

Institut für Didaktik und Ausbildungsforschung in der Medizin  
Klinik der Universität München  
Leitung: Prof. Dr. med. Martin Fischer, MME (Bern)



***Entwicklung und Validierung eines Situational  
Judgement Tests zum Umgang mit Emotionen  
für Medizinstudierende***

Dissertation  
zum Erwerb des Doktorgrades der Humanbiologie  
an der Medizinischen Fakultät der  
Ludwig-Maximilians-Universität zu München

vorgelegt von  
Tanja Graupe (geb. Pander)

aus  
München

2023

---

Mit Genehmigung der Medizinischen Fakultät der  
Universität München

Berichterstatter:	Prof. Dr. Martin Fischer
Mitberichterstatter:	PD Dr. Cornelius Schüle
	Prof. Dr. Claudia Bausewein
	Prof. Dr. Frank Padberg
Mitbetreuung durch den promovierten Mitarbeiter:	Prof. Dr. Jan-Willem Strijbos
	Prof. Dr. Claudia Kiessling
Dekan:	Prof. Dr. Thomas Gudermann
Tag der mündlichen Prüfung:	19.07.2023

## Affidavit



### Eidesstattliche Versicherung

Graupe, Tanja

Name, Vorname

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel:

Entwicklung und Validierung eines Situational Judgement Tests zum Umgang mit Emotionen für  
Medizinstudierende

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

München, 03.08.2023

Tanja Graupe

Ort, Datum

Unterschrift Doktorandin bzw. Doktorand

# Inhaltsverzeichnis

<b>Affidavit</b> .....	<b>3</b>
<b>Inhaltsverzeichnis</b> .....	<b>4</b>
<b>Abkürzungsverzeichnis</b> .....	<b>5</b>
<b>Publikationsliste</b> .....	<b>6</b>
<b>1. Beitrag zu den Veröffentlichungen</b> .....	<b>8</b>
1.1 Beitrag zu Paper I .....	8
1.2 Beitrag zu Paper II.....	8
<b>2. Einleitung</b> .....	<b>9</b>
2.1 Übergeordnete Fragestellungen und Forschungsvorhaben .....	9
2.2 Theoretischer Überblick und Relevanz ärztlicher kommunikativer Kompetenzen und Empathiefähigkeit.....	10
2.3 Prüfung kommunikativer Kompetenzen und Empathiefähigkeit .....	12
2.4 Entwicklung und Validierung eines Situational Judgement Test zur Messung von Empathiefähigkeit (Paper I) .....	14
2.5 Einsatz verschiedener Prüfungsinstrumente zur Messung von Empathiefähigkeit und deren Zusammenhang (Paper II) .....	16
2.6 Empfehlungen für den curricularen Einsatz von Prüfungsinstrumenten zur Empathiefähigkeit.....	17
<b>3. Zusammenfassung</b> .....	<b>18</b>
<b>4. Abstract (English)</b> .....	<b>19</b>
<b>5. Paper I: Developing and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of Emotional Sequences (VR-CoDES)</b> .....	<b>20</b>
<b>6. Paper II: The role of the emotive, moral, and cognitive components for the prediction of medical students' empathic behavior in an Objective Structured Clinical Examination (OSCE)</b> .....	<b>21</b>
<b>7. Literaturverzeichnis</b> .....	<b>22</b>
<b>Danksagung</b> .....	<b>27</b>
<b>Lebenslauf</b> .....	<b>Fehler! Textmarke nicht definiert.</b>

## Abkürzungsverzeichnis

BGR	Berlin Global Rating
EP	Explicit – Provide space
EPAAc	Explicit – Provide space – Affect – Acknowledge
EPAEm	Explicit – Provide space – Affect – Empathy
EPCAc	Explicit – Provide space – Content Acknowledge
EPCEx	Explicit – Provide space – Content – Explore
ER	Explicit – Reduce space
ERla	Explicit – Reduce space – Information advise
ERPp	Explicit – Reduce space – Post-poning
IRI	Interpersonal Reactivity Index
JSPE	Jefferson Scale of Physician Empathy
LMU Munich	Ludwig-Maximilians-Universität München
NP	Non-explicit – Provide space
NPBc	Non-explicit – Provide space – Back channel
NPIm	Non-explicit – Provide space – Implicit empathy
NPSi	Non-explicit – Provide space – Silence
NR	Non-explicit – Reduce space
OPTION	Observing-Patient-Involvement Scale
OSCE	Objective Structured Clinical Examination
OSVE	Objective Structured Video Examinations
PPOS	Patient-Provider-Oriented Scale
PSS	Providing-Space-based-Score
RIAS	Roter Interaction Analysis System
SJT	Situational Judgment Test
SJT-ES	Situational Judgment Test - Expert-based score
SJT-TS	Situational Judgment Test - Theory-based score
VR-CoDES	Verona Coding Definitions for Emotion Sequences

## Publikationsliste

**Graupe T**, Giemsa P, Schaefer K, Fischer MR, **Strijbos JW**, Kiessling C (2022). The role of the emotive, moral, and cognitive components for the prediction of medical students' empathic behavior in an Objective Structured Clinical Examination (OSCE). *Patient Education and Counseling*, 105(10), 3103–3109. doi: 0.1016/j.pec.2022.06.017.

Vogel C, Kiessling C, Fischer MR, **Graupe T** (2022). Does a sense of gratitude protect against empathy loss in medical students? An exploratory study. *GMS Journal for Medical Education*, 39(3), 1–17.

Giemsa P, Wübbolding C, Fischer MR, **Graupe T**, Härtl A, Lenz C, Sanftenberg L, Schelling J, Schüttpelz-Brauns K, Kiessling C (2020). What works best in a general practice specific OSCE for medical students: Mini-CEX or content-related checklists? *Medical Teacher*, 42(5), 578–584. doi: 10.1080/0142159X.2020.1721449.

**Graupe T**, Fischer, MR, Strijbos JW, Kiessling C (2020). Development and Piloting of a Situational Judgement Test for Emotion-Handling Skills Using the Verona Coding Definitions of Emotional Sequences (VR-CoDES). *Patient Education and Counseling*, 103(9), 1839–1845. doi: 10.1016/j.pec.2020.04.001.

Ortwein H, Benz A, Carl P, Huwendiek S, **Pander T**, Kiessling C (2017). Applying the Verona coding definitions of emotional sequences (VR-CoDES) to code medical students' written responses to written case scenarios: Some methodological and practical considerations. *Patient Education and Counseling*, 100(2), 305–312. doi:10.1016/j.pec.2016.08.026.

### Weitere Publikationen

Rechtien L, Gradel M, Fischer MR, **Graupe T**, Dimitriadis K (2022). A Mixed Methods Assessment of the Management Role of Physicians. *Advances in Medical Education and Practice*, 13, 1003–1017. <https://doi.org/10.2147/AMEP.S370245>.

Leunig A, Winkler M, Gernert JA, **Graupe T** & Dimitriadis K (2021). Management lessons through an interactive online discussion about hospital management during the COVID-19 pandemic. *GMS Journal for Medical Education*, 38(1). doi: 10.3205/zma001421.

Zibold J, Gernert JA, Reik LJU, Keidel LM, **Graupe T** & Dimitriadis K (2021). Adaptations to mentoring and peer mentor training at the medical faculty during the COVID-19 pandemic. *GMS Journal for Medical Education*, 38(1). doi: 10.3205/zma001404.

Gernert JA, Zibold J, Reik LJU, **Graupe T** & Dimitriadis K (2020). Restructuring career counselling ventures of a mentoring program for medical students in the course of the COVID-19 pandemic. *GMS Journal for Medical Education*, 37(3). doi: 10.3205/zma001366.

