere Studierende zumeist schon in der Familienplanung. Wie Al-Balas und Kollegen berichteten, profitieren Studierende von der hohen Flexibilität der Online-Lehre in Pandemiezeiten (Al-Balas et al. 2020). Durch die aufgezeichneten Seminare und Vorlesungspodcasts wurde ein höherer Grad an eigenständiger Zeitplanung erlangt. Für Studierende mit Kind oder erwerbstätige Studierende könnte sich damit die Möglichkeit ergeben, an mehr Lehrveranstaltungen im Rahmen der Online-Lehre teilzunehmen als es zu Zeiten vor der Pandemie in Präsenz möglich war.

Auf Basis der Annahmen, dass ältere Studierende mit mehr Berufserfahrung besser abschneiden (Simmenroth-Nayda et al. 2014) und die Flexibilität der Online-Veranstaltungen wesentlich erhöht wurde (Al-Balas et al. 2020), könnte es nahe liegen, dass ältere Studierende im ersten Pandemie-Semester besser abschneiden als ihre jüngeren Kommilitonen.

In der hier präsenten Analyse zeigte sich jedoch kein signifikanter Zusammenhang zwischen der Leistung der Studierenden und dem Alter. Weder in #clue noch den E-Fallseminaren oder der Modulabschlussklausur im Sommersemester 2020 konnte eine Korrelation nachgewiesen werden. Damit ließ sich die initiale Annahme, dass ältere Studierende im Corona-bedingten Online-Semester besser abschnitten, nicht bestätigen.

4.2.4 *Early-completer* und *non early-completer*

Um das weitere Studierendenverhalten in Bezug auf die Nutzung der fragenbasierten *tools* #clue und der E-Fallseminare zu untersuchen, wurde die Kohorte in *early-completer* und *non*

early-completer eingeteilt (für Details siehe Kapitel 2.2.3 und 3.2.2). Studierende, die als *early-completer* in den #clue Fragen galten, erlangten signifikant mehr Punkte, als Studierende, die die Bearbeitung der #clue-Fragen länger hinauszögerten. Der gleiche Trend spiegelte sich auch bei der Bearbeitung der E-Fallseminare wider. Studierende, die frühzeitig die E-Fallseminare bearbeiteten und damit zu den *early-completern* zählten, erzielten signifikant mehr Punkte in den E-Fallseminaren, als diejenigen Studierenden, die zu den *non early-completern* zählten. Ebenso zeigten die *early-completer* der E-Fallseminare signifikant bessere Klausurergebnisse in der Modulabschlussprüfung.

Die Themen der E-Fallseminare waren auf den Vorlesungen der vorangegangenen Woche abgestimmt. Zwischen dem Hören der Vorlesungen und dem Testen des aufgenommenen Wissens durch die E-Fallseminare lagen dementsprechend nur wenige Tage. Daraus resultiert möglicherweise, dass bei eifrigen Studierenden, die die E-Fallseminare kurz nach der Freischaltung lösten, das Wissen aus den Vorlesungen im Gedächtnis noch präsenter war und sie daher in den Tests mehr Punkte erzielten.

Ähnliches spiegelt sich in den Beobachtungen zur Abschlussklausur wider: Studierende, die die E-Fallseminare frühzeitig bearbeiteten schnitten in der Abschlussprüfung signifikant besser ab. Wie bereits Halamish und Bjork zeigten, führt anfängliches und wiederholtes Testen zu einer besseren Merkfähigkeit des Gelernten (Halamish und Bjork 2011). Diese Ergebnisse stimmen mit denen der hier vorliegenden Studie überein (Halamish und Bjork 2011). Das anfängliche, wiederholte Testen durch die E-Fallseminare könnte mit einer erhöhten Merkfähigkeit assoziiert sein, gemessen an den Leistungen der Abschlussprüfung und den erreichten Prozenten in dem Lehrformat.

Ein weiterer Erklärungsansatz für die beobachtete Konstellation, bietet auch die Betrachtung der verschiedenen Typen von Studierenden. Studierende, die die E-Fallseminare besonders schnell bearbeiteten, könnten möglicherweise überaus eifrige und ehrgeizige Studierende gewesen sein. Diese arbeiteten die Lehrinterventionen gegebenenfalls schneller ab und lernten eifriger.

4.2.5 Bearbeitungszeit der E-Fallseminare

Zusätzlich konnte in der Auswertung der Studiendaten erkannt werden, dass Studierende mit einer längeren Bearbeitungszeit der E-Fallseminare signifikant mehr Punkte in den E-Fallseminaren erreichten.

Ein höherer Zeitaufwand bedingt in diesem Fall ein besseres Lernergebnis. Einen möglichen Erklärungsansatz bietet das *time-on-task*-Modell. Dem Modell nach zu urteilen, hat eine optimierte und effektive Lernzeit einen maßgeblichen Einfluss auf den Lernerfolg (Carroll 1989; Gettinger 1984; Scholkmann et al. 2017). Bereits seit Jahrzehnten ist diese Modell bekannt und ist auch auf die hier vorliegende Studie anwendbar: Studierende, die mehr Zeit für die Bearbeitung der Fälle benötigten und sich intensiver mit den Fragen auseinandersetzten, hatten einen höheren Lernerfolg, gemessen an den erreichten Punkten in den E-Fallseminaren. Gleichzeitig ist zu vermuten, dass eine längere Bearbeitungszeit auf ein ausführlicheres Auseinandersetzen mit den Krankheitsbildern und Nachschlagen von fehlendem Wissen zusammenhängt. Vertiefendes Lernen wirkt sich damit positiv auf die Gesamtleistung aus. Womöglich kann auch hier wieder der eifrige Typ Student erkannt werden.

Da die Daten zur benötigten Bearbeitungszeit der #clue-Fragen nicht erhoben wurden, konnten derartige Zusammenhänge bei dem fragenbasierten *tool* #clue nicht berechnet werden.

4.2.6 Abschließende Beurteilung der Nutzungscharakteristika

Die nachgewiesene Korrelation zwischen den einzelnen Lehrformaten (#clue, E-Fallseminare, fallbasierte Aufgaben) und den MC-Prüfungen bzw. EMERGE waren nicht stark ausgeprägt. Daher liegt es nahe, dass die auf die Lehrformate angewendeten Bildungsziele Wissenserwerb und Wissensanwendung nicht scharf voneinander zu trennen sind. Ebenso ließ die signifikante Korrelation zwischen den in #clue erreichten Prozentwerten und den in den E-Fallseminaren erreichten Prozentwerten auf eine Unschärfe der Trennung der Bildungsziele hindeuten. #Clue verfolgt das Bildungsziel Wissenserwerb und -bewahrung, während mit den E-Fallseminaren eine Stärkung des *clinical reasoning* sowie die Besserung der Wissensanwendung angestrebt wird.

Abschließend ließ sich jedoch feststellen, dass Studierende sowohl mit den MC-Fragen als auch mit den *key-feature*-Fragen problemlos zurechtkamen. Diese Annahme beruht auf der Korrelation zwischen den Prozentwerten der Abschlussprüfung des Moduls M3.1 und den erreichten Prozentwerten in den E-Fallseminaren. Ebenso stimmte die Analyse dieser Studie mit Resultaten vorheriger Studien überein. Je mehr Intervention Studierende erhielten bzw. je mehr Aufgaben und Fragen sie erledigten, desto höher war der Lernerfolg und die erreichten Prozentwerte in der Abschlussprüfung (Details in Kapitel 4.2). Dabei gab es keine Unterschiede darin, welcher Fragetypus verwendet oder welches Bildungsziel verfolgt wurde.