Huber BC, Steffen J, Schlichtiger J, **Graupe T**, Deuster E, Strouvelle VP, Fischer MR, Massberg S, Brunner S (2020). Alteration of physical activity during COVID-19 pandemic lockdown in young adults. *Journal of Translational Medicine*, 18. doi: 10.1186/s12967-020-02591-7.

Thrien C, Fabry G, Härtl A, Kiessling C, **Graupe T**, Preusche I, Pruskil S, Schnabel KP, Sennekamp M, Rüttermann S, Wünsch A (2020). Feedback in medical education – a workshop report with practical examples and recommendations. *GMS Journal for Medical Education*, 37(5). doi: 10.3205/zma001339.

Gernert JA, Schaller J, **Graupe T**, Dimitriadis K (2019). 'Have a chat with...': an innovative format for mentoring of medical students. *Medical Education*, 53(5), 501–502. doi: <https://doi.org/10.1111/medu.13854>.

Nicolai L, Gradel M, Antón S, **Pander T**, Kalb A, Köhler L, Fischer MR, Dimitriadis K & von der Borch, P (2017). Die Doktabörse – ein innovatives Instrument zur Vermittlung von Promotionsarbeiten in der Medizin an der LMU München. *GMS Journal for Medical Education*, 34(3), Doc30.

Nicolai L, Schmidbauer M, Gradel M, Ferch S, Antón S, Hoppe B, **Pander T**, von der Borch P, Pinilla S, Fischer M, Dimitriadis K (2017). Facebook Groups as a Powerful and Dynamic Tool in Medical Education: Mixed-Method Study. *Journal of Medical Internet Research*, 19(12), e408. doi:10.2196/jmir.7990.

- Gradel M, Moder S, Nicolai L, **Pander T**, Hoppe B, Pinilla S, Von der Borch P, Fischer MR, Dimitriadis K (2016). Simulating the physician as healthcare manager: An innovative course to train for the manager role. *GMS Journal for Medical Education*, 33(3), Doc41. doi:10.3205/zma001040.
- Schäfer M, **Pander T**, Pinilla S, Fischer MR, von der Borch P, Dimitriadis K (2016). A prospective, randomised trial of different matching procedures for structured mentoring programmes in medical education. *Medical Teacher*, 1–9. doi:10.3109/0142159X.2015.1132834.
- Pinilla S, Nicolai L, Gradel M, **Pander T**, Fischer MR, von der Borch P, Dimitriadis K (2015). Undergraduate Medical Students Using Facebook as a Peer-Mentoring Platform: A Mixed-Methods Study. *JMIR Medical Education*, 1(2), e12. doi:10.2196/mededu.5063.
- Pinilla S, **Pander T**, von der Borch P, Fischer MR, Dimitriadis K(2015). 5 years of experience with a large-scale mentoring program for medical students. *GMS Zeitschrift für Medizinische Ausbildung*, 32(1). <https://doi.org/10.3205/zma000947>.
- Schäfer M, **Pander T**, Pinilla S, Fischer MR, von der Borch P, Dimitriadis K (2015). The Munich-Evaluation-of-Mentoring-Questionnaire (MEMeQ)—a novel instrument for evaluating protégés' satisfaction with mentoring relationships in medical education. *BMC Medical Education*, 15(1), 1. DOI 10.1186/s12909-015-0469-0.
- Steffen J, Grabbert M, **Pander T**, Gradel M, Köhler LM, Fischer MR, von der Borch P, Dimitriadis K (2015). Finding the right doctoral thesis—an innovative research fair for medical students. *GMS Zeitschrift für Medizinische Ausbildung*, 32(3). doi:10.3205/zma000971.
- Pander T**, Pinilla S, Dimitriadis K, Fischer MR (2014). Der Einsatz von Facebook in der medizinischen Ausbildung – Ein systematischer Literaturüberblick. *GMS Zeitschrift für Medizinische Ausbildung*, 31(3): Doc33.
- Welbergen L, Pinilla S, **Pander T**, Gradel M, von der Borch P, Fischer MR, Dimitriadis K (2014). Das FacharztDuell: Innovative Karriereplanung in der Medizin. *GMS Zeitschrift für Medizinische Ausbildung*, 31(2): Doc17.
- Welbergen L, von der Borch P, **Pander T**, Dimitriadis K (2013). Battle of the specialties: innovative career counselling for students. *Medical Education*, 47(5), 515–515. <https://doi.org/10.1111/medu.12159>.

# 1. Beitrag zu den Veröffentlichungen

## 1.1 Beitrag zu Paper I

*Graupe T, Fischer, MR, Strijbos JW, Kiessling C (2020). Development and Piloting of a Situational Judgement Test for Emotion-Handling Skills Using the Verona Coding Definitions of Emotional Sequences (VR-CoDES). Patient Education and Counseling, 103(9), 1839–1845.*

Konkretisierung Thema und Entwicklung Fragestellungen  
Konzeption Studiendesign  
Schreiben des Ethikantrages  
Durchführung qualitativer Interviews und Auswertung der qualitativen Daten  
Schreiben der Fallvignetten  
Koordination der Filmdreh-Tage  
Rekrutierung und Verwaltung der studentischen Stichprobe  
Durchführung und administrative Verwaltung der Studientage  
Archivierung und Dokumentation der Daten  
Auswertung der Daten  
Einpflegen des Situational Judgement Tests in CASUS  
Schulung der Experten  
Schreiben des Erstentwurfs des Artikels  
Koordination und Rücksprache mit den Co-Autoren  
Einreichen des Papers  
Bearbeiten der Reviewer-Kommentare  
Finalisierung des Papers

## 1.2 Beitrag zu Paper II

*Graupe T, Giemsa P, Schaefer K, Fischer MR, Strijbos JW, Kiessling C (2022). The role of the emotive, moral, and cognitive components for the prediction of medical students' empathic behavior in an Objective Structured Clinical Examination (OSCE). Patient Education and Counseling, 105(10), 3103–3109.*

Konkretisierung Thema und Entwicklung Fragestellungen  
Konzeption des Studiendesign  
Schreiben des Ethikantrages  
Durchführung und administrative Verwaltung der Studientage  
Archivierung und Dokumentation der Daten  
Auswertung der Daten  
Schreiben des Outlines  
Schreiben der Einleitung, des Methodenteils sowie der Diskussion des Papers  
Erstellung des Literaturverzeichnisses  
Überarbeitung des gesamten Papers  
Korrekturlesen des gesamten Papers  
Koordination und Rücksprache mit allen beteiligten Autoren  
Einreichen des Papers  
Bearbeiten der Reviewer-Kommentare  
Finalisierung des Papers



## 2. Einleitung

### 2.1 Übergeordnete Fragestellungen und Forschungsvorhaben

Die vorliegende Arbeit beschäftigt sich mit den übergeordneten Fragestellungen, wie ein *Situational Judgement Test* zur Prüfung des professionellen Umgangs mit Emotionen von Medizinstudierenden entwickelt und validiert werden kann und inwieweit dieses Prüfungsformat im Sinne einer konkurrenten Validität mit empathischem Verhalten korreliert.

In drei aufeinanderfolgenden Studien wurde ein computerbasierter *Situational Judgement Test* theoriegeleitet entwickelt, mit Medizinstudierenden der LMU München getestet und mit anderen Prüfungsformaten validiert.

**Studie 1:** Ziel dieser qualitativen Studie war die Identifikation typischer und realer Situationen aus dem praktischen Alltag einer Ärztin oder eines Arztes, in denen der Umgang mit Emotionen in der Kommunikation mit Patienten\*innen und Angehörigen eine besondere Rolle spielte. Hierfür wurden 16 halbstandardisierte Interviews nach der Methode der kritischen Ereignisse (Flanagan, 1954) mit praktisch tätigen Ärzt\*innen geführt. Die so gewonnenen authentischen Situationen wurden im Anschluss theoriegeleitet und durch den Einsatz mehrerer Expertenpanels in 29 Fallvignetten umgewandelt und mit (Laien-)Schauspieler\*innen verfilmt. Diese Fallvignetten bildeten die Grundlage zur Entwicklung des *Situational Judgement Tests*.

**Studie 2:** Ziel dieser Studie war die Entwicklung und Validierung eines computerbasierten *Situational Judgement Tests* zum Umgang mit Emotionen für Medizinstudierende. Die verfilmten Fallvignetten aus Studie 1 wurden in die Lernplattform CASUS eingebettet und theoriebasierte Antwortalternativen für jede Vignette mit einem weiteren Expertenpanel entwickelt. Als Antwortformat wurde der Schieberegler ausgewählt und zwei verschiedene Scores zur Bewertung der studentischen Antworten entwickelt. Durch ein single-trait-multi-method-Design wurde im Rahmen einer experimentellen Studie an der LMU München mit 88 Medizinstudierenden der *Situational Judgement Test* pilotiert und validiert. Außerdem sollte die „beste“ Version des Scorings identifiziert werden.

**Studie 3:** Ziel dieser Studie war die Validierung und die Überprüfung, ob und in welchem Ausmaß der entwickelte *Situational Judgement Test* mit zwei weiteren Bewertungsinstrumenten, die empathisches Verhalten von Medizinstudierenden in realitätsnahen Situationen im Rahmen eines *Objektivierten Strukturierten Klinischen Examen* prüfen, korreliert. Es wurde ebenfalls überprüft ob und inwieweit der *Situational Judgement Test* mit der studentischen Selbsteinschätzung hinsichtlich Empathiefähigkeit korreliert.

Die Ergebnisse von Studie 1 und 2 wurden in Paper I und die Ergebnisse von Studie 3 wurden in Paper II veröffentlicht.

In den nachfolgenden Kapiteln soll ein Überblick über die der Arbeit zugrunde liegenden theoretischen Konstrukte gegeben, ein thematischer Bogen gespannt, die Ergebnisse der durchgeführten Studien beschrieben, sowie Empfehlungen zum zukünftigen Einsatz der Studienergebnisse ausgesprochen werden.

## 2.2 Theoretischer Überblick und Relevanz ärztlicher kommunikativer Kompetenzen und Empathiefähigkeit

In den letzten Jahren ist die Bedeutung kommunikativer Kompetenzen für den ärztlichen Beruf mehr und mehr in den Vordergrund gerückt. Patientenbefragungen zeigen, dass diese sich Ärzt\*innen wünschen, die einfühlsam mit ihnen kommunizieren [1, 2]. Auch im Bereich der Patientensicherheit wird eine gelungene Kommunikation zwischen Ärzt\*innen und Patient\*innen als wichtiger positiver Einflussfaktor zur Fehlerverhinderung anerkannt [3, 4, 5]. Für Ärzt\*innen selbst bieten gute kommunikative Kompetenzen ebenfalls Vorteile, denn sie können sich so beispielsweise vor Burnout schützen und ihre persönliche Berufszufriedenheit steigern [6, 7].

Darüber hinaus verbessert eine gelungene klinische Kommunikation die Qualität der medizinischen Versorgung und sorgt beim medizinischen Kontakt zwischen Patient\*innen und Ärzt\*innen für effektivere Resultate, zum Beispiel in Bezug auf das Sammeln von Patientendaten, dem Verstehen der Patientenperspektiven, dem Bereitstellen von Informationen und Beratung, dem Treffen von Behandlungsentscheidungen und dem Umgang mit den Emotionen von Patient\*innen [8]. Im Rahmen dieser Arbeit soll das Hauptaugenmerk auf Letzteres – dem gelungenen Umgang mit Emotionen – gelegt werden. Denn diese so genannte Empathiefähigkeit stellt eine Schlüsselkomponente der ärztlichen professionellen kommunikativen Kompetenzen dar [9, 10]. Empathische Fähigkeiten bei Ärzt\*innen führen zu einer patientenzentrierten Medizin und erwiesenermaßen zu vielen behandlungstechnischen Vorteilen [11, 12, 13, 14, 15, 16, 17]. Ärzt\*innen identifizieren beispielsweise die Probleme der Patient\*innen sorgfältiger, genauer und schneller, da bei den Patient\*innen eine höhere Bereitschaft besteht sich zu öffnen und Informationen preiszugeben. Dadurch sind die Patient\*innen zufriedener, fühlen sich besser verstanden und stimmen der Behandlung häufiger zu [18, 19].

Obwohl die Rolle und Wichtigkeit der ärztlichen Empathiefähigkeit seit vielen Jahren untersucht und belegt werden, gibt es keine einheitliche Definition. Empathie ist ein multidimensionales Konstrukt und kann als eine Reihe professioneller Fähigkeiten oder Kompetenzen verstanden werden [20]. Eine häufig zitierte Definition ist die von Mercer und Reynolds [20]. Sie definieren die Empathiefähigkeit von Ärzt\*innen als die Fähigkeit die Situation, die Perspektive und die Gefühle der Patient\*innen (und die damit verbundenen Bedeutungen) zu verstehen, dieses Verständnis angemessen zu kommunizieren und im Umgang mit den Patient\*innen dementsprechend auf hilfreiche (therapeutische) Weise zu handeln [20]. Damit setzen sie einen Schwerpunkt auf kognitive und behaviorale empathische Fähigkeiten.

Eine erweiterte Definition bietet Morse, der vier Komponenten der Empathiefähigkeit beschreibt: emotional, moralisch, kognitiv, verhaltensorientiert [21]. Die emotionale Komponente umfasst die Fähigkeit, die Emotionen anderer zu beschreiben und zu teilen. Sie zeigt sich in der Wahrnehmung, emotional auf das emotionale Signal eines anderen zu reagieren [21]. Die moralische Komponente umfasst die innere altruistische Kraft, die das Praktizieren von Empathie motiviert und sich entweder in der Ausübung der eigenen Willenskraft, (...) sich auf einen anderen einzulassen oder in einer Haltung der Empfänglichkeit und Verfügbarkeit für andere manifestiert [21].

Die kognitive Komponente soll als die Fähigkeit verstanden werden, die Gefühle und Perspektiven eines anderen zu erkennen und zu verstehen und äußert sich in der Argumentation, Analyse oder kritischen Überprüfung von Prozessen oder in der bewussten kognitiven Verarbeitung der Situation [21]. Die verhaltensorientierte Komponente von Empathie ist die kommunikative Reaktion, um Verständnis für die Perspektive eines anderen zu vermitteln und zeigt sich in einem beobachtbaren empathischen Verhalten [21].