4.3 Beurteilung der Leistungsverläufe

Zur Beurteilung der Leistungsverläufe der Studierenden wurden die zweizeitig erbrachten Leistungen der Studierenden aus dem Wintersemester 2019/2020 (bzw. dem Sommersemester 2019) sowie aus der Modulabschlussklausur des Sommersemesters 2020 miteinander verglichen. In der vorliegenden Analyse konnte eine signifikante Korrelation zwischen den Ergebnissen der Modulabschlussklausur (Modul M3.1) sowie den Prüfungen des zweiten klinischen Semesters nachgewiesen werden. Dadurch ließ sich die Annahme bestätigen, dass beide Konstrukte ähnliche Kompetenzen abfragen.

Weiterhin konnten unterschiedlich starke Leistungen der Geschlechter in den Prüfungen erkannt werden. Während es in den Prüfungen des zweiten klinischen Semesters keinen signifikanten Unterschied zwischen Männern und Frauen gab, wurde in der Abschlussklausur des Moduls M3.1 ein signifikanter Unterschied nachgewiesen: Studentinnen schnitten in der Modulabschlussklausur signifikant schlechter ab als Studenten. Der gemessene signifikante Unterschied im ersten Corona-Semester könnte auf eine Umstellung der Präsenzlehre hin zu einer digitalen Lehre beruhen.

Mögliche Erklärungsansätze könnten sein, dass der fehlende soziale Kontakt eine entscheidende Rolle einnehmen könne. Die Pandemie und Isolation waren für die Bevölkerung eine große psychische Belastung (Robinson et al. 2021). Es gibt Hinweise darauf, dass Frauen unter der sozialen Isolation mehr litten als Männer (Elmer et al. 2020). Sie erkrankten während der Pandemie besonders häufig an psychischen Erkrankungen wie Depressionen (Elmer et al. 2020; Racine et al. 2021). Diese Auswirkungen spiegeln sich möglicherweise in der Motivation und der Leistung wider. Wie Judge und Kollegen erkannten, stehen psychische Zufriedenheit und Leistung bzw. Erfolg in einer engen Beziehung zueinander (Judge et al. 2001).

Nicht nur die psychische Belastung während der Pandemie könnte ein Grund für den Leistungsabfall der weiblichen Studierenden sein. Ebenso ließe sich der Abfall in der Änderung des Lernverhaltens/Lernmöglichkeiten begründen. Mit dem Beschluss der Kontaktbeschränkungen war ein Lernen in Gruppen nicht mehr möglich. Zudem schlossen Bibliotheken und weitere Einrichtungen, die von Studierenden zum Lernen genutzt wurden. Statt des Lernens in der *peer-groups* mussten Studierende alleine lernen oder auf Technologien wie „Sykpe“ und „Zoom“ zurückgreifen.

Neben dem veränderten Lernverhalten und der psychischen Belastung könnte auch die familiäre Situation einen Erklärungsansatz bieten. Woitowich und Kollegen sowie

Schiebinger und Gilmartin gaben an, dass Frauen mehr im Haushalt eingebunden sind als Männer (Schiebinger und Gilmartin 2010; Woitowich et al. 2021). Des Weiteren schrieben Jolly und Kollegen, dass sich Frauen eher frei nehmen als Männer, wenn es zu unerwarteten Problemen der Kinderbetreuung kommt, beispielsweise durch Krankheiten (Jolly et al. 2014). Die Einengung Zuhause und Ablenkung durch die Familie trägt möglicherweise zu einem schlechteren Lernprozess der Studentinnen bei, mit einem folglich ungünstigeren Lernergebnis (Dost et al. 2020). Zusätzlich führten die Schließungen von Schulen und Kitas im Rahmen der Covid-19-Pandemie zur erhöhten Belastung im häuslichen Umfeld.

Die genannten Faktoren könnten Hinweise auf die gefundene Assoziation zwischen Geschlecht und Prüfungserfolg liefern. Allerdings scheinen die Faktoren pandemiebedingt zu sein und spiegeln wahrscheinlich das Konzept einer seit Jahrzehnten bestehenden Rollenverteilung wider. Durch die Einschränkungen der Pandemie (fehlende Kinderbetreuung, etc.) entstanden Zusatzbelastungen zum Nachteil des Studiums. Da lediglich 5 % der Studierenden bereits ein Kind haben, betrifft diese Erklärung nur wenige Personen der betrachteten Kohorte (Middendorff et al. 2013).

Die bisherige Forschung deutet darauf hin, dass Studentinnen nicht ausschließlich durch die Eingliederung digitaler Lehrformate gegenüber ihren männlichen Kommilitonen benachteiligt sind. Das Ergebnis der Datenanalyse wird vermutlich von weiteren Störfaktoren beeinflusst, die pandemiebedingt sind (psychische Belastungen, verändertes Lernverhalten, etc.). Die nachgewiesenen Geschlechtsunterschiede müssen im Kontext dieser pandemischen Ausnahmesituation betrachtet werden und resultieren nicht allein aus der Umstellung des Lehrkonzepts.

4.4 Einflussfaktoren auf den Lernerfolg

4.4.1 Die studentische Prüfungsleistung

Zu Anfang der Studie wurde die Hypothese aufgestellt, dass eine intensivere Nutzung der digitalen Ressourcen mit dem Ziel Wissenserwerb und -bewahrung zu höheren Prozentwerten in der summativen MC-Prüfung führen.

Die Ergebnisse der univariaten Analyse wiesen auf, dass hohe Leistungen in den vorherigen Prüfungen des zweiten klinischen Semesters und das männliche Geschlecht positive Einflüsse auf die Modulabschlussklausur (im Modul M3.1) hatten. Des Weiteren scheinen hohe Prozentwerte in #clue sowie den E-Fallseminaren und eine geringe Verzögerung bei

der Bearbeitung der E-Fallseminare sich ebenfalls positiv auf die Ergebnisse der Modulabschlussprüfung auszuwirken. Im adjustierten multivariaten Modell blieben allerdings nur das männliche Geschlecht und die frühere Prüfungsleistung als prädiktive Faktoren für eine hohe Leistung in der Abschlussprüfung des Moduls M3.1 bestehen.

Die fehlenden Assoziationen zwischen den Nutzungscharakteristika und der Prüfungsleistung lassen vermuten, dass die digitalen Ressourcen nur einen geringen Einfluss auf den Lernerfolg haben. Frauen haben damit nicht per se durch die Einführung der digitalen *tools* einen Nachteil erfahren. Wie bereits in Kapitel 4.3 erwähnt, deuten bisherige Studien nicht daraufhin, dass Frauen allein durch die digitale Lehre benachteiligt sind. Vielmehr scheinen pandemiebedingte Störfaktoren, wie emotionaler Stress, Isolation oder Überbelastung im Haushalt, Gründe für eine geringere Prüfungsleistung der Frauen zu sein (Dost et al. 2020; Elmer et al. 2020; Woitowich et al. 2021).

Neben dem Geschlecht liefern auch die Ergebnisse aus den Abschlussprüfungen des zweiten klinischen Semesters eine Vorhersage über die Leistungen in der Abschlussprüfung des Moduls M3.1. Studierende, die im vorherigen Semester bereits gut waren, schnitten auch in der Prüfung des Moduls M3.1 gut ab. Gründe dafür könnten eine hohe Intelligenz oder hingegen eine gute Anpassung an das didaktische Konzept des (Medizin-)Studiums sein. Diese Faktoren sind allerdings individuell und kaum zu beeinflussen.

Abschließend lässt sich beurteilen, dass ein Verlust der Mehrheit der signifikanten Assoziationen in der multivariaten Analyse auf einen starken Einfluss des Geschlechts und der vorherigen Prüfungsleistungen in den univariaten Regressionsanalyse zurückzuführen ist. Die signifikanten Assoziationen in der univariaten Analyse waren vor allem durch die zwei Faktoren Geschlecht und vorherige Prüfungsleistung bestimmt. Damit sind das männliche Geschlecht sowie Prüfungsleistungen aus dem zweiten klinischen Semester in dieser Studie die einzigen Einflussfaktoren auf den Lernerfolg hinsichtlich des Faktenwissens (gemessen in einer summativen MC-Klausur). Diese Faktoren sind allerdings nicht beeinflussbar.

4.4.2 Studentische Leistung in EMERGE

Durch das *serious game* EMERGE wurde die Anwendung des erlernten Wissens durch virtuelle, real jedoch vorstellbare Patientenfälle abgefragt. Andere Studien untersuchten bereits den Lernerfolg durch solche computergestützten Spiele:

Schon 2013 berichteten Boeker und Kollegen, dass spielbasiertes *E-Learning* zu einem größeren Lernerfolg führe als herkömmliches Studieren mittels Skripten (Boeker et al. 2013).

Dazu führten sie eine Studie an Medizinstudierenden des Universitätsklinikums Freiburg durch, die entweder ein Training mittels Skript oder anhand eines computergestützten Spiels absolvierten. Im Anschluss wurde der Lernerfolg durch einen Test überprüft. Resultierend zeigte sich eine signifikant höhere Leistung im abschließenden Test in der Studienkohorte, die ein Training an dem computergestützten Spiel erhalten hat. Die Gruppe, die sich mit Hilfe eines Skripts auf den Test vorbereitete, schnitt signifikant schlechter ab (Gruppe Spielsitzung: 28,6 Punkte; Gruppe Skript: 26,0 Punkte; von insgesamt 34 Punkten). Neben dem Lernerfolg waren auch die Motivation und der Spaß-Faktor ein positives Ergebnis dieser Lehrintervention (Boeker et al. 2013). Auch Mansoori und Kollegen sowie Yang und Kollegen beschrieben die Attraktivität solcher digitaler Lehrformate (Mansoori et al. 2021; Yang et al. 2021).

Während andere Studien sich vor allem auf den Lernerfolg durch *serious games* und die Effektivität eines solchen Formats konzentrierten (Gentry et al. 2019; Hu et al. 2021; Tubelo et al. 2019), wurde in dieser Studie EMERGE als neues Prüfungsformat für die Anwendung des erlernten Wissens eingesetzt. Auch Aster und Kollegen untersuchten im Wintersemester 2019/2020, ob sich EMERGE neben der ressourcenintensiven OSCE (*objective structured clinical examination*)-Prüfung ebenso als Prüfungsformat eignet (Aster et al. 2022).

Wie Middeke und Kollegen beschrieben, können *serious games* wie EMERGE dazu beitragen, klinisch orientiertes Denken effektiv zu verbessern (Middeke et al. 2018). Sie zeigten, dass Studierende, die an EMERGE teilnahmen, vor allem hinsichtlich der richtigen Diagnose sowie therapeutischen Anordnung, die die Kernkompetenzen des *clinical reasoning* sind, bessere Ergebnisse erzielten (Middeke et al. 2018). In der Studie schnitten Studierende, die an EMERGE teilnahmen, signifikant besser in einer *key-feature*-Prüfung ab als diejenigen Studierende, die in Kleingruppensitzungen fallbasierte Aufgaben bearbeiteten (EMERGE: $62,5 \pm 17,7$ %; Kleingruppensitzungen: $54,2 \pm 21,9$ %). Auch die Arbeitsgruppe von Wu und Kollegen erkannte die durch computergestützte Spiele gegebenen Chancen (Wu et al. 2022): Virtuelle Simulationen können Handlungen unter Stresssituationen simulieren und klinische Fähigkeiten in komplexen Situationen verbessern. *Serious games* stellen eine Chance für Studierende (Krankenpflegeschüler*innen und weitere Berufsgruppen) dar, *clinical reasoning*-Kompetenzen zu stärken und das Verhalten im Umgang mit Stresssituationen im klinischen Alltag zu optimieren.

So sollte in dieser Studie das Format EMERGE dazu genutzt werden, die Anwendung des erlernten Wissens im Sinne des *clinical reasoning* zu prüfen. Dazu wurde analysiert, ob und wie sich die unterschiedlichen Ressourcen auf die studentische Leistung in EMERGE auswirken.

Durch die vorherige Einteilung der Ressourcen in Bildungsziele, wurde angenommen, dass sich diejenigen Bildungsziele, die sich auf die Anwendung des erlernten Wissens beziehen, die Leistungen in EMERGE beeinflussen.

Aus den gesammelten Daten ließ sich keine Korrelation zwischen EMERGE und dem fragenbasierten *tool #clue* nachweisen. Allerdings konnte eine positive Korrelation zwischen den *#clue*-Prozentwerten und den Prüfungsergebnissen erkannt werden. Dies entsprach auch den initialen Annahmen: Ein fragenbasiertes *tool* wie *#clue*, das als Bildungsziel Wissensbewahrung und -erwerb verfolgt, führt zu höheren Prozentwerten in der MC-Prüfung. Das erworbene Wissen wird dabei in den MC-Prüfungen passiv abgefragt. EMERGE hingegen stärkt die Fähigkeiten des *clinical reasoning* und damit die Wissensanwendung. Dementsprechend verfolgen EMERGE und *#clue* unterschiedliche Bildungsziele und messen unterschiedliche Kompetenzen der zukünftigen Ärzte*innen. Des Weiteren konnte kein Zusammenhang zwischen den Leistungen der Modulabschlussklausur und den in EMERGE erbrachten Leistungen festgestellt werden. Auch dieses Ergebnis entsprach den vorherigen Erwartungen, da diese beiden Prüfungsformate unterschiedliche Kompetenzen (Anwendung des erlernten Wissens und Wissenserwerb) der Studierenden abfragen.

In dieser Studie zeigte sich keine Assoziationen zwischen dem *serious game* und der Abschlussprüfung. Dennoch scheint es in der Literatur Hinweise darauf zu geben, dass virtuelle Spiele auch zu einer Verbesserung der Prüfungsleistung beitragen können (Boeker et al. 2013).

In den anschließenden Analysen wurden Prädiktoren für eine hohe Punktzahl in EMERGE gesucht. Um jene Prädiktoren zu identifizieren, wurden univariate lineare Regressionsanalysen durchgeführt. Es wurde erwartet, dass eine intensivere Nutzung derjenigen Ressourcen, die das Bildungsziel Wissensanwendung verfolgen, zu höheren Prozentwerten in dem *serious game* EMERGE führen. Die digitale Ressource *#clue* zeigte keine Assoziationen zu den Leistungen in EMERGE. Dies entsprach auch den Erwartungen, da unterschiedliche Bildungsziele verfolgt wurden. Während EMERGE die Wissensanwendung prüft, zielt *#clue* auf den Wissenserwerb sowie die Wissensbewahrung ab. Weiterhin war eine hohe Punktzahl in den E-Fallseminaren mit einer besseren Leistung in der Spielsitzung EMERGE assoziiert. Aufgrund der einzigen signifikanten Assoziation, wurde keine multivariate Analyse im Folgenden durchgeführt. Im Vorfeld bestand bereits die Annahme, dass sich eine signifikante Assoziation zwischen EMERGE und den E-Fallseminaren finden lässt, da beide auf das Bildungsziel Wissensanwendung abzielen.