Obwohl ärztliche Empathie, wie oben beschrieben, positive Auswirkungen auf medizinische Begegnungen haben kann [11, 12, 13, 14, 15, 16, 17] nutzen Ärzt\*innen nur 70–90% der Möglichkeiten, empathisch zu handeln (*empathic opportunities*) [22]. Ein möglicher Grund dafür könnte sein, dass sie die Emotionen der Patient\*innen nicht oder nicht angemessen erkennen und einschätzen können [23]. Patient\*innen drücken ihre Emotionen meist durch direkte oder indirekte Hinweise auf ein zugrundeliegendes Gefühl aus [24]. Ärzt\*innen haben dann die Möglichkeit zu entscheiden, diese indirekten und direkten Signale aufzunehmen und den Raum für Emotionen für die Patient\*innen mit verschiedenen kommunikativen Techniken zu öffnen oder zu schließen. Sowohl die patientenseitigen indirekten und direkten Hinweise auf Emotionen als auch die arztseitigen kommunikativen Techniken oder Reaktionen wurden in einem jahrelangen Konsentierungsprozess in den *Verona Coding Definitions for Emotional Sequences (VR-CoDES)* operationalisiert. Dieses theoretische Rahmenmodell soll als Grundlage dieser Arbeit dienen [24]. Laut *VR-CoDES* äußern Patient\*innen ihre Emotionen meist durch direkte oder indirekte Hinweise – so genannte *Cues* und *Concerns*. Ein *Concern* ist ein klarer und eindeutiger Ausdruck einer unangenehmen aktuellen oder kürzlich empfundenen Emotion, wobei das Gefühl oder die Sorge ausdrücklich verbalisiert wird. Ein Beispiel ist folgende Aussage eines Patienten: „Jetzt sind die Kopfschmerzen nicht mehr so schlimm...aber ich mache mir trotzdem Sorgen über die Ergebnisse des Tests“ [24].

Ein *Cue* ist ein verbaler oder nonverbaler indirekter Hinweis, der ein zu Grunde liegendes unangenehmes Gefühl andeutet [24]. Ein Beispiel ist folgende Aussage einer Patientin: „Wissen Sie, ich kann mich einfach nicht entspannen... (Pause) ...kann mich einfach nicht entspannen“ [24].

In einem weiteren Schritt bietet *VR-CoDES* ein Schema zur Kategorisierung der ärztlichen Reaktionen auf diese *Cues* und *Concerns*. Die Antworten der Ärzt\*innen werden in die Kategorien "explizit" und "nicht explizit" und in die Kategorien "raumöffnend" oder "raumschließend" unterteilt. Um als explizit kodiert zu werden, sollte eine Antwort entweder einen spezifischen oder einen direkten Verweis auf die Wörter im vorangehenden Hinweis oder Anliegen des Patienten oder der Patientin enthalten oder sehr klar sein, indem sie sich eindeutig auf den gegebenen Hinweis oder das erwähnte Anliegen bezieht. Eine nicht explizite Reaktion ist jede Reaktion der Ärztin oder des Arztes, die weder den Inhalt noch die Emotion des Hinweises der Patientin oder des Patienten ausdrücklich erwähnt oder mehrdeutig ist. Nonverbale Antworten der Ärztin oder des Arztes, die nicht von einer verbalen Antwort begleitet werden, werden immer als nicht explizit kodiert, zum Beispiel Nicken oder eine Handbewegung [25].

Antworten, die den Raum öffnen, laden die Patientin oder den Patienten dazu ein, mehr über ihre oder seine *Cues* und *Concerns* zu erzählen. Beispielhafte raumöffnende Strategien für Ärzt\*innen

sind Anerkennung, Explorieren und Empathie. Antworten werden als raumschließend eingestuft, wenn die Ärztin oder der Arzt beispielsweise den Hinweis der Patientin oder des Patienten ignoriert, einen Themenwechsel einführt, eine Diskussion der Sorge aufschiebt, aktiv blockiert oder ein anderes Verhalten anbietet, das die Gelegenheit für die Patient\*innen minimiert, mehr über ihre Sorgen zu erzählen [25]. Insgesamt bietet das Rahmenmodell 17 mögliche Strategien für raumöffnende oder raumschließende Verhaltensweisen [25].

Die Reaktion der Ärzt\*innen auf die *Cues* und *Concerns* der Patient\*innen beeinflusst das Ergebnis des medizinischen Zusammentreffens. Das Öffnen des Raumes für die Emotionen der Patient\*innen fördert die weitere Diskussion der Emotion, stärkt die therapeutische Allianz, begünstigt die bessere Bewältigung von Krankheiten, verringert Angstzustände und erhöht die Zufriedenheit der Patient\*innen, und auch die der Ärzt\*innen [26].

Eide et al. (2011) überprüften die Validität von *VR-CoDES* für das Erkennen von emotionalen Hinweisen von Patient\*innen. Sie empfehlen, dieses Rahmenmodell als Trainingsinstrument zu verwenden, um die empathische Genauigkeit der Ärzt\*innen zu fördern [27]. *VR-CoDES* können auch bereits für Medizinstudierende im Rahmen eines entsprechenden longitudinalen Curriculums eingesetzt werden, um deren empathische Reaktionsfähigkeit zu trainieren, zu prüfen und zu fördern.

### **2.3 Prüfung kommunikativer Kompetenzen und Empathiefähigkeit**

Viele internationale und nationale Organisationen zur medizinischen Ausbildung bezeichnen in ihren Lernzielkatalogen kommunikative Kompetenzen, wie Empathiefähigkeit, als Teil der medizinischen Expertise [28, 19, 30, 31, 31]. Mit der Novellierung der ärztlichen Approbationsordnung von 2012 wurde die ärztliche Gesprächsführung auch als Prüfungsinhalt in den zweiten Abschnitt der ärztlichen Prüfung aufgenommen, so dass auch in Deutschland von staatlicher Seite das Signal gesetzt wurde, dass diese Fähigkeit zu den zentralen ärztlichen Kompetenzen gehört [33]. Medizinische Curricula müssen ihre Lehr- und Lernstrategien ebenfalls dahingehend anpassen: Die Ausbildung und Förderung kommunikativer Kompetenzen, wie Empathiefähigkeit, müssen ein anerkannter Teil der medizinischen Ausbildung werden [34, 35].

An der Ludwig-Maximilians-Universität München (LMU) werden diese Kompetenzen bereits in zahlreichen Veranstaltungen vermittelt und trainiert. Doch fehlt es – gerade im deutschsprachigen Raum – an geeigneten Prüfungsformaten, um den Studierenden eine strukturierte Rückmeldung über ihren Lernfortschritt in diesem Bereich zu geben. Eine Prüfung im universitären Kontext kann formativ und summativ ablaufen. Die formative Prüfung hat es zum Hauptziel Studierenden eine direkte Rückmeldung über ihr Handeln oder Wissen zu geben und deren Lernprozess dadurch zu unterstützen. Formatives Feedback soll in den curricularen Ablauf integriert werden und regelmäßig stattfinden, damit die Studierenden die Möglichkeit haben die Ergebnisse zu verarbeiten, konstruktiv umzusetzen und so die Qualität ihrer Arbeit durch wiederholtes Beobachten, Reflektieren und Anwenden zu verbessern [36, 37]. Dies entspricht der Theorie der *deliberate practice* von Ericsson (2008), bei welcher davon ausgegangen wird, dass unmittelbares Feedback und

Wiederholung das Verhalten des Lernenden verändern und so kontinuierliches Lernen gesichert wird [38]. Die summative Prüfung findet am Ende einer Lehreinheit statt und entscheidet darüber, ob die Leistung der Studierenden genügt, um beispielsweise in einen nächsten Studienabschnitt oder in ein höheres Semester überzuwechseln. Alle Prüfungsformate – ob formativ oder summativ – müssen den allgemeinen Testgütekriterien gerecht werden: Die Leistung der Studierenden muss fair und objektiv geprüft werden, die Fragen müssen reliabel und valide sein und das Wissen oder Verhalten abfragen und erzeugen, welches am Ende einer Lerneinheit von den Studierenden (von den Lehrenden oder der Fakultät) erwartet wird [39].