Allerdings konnte in den Analysen auch eine signifikante Assoziation zwischen den Prozentwerten in den E-Fallseminaren und den Prozentwerten der summativen Abschlussprüfung gefunden werden. Diese beiden Formate zielen nicht auf das gleiche Bildungsziel ab. Daher liegt es nahe, dass eine gute Leistung in den E-Fallseminaren keine direkte Auswirkung auf die studentische Leistung in der virtuellen Notaufnahme EMERGE hat. Vielmehr kann angenommen werden, dass leistungsstarke Studierende mit den E-Fallseminaren, der summativen MC-Prüfung sowie mit EMERGE problemlos zurechtkamen ganz unabhängig von dem Bildungsziel, das sich dahinter verbirgt.

4.5 Wertung der Prädiktoren für negative Leistungsveränderungen

Neben der in Kapitel 4.4 erläuterten Annahme, gab es eine weitere Hypothese: Die Hypothese beschrieb, dass eine geringe Verzögerung bei der Abgabe der ersten fallbasierten Aufgabe, der Bearbeitung des ersten E-Fallseminars und der ersten zehn #clue-Fragen zu hohen Leistungen in der Modulabschlussklausur führten. Gleichzeitig wurde erwartet, dass Studierende, die hohe Prozentwerte in dem ersten E-Fallseminar und den ersten zehn #clue-Fragen erreichten, ebenso gute Ergebnisse in der Prüfung hervorbringen.

Zur Bestätigung der Hypothese, wurde nach Vorhersagewerten gesucht, die am Anfang des Moduls erfasst werden können und auf die spätere Leistung am Ende des Moduls deuten lassen. Eine gute Leistung im ersten E-Fallseminar, gekennzeichnet durch einen höheren Prozentwert, war ein signifikanter Vorhersagewert für die abschließende Prüfungsleistung. Dies ging aus der univariaten Regressionsanalyse hervor und ließ sich durch die multivariate Regressionsanalyse bestätigen. Bereits Marden und Kollegen postulierten, dass Ergebnisse einer Lehrintervention auf anschließende Prüfungsleistungen hindeuten können (Marden et al. 2013). Da es sich bei den E-Fallseminaren und der Abschlussprüfung um unterschiedliche Fragetypen handelt, liegt es nahe, dass leistungsstarke Studierende beide Fragetypen problemlos bearbeiten können.

Die folgende Analyse hingegen konnte nicht bestätigen, dass die Verzögerung der ersten Aufgabe sowie die Verzögerung der Bearbeitung der ersten zehn #clue-Fragen in Zusammenhang mit der Prüfungsleistung stehen. Weiterhin ließ sich keine Assoziation von der Fragenanzahl bzw. Seminaranzahl und der aufgewandten Zeit zu den Prüfungsergebnissen erkennen. Die univariaten Regressionsanalyse zeigte einen Zusammenhang zwischen einer schnellen Bearbeitung des ersten E-Fallseminars (sowie Prozentwerten in den ersten zehn #clue-Fragen) mit den Prüfungsleistungen. Dieser Effekt verschwand jedoch in der multivariaten Regression. Ein möglicher Grund für das Fehlen der

Assoziationen in der multivariaten Regression könnte ein übergeordneter Effekt des Geschlechts sowie der Prüfungsleistungen aus dem vorherigen Semester sein.

Wie in der Auswertung zu sehen, war die studentische Leistung im ersten E-Fallseminaren der einzige Vorhersagewert. Es konnte nachgewiesen werden, dass Studierende, die eine hohe Leistung im ersten E-Fallseminar zeigten, ebenso eine hohe Prüfungsleistung erbrachten. Dementsprechend ist es möglich, gefährdete Studierende, die weniger leistungsstark im ersten E-Fallseminar sind, mit zusätzlichen Lehrinterventionen zu unterstützen. Dringlich ist es, solche Studierende frühzeitig zu erkennen, da durch eine frühe Intervention schwindende Motivation, (Versagens-)Ängste und gar Studienabbrüche reduziert werden können (Heublein und Wolter 2011; Ludwig AB et al. 2015). Studierenden am Anfang des Moduls Unterstützung für ein besseres Verständnis von Basisthemen (wie Medikamentenwirkungen oder Stoffwechselfvorgänge) anzubieten, ist besonders sinnvoll in Hinblick auf die folgenden, aufbauenden Lehrinhalte.

4.6 Stärken und Limitationen

Bislang wurde die Nutzung der Lernressourcen in einem rein digitalen Curriculum und deren Auswirkung auf den Lernerfolg der Studierenden noch nicht näher untersucht. Mit dieser Studie zu Beginn der Corona-Pandemie wurden erstmals die Auswirkungen einer reinen Online-Lehre auf die Leistungen der Studierenden untersucht. Besonders hervorzuheben ist der große Stichprobenumfang der Studie sowie die hohe Rücklaufquote von 95,6 %. Mit der Implementierung verschiedener Lehrformate, deren Entwicklung und Design an Bildungsziele angepasst wurden, konnte ein umfangreicher Datensatz generiert werden. Dadurch war es nicht nur möglich, die Leistungen in den einzelnen Lehrformaten zu beurteilen, sondern auch das Lernverhalten der Studierenden zu untersuchen.

Der Pandemie geschuldet, gab es jedoch keine Kontrollgruppe. Jeder Studierende im Sommersemester 2020 durchlief das gleiche Online-Modul. Trotz des umfangreichen Datensatzes war nicht bekannt, wie viel Zeit die Studierenden in das Selbststudium investierten und welche Ressourcen die Studierenden zusätzlich genutzt haben. Durch einen Austausch der Lösungen der einzelnen Fragen, Fallseminare und fallbasierten Aufgaben unter den Studierenden kann eine Verschiebung der Daten nicht ausgeschlossen werden. Weiterhin erhob die Studie keine Informationen über die Lernstrategie der Studierenden. Zur Beurteilung der Leistungsverläufe wären diesen Daten durchaus hilfreich.

4.7 Weiterführende Forschung

Um spezifische Gründe für die Unterschiede der Leistungen bei Männern und Frauen identifizieren zu können, wird in Zukunft mehr qualitative Forschung nötig sein. Zudem sollte eine Analyse auch auf weitere Bildungsinstitutionen erweitert werden.

Gelingt es, gefährdete Studierende frühzeitig zu erkennen, müssen gezielte Strategien zur Unterstützung und Förderung entwickelt werden. Die stetige Evaluation der Strategien ist dabei eingeschlossen. Beispielsweise könnten gefährdete Studierende weitere vertiefende Fälle, Vorlesungen oder persönliche Meetings angeboten werden, die gezielt auf die Schwächen der Studierenden abgestimmt sind. Dazu ist es notwendig, dass Frühwarnsysteme genau die Themen erkennen, die zu den Schwächen des Studierenden zählen.

Zusätzlich könnten weitere individuelle Charakteristika der Studierenden mit in die Studie aufgenommen werden. Dazu gehören gezielte Befragungen der Studierenden zu ihrer Wohnsituation sowie familiären Struktur. Des Weiteren könnten vorherige psychische Erkrankungen und das subjektive Erleben des Lockdowns erhoben werden. Für die Beurteilung des Lernverhaltens der Studierenden wäre es zudem wichtig, die vorherige Lernstrategie (Gruppen-Lernen oder Einzel-Lernen) sowie zu Zeiten der Pandemie zu erfragen. Dies könnte im Rahmen von Interviews oder Fragebögen geschehen. Wichtig wäre zu erkennen, ob es im Lernverhalten Geschlechtsunterschiede gibt, die Rückschlüsse auf die Prüfungsleistung ziehen lassen.

5 Zusammenfassung

Mit dem Beginn der Corona-Pandemie wurde die universitäre Lehre maßgeblich umstrukturiert: Da konservative Präsenzveranstaltungen nach den Bestimmungen des Infektionsschutzgesetzes nicht in gewohntem Maße stattfinden konnten, wurden vermehrt digitale Ersatzformate eingesetzt. Für das Modul M3.1 „Erkrankungen des Herz-Kreislauf-Systems und der Lunge“ an der Universitätsmedizin Göttingen standen den Studierenden der Humanmedizin neben aufgezeichneten Vorlesungen auch fallbasierte Aufgaben, E-Fallseminare, #clue-Fragen sowie die virtuelle Notaufnahme EMERGE zur Verfügung.

Ziel der Studie war es, (1) die studentische Leistung vor der Pandemie und im ersten Semester nach Beginn der Pandemie zu vergleichen, um Änderungen im Lernerfolg aufzudecken. Ebenso (2) sollten Zusammenhänge zwischen den Nutzungscharakteristika der digitalen Lehrangebote und dem Lernerfolg identifiziert werden, sowie (3) Prädiktoren ausfindig gemacht werden, die eine ungünstige Prüfungsleistung vorhersagen. Dazu wurden T-Tests, Korrelationen und Regressionsanalysen durchgeführt.

Die hier vorliegende prospektive Kohortenstudie deutet auf eine unterschiedliche Nutzung der digitalen Lernressourcen hin. Wie aus den Beobachtungen zu schlussfolgern ist, sind Studentinnen in dem überwiegend digitalen Lehrplan gegenüber ihren männlichen Kommilitonen benachteiligt. Die adjustierte Analyse gab jedoch keinen Hinweis darauf, dass eine Benachteiligung ausschließlich durch den Einsatz der digitalen Lehrformate zu erklären ist. Vielmehr müssen weitere zugrundeliegende Ursachen näher untersucht und eventuelle Störfaktoren, die größtenteils pandemiebedingt sind, identifiziert werden. Ebenso wiesen die Analysen einen Zusammenhang zwischen den Leistungen des ersten E-Fallseminars und den Prüfungsleistungen nach. Damit konnte ein Prädiktor identifiziert werden, der gefährdete Studierende frühzeitig erkennen lässt. Eine rechtzeitige Unterstützung Derselben wird dadurch ermöglicht.

Neben der Förderung des individuellen Lernstils durch die unterschiedlichen Lernangebote, kann die Umstrukturierung wesentlich zur Ressourcenallokation innerhalb der medizinischen Ausbildung beitragen. Dementsprechend werden die digitalen Formate mit großer Wahrscheinlichkeit auch nach der Pandemie noch einen hohen Stellenwert in der Ausbildung der zukünftigen Ärzte*innen einnehmen.

6 Literaturverzeichnis

- Ackermann O, Siemann H, Schwarting T, Ruchholtz S (2010): Effektives training der chirurgischen röntgenbefundung durch E-learning. *Z Orthop Unfall* 148, 348-352
- Al-Balas M, Al-Balas HI, Jaber HM, Obeidat K, Al-Balas H, Aborajoo EA, Al-Taher R, Al-Balas B (2020): Distance learning in clinical medical education amid COVID-19 pandemic in Jordan: current situation, challenges, and perspectives. *BMC Med Educ* 20, 341
- Aldeman NLS, de Sá Urtiga Aita KM, Machado VP, da Mata Sousa LCD, Coelho AGB, da Silva AS, da Silva Mendes AP, de Oliveira Neres FJ, do Monte SJH (2021): Smartpath(k): a platform for teaching glomerulopathies using machine learning. *BMC Med Educ* 21, 248
- Ali A, Subhi Y, Ringsted C, Konge L (2015): Gender differences in the acquisition of surgical skills: a systematic review. *Surg Endosc* 29, 3065-3073
- Aristovnik A, Keržič D, Ravšelj D, Tomažević N, Umek L (2020): Impacts of the COVID-19 Pandemic on Life of Higher Education Students: A Global Perspective. *Sustainability* 12, 8438
- Aster A, Scheithauer S, Middeke AC, Zegota S, Clauberg S, Artelt T, Schuelper N, Raupach T (2022): Use of a Serious Game to Teach Infectious Disease Management in Medical School: Effectiveness and Transfer to a Clinical Examination. *Front Med (Lausanne)* 2, 863764
- Augustin M (2014): How to learn effectively in medical school: test yourself, learn actively, and repeat in intervals. *Yale J Biol Med* 87, 207-212
- Austin A, Rudolf F, Fernandez J, Ishimine P, Murray M, Suresh P, McDaniel M, Shishlov K, Oyama L (2021): COVID-19 educational innovation: Hybrid in-person and virtual simulation for emergency medicine trainees. *AEM Educ Train* 5, e10593
- Bączek M, Zagańczyk-Bączek M, Szpringer M, Jaroszyński A, Woźakowska-Kapłon B (2021): Students' perception of online learning during the COVID-19 pandemic: A survey study of Polish medical students. *Medicine (Baltimore)* 100, e24821
- Barber C, Hammond R, Gula L, Tithcott G, Chahine S (2018): In Search of Black Swans: Identifying Students at Risk of Failing Licensing Examinations. *Acad Med* 93, 478-485
- Biggs J (1996): Enhancing teaching through constructive alignment. *High Educ (Dordr)* 32, 347-364
- Boeker M, Andel P, Vach W, Frankenschmidt A (2013): Game-based e-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. *PLoS One* 8, e82328

- Brusamento S, Kyaw BM, Whiting P, Li L, Tudor Car L (2019): Digital Health Professions Education in the Field of Pediatrics: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. *J Med Internet Res* 21, e14231
- Bundesregierung D (Hrsg.): Besprechung der Bundeskanzlerin mit den Regierungschefinnen und Regierungschefs der Länder vom 22.03.2020. Die Bundeskanzlerin und die Regierungschefinnen und Regierungschefs der Länder fassen am 22. März 2020 folgenden Beschluss. Bundesregierung Deutschland, Berlin 2020;
<https://www.bundesregierung.de/breg-de/themen/coronavirus/besprechung-der-bundeskanzlerin-mit-den-regierungschefinnen-und-regierungschefs-der-laender-vom-22-03-2020-1733248>; abgerufen am 22. Oktober 2022
- Butler AC (2010): Repeated testing produces superior transfer of learning relative to repeated studying. *J Exp Psychol Learn Mem Cogn* 36, 1118-1133
- Carpenter SK (2009): Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *J Exp Psychol Learn Mem Cogn* 35, 1563-1569
- Carroll JB (1989): The Carroll Model: A 25-Year Retrospective and Prospective View. *Educ Res* 18, 26-31
- Chang WJ, Jiang YD, Xu JM (2020): [Experience of teaching and training for medical students at gastrointestinal surgery department under COVID-19 epidemic situation]. *Zhonghua Wei Chang Wai Ke Za Zhi* 23, 616-618
- Chumley-Jones HS, Dobbie A, Alford CL (2002): Web-based learning: sound educational method or hype? A review of the evaluation literature. *Acad Med* 77, S86-S93
- Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM (2008): Internet-based learning in the health professions: a meta-analysis. *JAMA* 300, 1181-1196
- Cook DA, Levinson AJ, Garside S, Dupras DM, Erwin PJ, Montori VM (2010): Instructional design variations in internet-based learning for health professions education: a systematic review and meta-analysis. *Acad Med* 85, 909-922
- Darici D, Reissner C, Brockhaus J, Missler M (2021): Implementation of a fully digital histology course in the anatomical teaching curriculum during COVID-19 pandemic. *Ann Anat* 236, 151718
- De Ponti R, Marazzato J, Maresca AM, Rovera F, Carcano G, Ferrario MM (2020): Pre-graduation medical training including virtual reality during COVID-19 pandemic: a report on students' perception. *BMC Med Educ* 20, 332