In Bezug auf Prüfungsformate der ärztlichen Empathiefähigkeit soll hier die Klassifikation von Hemmerdinger (2007) als Grundlage dienen. Er unterteilt diese Prüfungsformate in 1) Beurteilungen durch die erste Person (*first person assessment*), 2) Beurteilung durch die zweite Person (*second person assessment*) und 3) Beurteilung durch die dritte Person (*third person assessment*) [40].

Die Beurteilung durch die erste Person umfasst standardisierte Selbstbewertungsinstrumente, meist in Form von Fragebögen, welche Selbstwahrnehmung, Selbstvertrauen oder Einstellungen der Befragten messen [40]. Prominente Beispiele im Bereich der ärztlichen Empathiefähigkeit sind der *Interpersonal Reactivity Index* (IRI), der die selbst eingeschätzten kognitiven Fähigkeiten und die emotionale Reaktivität von Personen misst [41] und die *Jefferson Scale of Physician Empathy* (JSPE), die entwickelt wurde, um die Orientierungen und Einstellungen von im Gesundheitswesen Tätigen oder Medizinstudierenden in Bezug auf Empathie im Rahmen der Patientenversorgung zu messen [42].

Die Beurteilung durch die zweite Person umfasst Fragebögen, die von Patient\*innen beantwortet werden, um das empathische Verhalten der ihnen zugeteilten Ärzt\*innen im Rahmen ihrer individuellen Behandlung zu bewerten. Ein prominentes Beispiel ist das *CARE*-Instrument [43].

Die Beurteilung durch die dritte Person umfasst standardisierte Instrumente, die von Beobachtern verwendet werden, um das Verhalten der Lernenden in realen oder simulierten klinischen Szenarien zu bewerten. Empathiefähigkeit in realen klinischen Situationen kann durch direkte Beobachtung oder durch die Analyse von Video- oder Audioaufnahmen bewertet werden. Hierbei können Kodierungsinstrumente für die Sequenzanalyse verwendet werden, wie das *Roter Interaction Analysis System* (RIAS), welches aber ein allgemeineres Instrument und weniger für die Beobachtung emphatischer Interaktionen geeignet ist [52]. Das bereits oben beschriebene Verona Coding System (VR-CoDES) [24] kann hier ebenfalls eingesetzt werden.

Im Medizinstudium hat sich als angemessene Methode für die Überprüfung kommunikativer Kompetenzen in den letzten Jahren das *Objektivierte Strukturierte Klinische Examen* (OSCE) etabliert. Studierende durchlaufen kontextreiche, realistische Szenarien – so genannte Stationen – aus der ärztlichen Praxis, welche meist mit Schauspielpatient\*innen nachgespielt werden. Die Studierenden werden dabei von geschulten Prüfer\*innen nach bestimmten Qualitätskriterien beurteilt [44, 45]. Zur Messung der Kommunikationsfähigkeiten in einem OSCE wurden bereits zahlreiche Instrumente entwickelt, entweder in Form von stationsbezogenen Checklisten oder globalen Ratings [46, 47, 48, 49, 50], wie dem *Berlin Global Rating* (BGR) [51].

Neben vielen Vorteilen des OSCEs ist ein zentraler Nachteil der hohe Ressourcenaufwand für die Vorbereitung und Durchführung der Prüfung. Ein kontinuierlicher Einsatz, auch zu formativen Zwecken, ist aufgrund finanzieller, zeitlicher und personeller Beschränkungen für viele Fakultäten mit hohen Studierendenzahlen daher kaum leistbar.

Bislang enthält die Klassifikation von Hemmerdinger (2007) für die Überprüfung von Empathie keine Bewertung mit schriftlichen oder videobasierten Tests. Eine bisher offene Forschungsfrage ist, inwieweit schriftliche oder computerbasierte Prüfungsformate komplementär zu einem OSCE genutzt werden können. Gekoppelt damit ist die grundlegendere Frage, welche Rolle kognitive Aspekte für eine gelungene ärztliche Kommunikation spielen. Grundsätzlich lässt sich kognitives Wissen – Faktenwissen, prozedurales und konzeptionelles Wissen – mit schriftlichen und computerbasierten Tests messen [39]. Diese Tests können aufgrund des Kosten-Nutzen-Verhältnisses sowie ihres Kontextreichtums und der Möglichkeit den Kompetenzerwerb gezielt zu fördern, zum Beispiel durch die Übung in einer kontrollierten Umgebung ohne mögliche Stör-Variablen oder durch die stufenweise und kontrollierte Zunahme der Komplexität oder Intensität der Erfahrung, eine effektive Alternative sein.

In den letzten Jahren wurden verschiedene schriftliche Prüfungsformate entwickelt, um über reines Faktenwissen hinaus klinisches Anwendungswissen zu prüfen.

Der *Skript Konkordanz Test* [55] eignet sich gut, um klinisch orientiertes logisches Denken (*clinical reasoning*) zu prüfen. Wesentlich ist hier vor allem, dass sich die Lernenden in klinisch relevante Alltagssituationen eindenken müssen. Hier bekommen sie nachfolgend wichtige Informationen und müssen daraufhin jeweils neue Entscheidungen treffen. Die Antworten werden nicht klassisch im Sinne einer Einfachauswahl (richtig/falsch) bewertet, sondern anhand eines Vergleichs mit den Antworten eines Experten-Panels von Kliniker\*innen.

*Key Feature Tests* konzentrieren sich meist auf einen schwierigen Aspekt in der medizinischen Interaktion, bei dem die Lernenden am ehesten Fehler machen (könnten). Ein *Key Feature* ist definiert als ein wichtiger Schritt bei der Lösung eben dieses schwierigen Aspekts. Dieser Prüfungsansatz ist grundsätzlich nicht durch ein bestimmtes Format definiert, wird jedoch meistens im Rahmen schriftlicher Test verwendet. Hierbei werden zu einer klinisch relevanten (Problem-) Situation aufbauend mehrere Fragen gestellt. Nach der Beantwortung einer Frage bekommen die Lernenden die korrekte Antwort präsentiert, können aber ihre eigene Antwort nicht mehr verändern, aber das neu erworbene Wissen zur Beantwortung der Folgefrage verwenden [56].

*Situational Judgement Tests (SJT)* stellen eine weitere Form dieses Prüfungsansatzes dar.

## **2.4 Entwicklung und Validierung eines Situational Judgement Test zur Messung von Empathiefähigkeit (Paper I)**

*Situational Judgement Tests (SJT)* stellen eine innovative Methode dar „nicht-akademische“ Fähigkeiten, wie Empathiefähigkeit, bei Medizinstudierenden zu beurteilen. In einem *SJT* werden die Teilnehmer\*innen mit schriftlichen oder videobasierten hypothetischen arbeitsbezogenen

Szenarien konfrontiert und aufgefordert, alternative Reaktionen innerhalb dieser Szenarien zu bewerten [57]. Antworten können wissensbasiert oder verhaltensbasiert sein [58, 59] und können von Einfachwahl zu Mehrfachantworten, bis hin zu Ranking-Antwortformaten variieren [60, 61]. *SJTs* vergleichen die Antworten der Lernenden mit den Ergebnissen eines Expertenpanels. *SJTs* basieren auf der Verhaltenskonsistenztheorie: Voraussichtliches Verhalten kann zukünftiges Verhalten vorhersagen [62]. Es gibt auch zunehmend Belege dafür, dass Lernende während *SJTs* Überzeugungen über die Wirksamkeit verschiedener Verhaltensweisen entwickeln [63]. Schließlich scheinen *SJTs* in der Praxis wirksame Prädiktoren für das tatsächlich gezeigte Verhalten der Lernenden in Realsituation zu sein [64, 65, 66].

*SJTs* im medizinischen Kontext weisen ein moderates bis gutes Maß an Zuverlässigkeit [57, 64, 67, 68, 69, 70] sowie ein hohes Maß an prädiktiver Validität im Gesundheitswesen auf [60, 61, 64, 70, 71]. *SJTs* haben im Vergleich zu anderen Prüfungsinstrumenten, wie kognitiven Wissens-tests, weniger nachteilige Auswirkungen auf ethnische Zugehörigkeit und Geschlecht [72] und die Reaktionen der Teilnehmer\*innen auf *SJTs* sind positiv [68, 70, 73, 74].

Videobasierte *SJTs* rufen günstigere Reaktionen der Lernenden hervor und weisen einen mittleren Grad an Wiedergabetreue auf als textbasierte *SJTs* mit geringer Wiedergabetreue [70]. Die anfänglichen Entwicklungskosten für videobasierte *SJTs* sind im Vergleich zu Fragebögen und *OSCEs* höher. Da sie jedoch langfristig ohne wiederkehrenden Einsatz von Schauspielpatient\*innen funktionieren und leicht wiederverwendet werden können, sinken die Kosten im Laufe der Zeit [75].

In Paper I wurde die Entwicklung und Testung eines *SJTs* zum Umgang mit Emotionen dargestellt und diskutiert (Studien 1 und 2). Die wichtigsten Ergebnisse in Bezug auf das Antwortverhalten der Teilnehmenden hierbei waren, dass in dem von uns entwickelten *SJT* auf Grundlage der *VR-CoDES*, die Expert\*innen und Studierenden, welche den Test im Rahmen unserer experimentellen Studie absolvierten, meistens raumschließende Antworten als unangemessen bewerteten und explizite Antworten bevorzugten. Frauen schnitten besser ab als Männer, und es gab keinen Rückgang der Empathiefähigkeit in Abhängigkeit vom Studienjahr der Studierenden. Es gab mittlere Korrelationen mit den eingesetzten Selbstbewertungsinstrumenten. Die Akzeptanz des *SJTs* bei den Studierenden war hoch.

Der Einsatz von *VR-CoDES* als theoretisches Rahmenmodell sowie die Verwendung von authentischen Vignetten, Videos und Expertenpanels trugen positiv zur Entwicklung des *SJTs* bei. Die Entwicklungskosten waren hoch, können aber im Laufe der Zeit relativiert werden. Das Festlegen eines angemessenen Scores zur Bewertung sowie das Anbieten einer geeigneten Feedback-Struktur scheinen sinnvoll und wichtig zu sein.

## 2.5 Einsatz verschiedener Prüfungsinstrumente zur Messung von Empathiefähigkeit und deren Zusammenhang (Paper II)

Wie bereits oben beschrieben, setzt sich ärztliche Empathiefähigkeit nach Morse (1992) aus den folgenden vier Komponenten zusammen: emotional, moralisch, kognitiv, verhaltensorientiert [21]. Unserer Kenntnis nach gibt es kaum Belege dafür, inwieweit die emotionalen, moralischen und kognitiven Komponenten von Empathie, empathisches Verhalten in realen Situationen vorhersagen. Verschiedene Prüfungsformate messen unterschiedliche Komponenten von Empathiefähigkeit. Der *Interpersonal Reactivity Index (IRI)* kann dazu verwendet werden die kognitive und emotionale Komponente von Empathie zu messen [41], die *Jefferson Scale of Physician Empathy (JSPE)* misst die moralische Komponente [42] und *SJT*s messen prozedurales Wissen über Empathie, das nach Morse et al. mit der kognitiven Komponente von Empathie gleichgesetzt werden kann [21]. Empathisches Verhalten in simulierten und realen Situationen kann im Rahmen eines *OSCE*s gemessen werden.

Unser Ziel war es durch eine experimentelle Studie an der LMU München mit Medizinstudierenden, den Zusammenhang zwischen den oben genannten Komponenten nach Morse (1992) – also Wissen, Einstellungen und Verhalten im Umgang mit Emotionen – zu untersuchen. Diese Komponenten werden gemessen anhand der Ergebnisse eines selbst-entwickelten *SJT*s [76], der Selbstwahrnehmung eines Studierenden anhand von *IRI* [41] und *JSPE-S* [42] und einer Beobachterbewertung im Rahmen eines *OSCE*s [77]. Als Bewertungsinstrumente des studentischen Verhaltens im *OSCE* wurden das *BGR* [51] und *VR-CoDES* [24] herangezogen.

Paper II der vorliegenden Doktorarbeit zeigt und diskutiert die Ergebnisse dieser experimentellen Studie (Studie 3).

Die Ergebnisse zeigten, dass das Instrument zur Messung der emotionalen Komponente (*IRI*) nicht signifikant mit empathischem Verhalten im *OSCE* korrelierte, während die studentische Selbsteinschätzung der Empathiefähigkeit, gemessen mit der *JSPE-S* Skala, ebenso wie die Ergebnisse des *SJT*s mit dem empathischen Verhalten der Studierenden im *OSCE* signifikant korrelierte. Alle Korrelationskoeffizienten waren niedrig bis mittel. Die höchste Varianzaufklärung für empathisches Verhalten, gemessen durch *VR-CoDES*, wurde für den Experten-basierten Score *SJT-ES* ( $R^2 = 0,125$ ) und den theoriebasierten Score *SJT-TS* ( $R^2 = 0,131$ ) gefunden. *JSPE-S* ( $R^2 = 0,11$ ) und *SJT-ES* ( $R^2 = 0,10$ ) erklärten den größten Anteil der Varianz im empathischen Verhalten, gemessen durch *BGR*. Die schrittweise multiple Regression verbesserte das Modell für *BGR* durch Einbeziehung von *SJT-ES* und *JSPE-S* und erklärte 16,2 % der Varianz. So kann geschlussfolgert werden, dass das Instrument zur Messung der emotionalen Komponente (*IRI*) empathisches Verhalten nicht signifikant voraussagt, während die Instrumente zur Messung der moralischen (*JSPE-S*) und kognitiven Komponenten (*SJT*) empathisches Verhalten signifikant vorhersagen.



## 2.6 Empfehlungen für den curricularen Einsatz von Prüfungsinstrumenten zur Empathiefähigkeit

Die ärztliche Empathiefähigkeit ist ein multidimensionales Konstrukt mit unterschiedlichen Komponenten, die miteinander interkorrelieren und nicht getrennt voneinander auftreten. Dies muss im Rahmen eines Curriculums berücksichtigt werden. Kein einzelnes Prüfungsinstrument kann alle unterschiedlichen Facetten von Empathie bewerten. Daher schlagen wir ein longitudinales Prüfungskonzept. Hierbei können Informationen über die Lernenden im Sinne einer Triangulierung mit Hilfe verschiedener Instrumente gesammelt werden, um einen umfassenden Eindruck von ihren empathischen Einstellungen, ihrem Wissen und ihrem Verhalten zu erhalten und empathische Fähigkeiten so valide wie möglich vorherzusagen.

Es zeigt sich, dass sich *Situational Judgement Tests* gut eignen, um prozedurales Wissen über Empathie von Medizinstudierenden zu messen und daher einen berechtigten Platz in solch einem longitudinalen Prüfungskonzept haben. Bei der Entwicklung eines *SJT*s sollte man darauf achten ein theoriegeleitetes Rahmenmodell, authentische Fallvignetten, Videos und Expertenpanels zu verwenden. Dies alles trägt zu einer erfolgreichen und validen Testentwicklung bei. Das Festsetzen eines Scores zur Beurteilung der Lernenden und die Implementierung einer angemessenen Feedback-Struktur scheinen sinnvoll zu sein. *SJT*s können in großen Gruppen von Studierenden eingesetzt werden und sind weniger ressourcenintensiv als *OSCE*s. Die Entwicklungskosten eines videobasierten *SJT*s sind hoch, amortisieren sich aber im Laufe der Zeit.