- Djermester P, Gröschke C, Gintrowicz R, Peters H, Degel A (2021): Bedside teaching without bedside - an introduction to clinical reasoning in COVID-19 times. *GMS J Med Educ* 38, Doc14
- Dost S, Hossain A, Shehab M, Abdelwahed A, Al-Nusair L (2020): Perceptions of medical students towards online teaching during the COVID-19 pandemic: a national cross-sectional survey of 2721 UK medical students. *BMJ Open* 10, e042378
- Elmer T, Mephram K, Stadtfeld C (2020): Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. *PLoS One* 15, e0236337
- Elshami W, Taha MH, Abuzaid M, Saravanan C, Al Kawas S, Abdalla ME (2021): Satisfaction with online learning in the new normal: perspective of students and faculty at medical and health sciences colleges. *Med Educ Online* 26, 1920090
- Emmanouil P, Pavlos M, Avgerinos ED, Sofia A, Christos T (2008): Evolution of medical education in ancient Greece. *Chin Med J (Engl)* 121, 2202-2206
- Fairén M, Moyés J, Insa E (2020): VR4Health: Personalized teaching and learning anatomy using VR. *J Med Syst* 44, 94
- Foohy S, Nagji A, Yilmaz Y, Sibbald M, Monteiro S, Chan TM (2022): Developing the Virtual Resus Room: Fidelity, Usability, Acceptability, and Applicability of a Virtual Simulation for Teaching and Learning. *Acad Med* 97, 679-683
- Fransen F, Martens H, Nagtzaam I, Heeneman S (2018): Use of e-learning in clinical clerkships: effects on acquisition of dermatological knowledge and learning processes. *Int J Med Educ* 2, 11-17
- Gentry SV, Gauthier A, L'Estrade Ehrstrom B, Wortley D, Lilienthal A, Tudor Car L, Dauwels-Okutsu S, Nikolaou CK, Zary N, Campbell J, et al. (2019): Serious Gaming and Gamification Education in Health Professions: Systematic Review. *J Med Internet Res* 21, e12994
- George PP, Zhabenko O, Kyaw BM, Antoniou P, Posadzki P, Saxena N, Semwal M, Tudor Car L, Zary N, Lockwood C, et al. (2019): Online Digital Education for Postregistration Training of Medical Doctors: Systematic Review by the Digital Health Education Collaboration. *J Med Internet Res* 21, e13269
- Gettinger M (1984): Individual differences in time needed for learning: A review of literature. *Educ Psychol* 19, 15-29
- Halamish V, Bjork RA (2011): When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *J Exp Psychol Learn Mem Cogn* 37, 801-812

- Heinzmann A, Bode S, Forster J, Berger J (2021): Interactive, case-based seminars in the digitized pediatrics block internship from the students' perspective. *GMS J Med Educ* 38, Doc24
- Heublein U, Wolter A (2011): Studienabbruch in Deutschland. Definition, Häufigkeit, Ursachen, Maßnahmen. *Z Padagogik* 57, 214-236
- Hew KF, Lo CK (2018): Flipped classroom improves student learning in health professions education: a meta-analysis. *BMC Med Educ* 18, 38
- Hoff E, Haberstroh N, Sostmann K, Perka C, Putzier M, Schmidmaier G, Back DA (2014): [E-learning in orthopedics and traumatology. A comparative pilot study on acceptance and knowledge acquisition among users and non-users]. *Orthopade* 43, 674-680
- Hrynychak P, Takahashi SG, Nayer M (2014): Key-feature questions for assessment of clinical reasoning: a literature review. *Med Educ* 48, 870-883
- Hu L, Zhang L, Yin R, Li Z, Shen J, Tan H, Wu J, Zhou W (2021): NEOGAMES: A Serious Computer Game That Improves Long-Term Knowledge Retention of Neonatal Resuscitation in Undergraduate Medical Students. *Front Pediatr* 9, 645776
- Huber J, Wittl M, Schunk M, Fischer MR, Tolks D (2021): The use of the online Inverted Classroom Model for digital teaching with gamification in medical studies. *GMS J Med Educ* 38, Doc3
- Irby DM, Wilkerson L (2003): Educational innovations in academic medicine and environmental trends. *J Gen Intern Med* 18, 370-376
- Islam MA, Barna SD, Raihan H, Khan MNA, Hossain MT (2020): Depression and anxiety among university students during the COVID-19 pandemic in Bangladesh: A web-based cross-sectional survey. *PLoS One* 15, e0238162
- Jawaid M, Bakhtiar N, Masood Z, Mehar AK (2019): Effect of Paper- and Computer-based Simulated Instructions on Clinical Reasoning Skills of Undergraduate Medical Students: A Randomized Control Trial. *Cureus* 11, e6071
- Jolly S, Griffith KA, DeCastro R, Stewart A, Ubel P, Jaggi R (2014): Gender differences in time spent on parenting and domestic responsibilities by high-achieving young physician-researchers. *Ann Intern Med* 160, 344-353
- Judge TA, Thoresen CJ, Bono JE, Patton GK (2001): The job satisfaction-job performance relationship: a qualitative and quantitative review. *Psychol Bull* 127, 376-407
- Kang SY, Kim SH, Kwon YE, Kim TB, Park HK, Park HW, Chang YS, Jee YK, Moon HB, Min KU, et al. (2018): The virtual asthma guideline e-learning program: learning effectiveness and user satisfaction. *Korean J Intern Med* 33, 604-611

- Khalil R, Mansour AE, Fadda WA, Almisnid K, Aldamegh M, Al-Nafeesah A, Alkhalifah A, Al-Wutayd O (2020): The sudden transition to synchronized online learning during the COVID-19 pandemic in Saudi Arabia: a qualitative study exploring medical students' perspectives. *BMC Med Educ* 20, 285
- Kim RH, Brinster NK, Meehan SA (2020): Dermatopathology education during the COVID-19 pandemic: Virtual simulation of the multiheaded microscope. *J Am Acad Dermatol* 83, e243-e244
- Knie K, Schwarz L, Frehle C, Schulte H, Taetz-Harrer A, Kiessling C (2020): To zoom or not to zoom - the training of communicative competencies in times of Covid 19 at Witten/Herdecke University illustrated by the example of "sharing information". *GMS J Med Educ* 37, Doc83
- Kopp V, Möltner A, Fischer MR (2006): Key-Feature-Probleme zum Prüfen von prozeduralem Wissen: Ein Praxisleitfaden. *Gms Z Med Ausbild* 23, 2006-2023
- Kromann CB, Jensen ML, Ringsted C (2011): Test-enhanced learning may be a gender-related phenomenon explained by changes in cortisol level. *Med Educ* 45, 192-199
- Kühl SJ, Toberer M, Keis O, Tolks D, Fischer MR, Kühl M (2017): Concept and benefits of the Inverted Classroom method for a competency-based biochemistry course in the pre-clinical stage of a human medicine course of studies. *GMS J Med Educ* 34, Doc31
- Kuo KH, Leo JM (2019): Optical Versus Virtual Microscope for Medical Education: A Systematic Review. *Anat Sci Educ* 12, 678-685
- Larsen DP, Butler AC, Roediger HL 3rd (2008): Test-enhanced learning in medical education. *Med Educ* 42, 959-966
- Larsen DP, Butler AC, Roediger HL 3rd (2009): Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med Educ* 43, 1174-1181
- Larsen DP, Butler AC, Roediger HL 3rd (2013a): Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Med Educ* 47, 674-682
- Larsen DP, Butler AC, Lawson AL, Roediger HL 3rd (2013b): The importance of seeing the patient: test-enhanced learning with standardized patients and written tests improves clinical application of knowledge. *Adv Health Sci Educ Theory Pract* 18, 409-425
- Lee J, Solomon M, Stead T, Kwon B, Ganti L (2021): Impact of COVID-19 on the mental health of US college students. *BMC Psychol* 2, 95
- Leong S, Mc Laughlin P, O'Connor OJ, O'Flynn S, Maher MM (2012): An assessment of the feasibility and effectiveness of an e-learning module in delivering a curriculum in radiation protection to undergraduate medical students. *J Am Coll Radiol* 2, 203-209

- Ludwig AB, Burton W, Weingarten J, Milan F, Myers DC, Kligler B (2015): Depression and stress amongst undergraduate medical students. *BMC Med Educ* 15, 141
- Ludwig S, Schuelper N, Brown J, Anders S, Raupach T (2018): How can we teach medical students to choose wisely? A randomised controlled cross-over study of video- versus text-based case scenarios. *BMC Med* 16, 107
- Lupien SJ, Maheu F, Tu M, Fiocco A, Schramek TE (2007): The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain Cogn* 65, 209-237
- Mansoor MS, Khazaei MR, Azizi SM, Niromand E (2021): Comparison of the effectiveness of lecture instruction and virtual reality-based serious gaming instruction on the medical students' learning outcome about approach to coma. *BMC Med Educ* 21, 347
- Marden NY, Ulman LG, Wilson FS, Velan GM (2013): Online feedback assessments in physiology: effects on students' learning experiences and outcomes. *Adv Physiol Educ* 37, 192-200
- Masys DR (1998): Advances in information technology. Implications for medical education. *West J Med* 168, 341-347
- McConnell MM, St-Onge C, Young ME (2015): The benefits of testing for learning on later performance. *Adv Health Sci Educ Theory Pract* 20, 305-320
- McCoy L, Lewis JH, Dalton D (2016): Gamification and Multimedia for Medical Education: A Landscape Review. *J Am Osteopath Assoc* 116, 22-34
- Middeke A, Anders S, Schuelper M, Raupach T, Schuelper N (2018): Training of clinical reasoning with a Serious Game versus small-group problem-based learning: A prospective study. *PLoS One* 13, e0203851
- Middendorff E, Apolinarski B, Poskowsky J, Kandulla M, Netz N: Die wirtschaftliche und soziale Lage der Studierenden in Deutschland 2012. 20. Sozialerhebung des Deutschen Studentenwerks durchgeführt durch das HIS-Institut für Hochschulforschung; Bundesministerium für Bildung und Forschung, Berlin 2013
- Nathaniel TI, Goodwin RL, Fowler L, McPhail B, Black AC, Jr. (2021): An Adaptive Blended Learning Model for the Implementation of an Integrated Medical Neuroscience Course During the Covid-19 Pandemic. *Anat Sci Educ* 14, 699-710
- Niamtu J, 3rd (2001): The power of PowerPoint. *Plast Reconstr Surg* 108, 466-484
- O'Flaherty J, Phillips C (2015): The use of flipped classrooms in higher education: A scoping review. *Internet High Educ* 25, 85-95

- Olmes GL, Zimmermann JSM, Stotz L, Takacs FZ, Hamza A, Radosa MP, Findeklee S, Solomayer EF, Radosa JC (2021): Students' attitudes toward digital learning during the COVID-19 pandemic: a survey conducted following an online course in gynecology and obstetrics. *Arch Gynecol Obstet* 304, 957-963
- Oussi N, Renman P, Georgiou K, Enochsson L (2021): Baseline characteristics in laparoscopic simulator performance: The impact of personal computer (PC)-gaming experience and visuospatial ability. *Surg Open Sci* 4, 19-25
- Patel PM, Tsui CL, Varma A, Levitt J (2020): Remote learning for medical student-level dermatology during the COVID-19 pandemic. *J Am Acad Dermatol* 83, e469-e470
- Peine A, Kabino K, Spreckelsen C (2016): Self-directed learning can outperform direct instruction in the course of a modern German medical curriculum - results of a mixed methods trial. *BMC Med Educ* 16, 158
- Puljak L, Čivljak M, Haramina A, Mališa S, Čavić D, Klinec D, Aranza D, Mesarić J, Skitarelić N, Zoranić S, et al. (2020): Attitudes and concerns of undergraduate university health sciences students in Croatia regarding complete switch to e-learning during COVID-19 pandemic: a survey. *BMC Med Educ* 20, 416
- Racine N, Hetherington E, McArthur BA, McDonald S, Edwards S, Tough S, Madigan S (2021): Maternal depressive and anxiety symptoms before and during the COVID-19 pandemic in Canada: a longitudinal analysis. *Lancet Psychiatry* 8, 405-415
- Raupach T, Andresen JC, Meyer K, Strobel L, Koziolok M, Jung W, Brown J, Anders S (2016): Test-enhanced learning of clinical reasoning: a crossover randomised trial. *Med Educ* 50, 711-720
- Redinger JW, Cornia PB, Albert TJ (2020): Teaching During a Pandemic. *J Grad Med Educ* 12, 403-405
- Robinson LJ, Engelson BJ, Hayes SN (2021): Who Is Caring for Health Care Workers' Families Amid COVID-19? *Acad Med* 96, 1254-1258
- Roediger HL 3rd, Karpicke JD (2006): Test-enhanced learning: Taking memory tests improves long-term retention. *Psychol Sci* 17, 249-255
- Röhle A, Horneff H, Willemer MC (2021): Practical teaching in undergraduate human and dental medical training during the COVID-19 crisis. Report on the COVID-19-related transformation of peer-based teaching in the Skills Lab using an Inverted Classroom Model. *GMS J Med Educ* 38, Doc2
- Rose S (2020): Medical Student Education in the Time of COVID-19. *JAMA* 323, 2131-2132