*SJT*s und Selbsteinschätzungstests können leistungsbezogene Beurteilungen, wie *OSCE*s, in einem Curriculum zwar nicht ersetzen, sie können diese jedoch sinnvoll unterstützen und zum Beispiel hilfreich sein, um Studierende zu identifizieren, die Defizite in ihren empathischen Fähigkeiten haben. So können ihnen zusätzliche Schulungs- und Lernmöglichkeiten angeboten werden. Dies führt zu einem vollständigeren Bild der Fähigkeiten der Studierenden aus verschiedenen Blickwinkeln und zu einem besseren Lehrstandard. *SJT*s können flexibel an das Kompetenz- und Ausbildungsniveau der Lernenden angepasst werden.

Wir gehen davon aus, dass für den Erwerb von Empathiefähigkeit alle Komponenten von Empathie nach Morse [21] trainiert und erlernt werden müssen. Prozedurales und konzeptionelles Wissen, die Fähigkeit, die Emotionen anderer zu teilen und die Bereitschaft, empathisch auf andere zuzugehen, führen zu einem besseren empathischen Verhalten gegenüber Patient\*innen in realen Situationen. Im Rahmen eines longitudinalen Curriculums sollte eine Triangulation von Prüfungsinstrumenten zur Messung der Empathiefähigkeit eingesetzt werden.

### 3. Zusammenfassung

Die vorliegende Doktorarbeit widmet sich den übergeordneten Fragestellungen, wie ein *Situational Judgement Test* zur Prüfung des professionellen Umgangs mit Emotionen von Medizinstudierenden entwickelt und validiert werden kann und inwieweit dieses Prüfungsformat empathisches Verhalten voraussagen kann. In drei aufeinanderfolgenden Studien wurde ein computerbasierter *Situational Judgement Test* theoriegeleitet entwickelt (Studie 1), mit Medizinstudierenden der LMU München im Rahmen einer experimentellen Studie getestet (Studie 2), in Relation zu anderen Prüfungsformaten validiert und Empfehlungen für den curricularen Einsatz des *Situational Judgement Test* ausgesprochen (Studie 3).

Die Ergebnisse der Studien 1 und 2 (Paper I) zeigen, dass *VR-CoDES* ein praktikables Rahmenmodell zur Entwicklung eines *SJT*s zur Messung der Empathiefähigkeit von Medizinstudierenden ist. Um Medizinstudierenden dabei zu helfen, professionelles empathisches Verhalten zu entwickeln, muss die Prüfungssituation realistische Kontexte nachbilden, welches wir mit einer vorge-schalteten qualitativen Interviewstudie erreichten. Wie sich des Weiteren herausstellte waren die Verwendung von authentischen Szenarien als Grundlage für die Fallvignetten, der Einsatz von Videos in einer online-basierten Lernplattform und das Heranziehen von Expertenpanels zur Entwicklung und Validierung des Tests sowie eines entsprechenden Scores, wichtige Komponenten zur erfolgreichen Entwicklung und Implementierung eines *SJT*s.

Durch Studie 3 (Paper II) konnten wir uns einem Verständnis von ärztlicher Empathiefähigkeit und ihrer Messung annähern, indem wir verschiedene Instrumente, die die einzelnen Komponenten von Empathie abdecken, kombinierten. Wir stützten unsere Studie auf Definitionen und auf Instrumente, die im Bereich der medizinischen Ausbildung gut etabliert und validiert sind. Wir konnten zeigen, dass die Komponenten des multidimensionalen Empathie-Modells interkorreliert sind und nicht getrennt voneinander auftreten und, dass sich der von uns entwickelte *SJT* gut dazu eignet, die prozeduralen Wissensanteile von Empathie bei Medizinstudierenden zu messen. Die zwei von uns eingesetzten Scores (Experten-basiert und theoriebasiert) des *SJT*s sagten empathisches Verhalten im *OSCE* voraus.

Durch unsere Empfehlungen für den Einsatz von verschiedenen Prüfungsinstrumenten zur Messung medizinischer Empathiefähigkeit konnten wir einen sinnvollen Beitrag zur zukünftigen Gestaltung von longitudinalen Medizincurricula leisten. Hier findet auch der von uns entwickelte *SJT* einen berechtigten Platz und kann einen wichtigen Anteil des multidimensionalen Konstrukts der ärztlichen Empathiefähigkeit valide abbilden.

## 4. Abstract (English)

This dissertation addresses the overarching questions of how a *Situational Judgement Test* can be developed and validated for testing medical students' professional handling of emotions and to what extent this assessment format can forecast empathic behavior in real-life situations.

In three consecutive studies, a computer-based *Situational Judgement Test* was developed theory-based (study 1), validated within an experimental study with medical students at LMU Munich (study 2), and compared to other assessment formats (study 3).

The results of studies 1 and 2 (paper I) demonstrate that *VR-CoDES* is a feasible framework model for developing an *SJT* to measure empathy in medical students. To help medical students develop professional empathic behavior, the examination situation must replicate realistic contexts, which we achieved with an upstream qualitative interview study. Furthermore, the use of authentic scenarios as the basis for the case vignettes, the use of videos in an online-based learning platform, and the use of expert panels to develop and validate the test and a corresponding score were important components in our studies to successfully achieve this goal.

Through study 3 (paper II), we were able to approach an understanding of physicians' empathy and its measurement by combining different assessment instruments covering each component of empathy. We built our study on definitions and instruments that are well-known and validated in the field of medical education. We demonstrated that the components of the multidimensional empathy model are intercorrelated and not separate entities, and that the *SJT* we developed is well suited to scale the procedural knowledge components of empathy of medical students. The two scores of the *SJTs* used (*expert-based and theory-based*) predicted empathic behavior in an *OSCE*.

Through our recommendations for the use of different examination instruments to measure medical empathy skills, we were able to make a meaningful contribution to the future design of longitudinal medical curricula. Here, the *SJT* we developed also finds a legitimate place and can validly represent an important part of the multidimensional construct of medical empathy ability.

## **5. Paper I: Developing and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of Emotional Sequences (VR-CoDES)**

Graupe, T, Fischer, MR, Strijbos, JW & Kiessling, C (2020). Development and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of Emotional Sequences (VR-CoDES). *Patient Education and Counseling*, 103(9), 1839-1845. <https://doi.org/10.1016/j.pec.2020.04.001>.

**6. Paper II: The role of the emotive, moral, and cognitive components for the prediction of medical students' empathic behavior in an Objective Structured Clinical Examination (OSCE)**

Graupe, T, Giemsa, P, Schaefer, K, Fischer, MR, Strijbos, JW & Kiessling, C (2022). The role of the emotive, moral, and cognitive components for the prediction of medical students' empathic behavior in an Objective Structured Clinical Examination (OSCE). *Patient Education and Counseling*, 105(10), 3103-3109. <https://doi.org/10.1016/j.pec.2022.06.017>.