- Ruiz JG, Mintzer MJ, Leipzig RM (2006): The impact of E-learning in medical education. *Acad Med* 81, 207-212
- Ruiz-Grao MC, Cebada-Sánchez S, Ortega-Martínez C, Alfaro-Espín A, Candel-Parra E, García-Alcaraz F, Molina-Alarcón M, Delicado-Useros V (2022): Nursing Student Satisfaction with the Teaching Methodology Followed during the COVID-19 Pandemic. *Healthcare (Basel)* 10, 597
- Rutledge C, Walsh CM, Swinger N, Auerbach M, Castro D, Dewan M, Khattab M, Rake A, Harwayne-Gidansky I, Raymond TT, et al. (2018): Gamification in Action: Theoretical and Practical Considerations for Medical Educators. *Acad Med* 93, 1014-1020
- Sauer N, Salek A, Szlasa W, Ciecieląg T, Obara J, Gawel S, Marciniak D, Karłowicz-Bodalska K (2022): The Impact of COVID-19 on the Mental Well-Being of College Students. *Int J Environ Res Public Health* 19, 5089
- Schiebinger L, Gilmartin SK (2010): Housework Is an Academic Issue. *Academe* 96, 39-44
- Schittek Janda M, Tani Botticelli A, Mattheos N, Nebel D, Wagner A, Nattestad A, Attström R (2005): Computer-mediated instructional video: a randomised controlled trial comparing a sequential and a segmented instructional video in surgical hand wash. *Eur J Dent Educ* 9, 53-58
- Scholkmann A, Siemon J, Boom K-D, Knigge M (2017): Time on task during game-based learning. The effects of cognitive skills, goal orientations and the characteristics of learning partners – an analysis based on video data. *Z Erziehungswiss* 20, 651-669
- Seifert T, Becker T, Büttcher AF, Herwig N, Raupach T (2021): Restructuring the clinical curriculum at University Medical Center Göttingen: effects of distance teaching on students' satisfaction and learning outcome. *GMS J Med Educ* 38, Doc1
- Servin-Rojas M, Olivas-Martínez A, Dithurbide-Hernandez M, Chavez-Vela J, Petricevich VL, García-Juárez I, Gallo de Moraes A, Zendejas B (2022): Impact of the COVID-19 pandemic on the clinical training of last year medical students in Mexico: a cross-sectional nationwide study. *BMC Med Educ* 22, 24
- Shahrivini B, Baxter SL, Coffey CS, MacDonald BV, Lander L (2021): Pre-clinical remote undergraduate medical education during the COVID-19 pandemic: a survey study. *BMC Med Educ* 21, 13
- Simmenroth-Nayda A, Meskauskas E, Burckhardt G, Görlich Y (2014): [Medical school admission test at the University of Goettingen - which applicants will benefit?]. *Z Evid Fortbild Qual Gesundheitswes* 108, 609-617

- Son C, Hegde S, Smith A, Wang X, Sasangohar F (2020): Effects of COVID-19 on College Students' Mental Health in the United States: Interview Survey Study. *J Med Internet Res* 22, e21279
- Tarpada SP, Morris MT, Burton DA (2016): E-learning in orthopedic surgery training: A systematic review. *J Orthop* 13, 425-430
- Taveira-Gomes T, Ferreira P, Taveira-Gomes I, Severo M, Ferreira MA (2016): What Are We Looking for in Computer-Based Learning Interventions in Medical Education? A Systematic Review. *J Med Internet Res* 18, e204
- Tolks D, Schafer C, Raupach T, Kruse L, Sarikas A, Gerhardt-Szep S, Kllauer G, Lemos M, Fischer MR, Eichner B, et al. (2016): An Introduction to the Inverted/Flipped Classroom Model in Education and Advanced Training in Medicine and in the Healthcare Professions. *GMS J Med Educ* 33, Doc46
- Tsurugano S, Nishikitani M, Inoue M, Yano E (2021): Impact of the COVID-19 pandemic on working students: Results from the Labour Force Survey and the student lifestyle survey. *J Occup Health* 63, e12209
- Tubelo RA, Portella FF, Gelain MA, de Oliveira MMC, de Oliveira AEF, Dahmer A, Pinto MEB (2019): Serious game is an effective learning method for primary health care education of medical students: A randomized controlled trial. *Int J Med Inform* 130, 103944
- van der Keylen P, Lippert N, Kunisch R, Kühlein T, Roos M (2020): Asynchronous, digital teaching in times of COVID-19: a teaching example from general practice. *GMS J Med Educ* 37, Doc98
- Wahlgren CF, Edelbring S, Fors U, Hindbeck H, Ståhle M (2006): Evaluation of an interactive case simulation system in dermatology and venereology for medical students. *BMC Med Educ* 6, 40
- Walsh K (2015): Mobile Learning in Medical Education: Review. *Ethiop J Health Sci* 25, 363-366
- Wang X, Hegde S, Son C, Keller B, Smith A, Sasangohar F (2020): Investigating Mental Health of US College Students During the COVID-19 Pandemic: Cross-Sectional Survey Study. *J Med Internet Res* 22, e22817
- Weissmann Y, Useini M, Goldhahn J (2021): COVID-19 as a chance for hybrid teaching concepts. *GMS J Med Educ* 38, Doc12
- Wiedenmann C, Wacker K, Bohringer D, Maier P, Reinhard T (2021): [Online examination course instead of classroom teaching: adaptation of medical student teaching during the COVID-19 pandemic]. *Ophthalmologie* 119, 11-18

- Woitowich NC, Jain S, Arora VM, Joffe H (2021): COVID-19 Threatens Progress Toward Gender Equity Within Academic Medicine. *Acad Med* 96, 813-816
- Wu Q, Wang Y, Lu L, Chen Y, Long H, Wang J (2022): Virtual Simulation in Undergraduate Medical Education: A Scoping Review of Recent Practice. *Front Med (Lausanne)* 9, 855403
- Yang T, Buck S, Evans L, Auerbach M (2021): A Telesimulation Elective to Provide Medical Students With Pediatric Patient Care Experiences During the COVID Pandemic. *Pediatr Emerg Care* 37, 119-122
- Yapjakis C (2009): Hippocrates of Kos, the father of clinical medicine, and Asclepiades of Bithynia, the father of molecular medicine. Review. *In Vivo* 23, 507-514
- Yeh DD, Park YS (2015): Improving Learning Efficiency of Factual Knowledge in Medical Education. *J Surg Educ* 72, 882-889
- Zawacki-Richter O (2021): The current state and impact of Covid-19 on digital higher education in Germany. *Hum Behav Emerg Technol* 3, 218-226
- Zis P, Artemiadis A, Bargiotas P, Nteveros A, Hadjigeorgiou GM (2021): Medical Studies during the COVID-19 Pandemic: The Impact of Digital Learning on Medical Students' Burnout and Mental Health. *Int J Environ Res Public Health* 18, 349
- Zolotov Y, Reznik A, Bender S, Isralowitz R (2020): COVID-19 Fear, Mental Health, and Substance Use Among Israeli University Students. *Int J Ment Health Addict* 20, 230-236
- Zuo L, Dillman D, Miller Juvé A (2020): Learning at home during COVID-19: A multi-institutional virtual learning collaboration. *Med Educ* 54, 664-665

Danksagung

An dieser Stelle möchte ich mich bei denjenigen Personen bedanken, die mir die Anfertigung der Dissertation ermöglichten und mich in dem Vorhaben stetig unterstützten.

Mein besonderer Dank gilt Herrn Prof. Dr. med. Tobias Raupach für die Stellung des Themas, die ausgezeichnete Betreuung und die stetige Bereitschaft zum Diskurs. Auch für das zeitintensive Korrekturlesen und die konstruktiven Anmerkungen möchte ich mich herzlich bedanken.

Ebenso möchte ich Frau PD. Dr. med. dent. Sabine Sennhenn-Kirchner meinen Dank aussprechen für die stetige Betreuung und Unterstützung bis zur letztlich Vollendung der Arbeit.