## 7. Literaturverzeichnis

- 1 Little P, Everitt H, Williamson I, Warner G, Moore M, Gould C (2001). Preferences of patients for patient-centered approach to consultation in primary care: observational study. *British Medical Journal*, 322, 1–7. doi: <https://doi.org/10.1136/bmj.322.7284.468>.
- 2 Coulter A (2002). Patients' views of the good doctor. Doctors have to earn patients' trust. *British Medical Journal*, 325, 668–669. doi: <https://doi.org/10.1136/bmj.325.7366.668>.
- 3 Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM (1997). Physician-Patient Communication. The Relationship With Malpractice Claims Among Primary Care Physicians and Surgeons. *Journal of the American Medical Association*, 277, 553–559. doi:10.1001/jama.1997.03540310051034.
- 4 Lefevre FX, Waters TM, Budetti PP (2000). A Survey of Physician Training Programs in Risk Management and Communication Skills for Malpractice Prevention. *Journal of Law, Medicine & Ethics*, 28, 258–266. <https://doi.org/10.1111/j.1748-720X.2000.tb00669.x>.
- 5 Tamblyn R, Abrahamowicz M, Dauphinee D, Wenghofer E, Jacques A, Klass D, Smee S, Blackmore D, Winslade N, Girard N, Du Berger R, Bartman I, Buckeridge DL, Hanley JA (2007). Physician scores on a national clinical skills examination as predictors of complaints to medical regulatory authorities. *Journal of the American Medical Association*, 298, 993–1001. doi:10.1001/jama.298.9.993.
- 6 Fallowfield L, Jenkins V (1999). Effective Communication Skills are the Key to Good Cancer Care. *European Journal of Cancer*, 35, 1592–1597. [https://doi.org/10.1016/S0959-8049\(99\)00212-9](https://doi.org/10.1016/S0959-8049(99)00212-9).
- 7 Maguire P, Pitceathly C (2002). Key communication skills and how to acquire them. *British Medical Journal*, 325, 697–700. doi: <https://doi.org/10.1136/bmj.325.7366.697>.
- 8 L. Zandbelt, Patient-centred Communication in the Medical Specialist Consultation, University of Amsterdam, Amsterdam, 2006. [https://pure.uva.nl/ws/files/3779195/45880\\_Zandbelt.pdf](https://pure.uva.nl/ws/files/3779195/45880_Zandbelt.pdf).
- 9 Feldman-Stewart D, Brundage M, Tishelman C (2005). A conceptual framework for patient–professional communication: an application to the cancer context. *Psychooncology*, 14 (10), 801–809. <https://doi.org/10.1002/pon.950>.
- 10 Norfolk T, Birdi K, Walsh D (2007). The role of empathy in establishing rapport in the consultation: a new model. *Medical Education*, 41, 690–697. <https://doi.org/10.1111/j.1365-2923.2007.02789.x>.
- 11 Lelorain S, Brédart A, Dolbeault S, Sultan S (2012). A systematic review of the associations between empathy measures and patient outcomes in cancer care. *Psychooncology*, 21, 1255–1264. <https://doi.org/10.1002/pon.2115>.
- 12 Neumann M, Edelhäuser F, Tauschel D, Fischer MR, Wirtz M, Woopen C (2011). Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Academic Medicine*, 86, 996–1009. <https://doi.org/10.1097/ACM.0b013e318221e615>.
- 13 Neumann M, Scheffer C, Tauschel D, Lutz, G, Wirtz M, Edelhäuser F (2012). Physician empathy: definition, outcome-relevance and its measurement in patient care and medical education, *GMS Journal of Medical Education*, 29. <https://doi.org/10.3205/zma000781>.
- 14 Elliott R, Bohart AC, Watson JC, Murphy D (2018). Therapist empathy and client outcome: an updated meta-analysis. *Psychotherapy*, 55, 399–410. <https://doi.org/10.1037/pst0000175>.
- 15 Ogle J, Bushnell J, Caputi P (2013). Empathy is related to clinical competence in medical care, *Medical Education*, 47, 824–831. <https://doi.org/10.1111/medu.12232>.
- 16 Derksen F, Bensing J, Lagro-Janssen A (2013). Effectiveness of empathy in general practice: a systematic review. *British Journal of General Practice*, 63, e76–e84. <https://doi.org/10.3399/bjgp13X660814>.

- 17 Shanafelt T, West C, Zhao X (2005). Relationship between increased personal well being and enhanced empathy among internal medicine residents, *Journal of General Internal Medicine*, 20, 559–564. <https://doi.org/10.1111/j.1525-1497.2005.0108.x>.
- 18 Stewart M (1995). Effective physician-patient communication and health outcomes: a review. *Canadian Medical Association Journal*, 152, 1423–1433. PMID: PMC1337906.
- 19 Ong LM, de Haes JC, Hoos AM, Lammes FB (1995). Doctor-patient communication: a review of the literature. *Social Science & Medicine*, 40, 903–918. [https://doi.org/10.1016/0277-9536\(94\)00155-M](https://doi.org/10.1016/0277-9536(94)00155-M).
- 20 Mercer, SW, Reynolds WJ (2002). Empathy and quality of care, *British Journal of General Practice*, 52, 9–12.
- 21 Morse JM, Anderson G, Bottorff JL, Yonge O, O'Brien B, Solberg SM (1992). Exploring empathy: a conceptual fit for nursing practice, *Image-the journal of nursing scholarship*, 24, 273–280. <https://doi.org/10.1111/j.1547-5069.1992.tb00733.x>.
- 22 Hsu I, Saha S, Korhuis PT, Sharp V, Cohn J, Moore RD, Beach MC (2012). Providing support to patients in emotional encounters: a new perspective on missed empathic opportunities. *Patient Education and Counseling*, 88 (3), 436–442. doi: <http://dx.doi.org/10.1016/j.pec.2012.06.015>.
- 23 Coulehan JL, Platt FW, Egener B, Frankel R, Lin CT, Lown B, Salazar WH (2001). "Let me see if I have this right . . .": words that help build empathy, *Annals of Internal Medicine*, 135 (3), 221–227. doi:<http://dx.doi.org/10.7326/0003-4819-135-3-200108070-00022>.
- 24 Zimmermann C, Del Piccolo L, Bensing J, Bergvik S, De Haes H, Eide H, Fletcher I, Goss C, Heaven C, Humphris G, Kim YM, Langewitz W, Meeuwesen L, Nuebling M, Rimondin M, Salmon P, Dulmen S, Wissow L, Zandbelt L, Finset A (2011). Coding patient emotional cues and concerns in medical consultations: the Verona coding definitions of emotional sequences (VR- CoDES). *Patient Education and Counseling*. 82 (2), 141–148. doi:<http://dx.doi.org/10.1016/j.pec.2010.03.017>.
- 25 Del Piccolo L, Finset A, Mellblom AV, Figueiredo-Braga M, Korsvold L, Zhou Y, Zimmermann C, Humphris G (2017). Verona coding definitions of emotional sequences (VR-CoDES): conceptual framework and future directions. *Patient Education and Counseling*, 100 (12), 2303–2311. doi:<http://dx.doi.org/10.1016/j.pec.2017.06.026>.
- 26 Adams K, Cimino J, Arnold R, Anderson W (2012). "Why should I talk about emotion? Communication patterns associated with physician discussion of patient expressions of negative emotion in hospital admission encounters. *Patient Education and Counseling*, 89(1), 44–50. <https://doi.org/10.1016/j.pec.2012.04.005>.
- 27 Eide H, Eide T, Rustøen T, Finset A (2011). Patient validation of cues and concerns identified according to Verona coding definitions of emotional sequences (VR- CoDES): a video- and interview-based approach. *Patient Education and Counseling*, 82 (2), 156–162. doi:<http://dx.doi.org/10.1016/j.pec.2010.04.036>.
- 28 Accreditation Council for Graduate Medical Education (2005). The ACGME Outcome Project: an Introduction. <http://cmfd.csmu.edu.tw/ezcatfiles/cmfd/img/img/447/ACGMEOutcomeProject.pdf>.
- 29 Frank JR (2005). The CanMEDS 2005 Physician Competency Framework Ottawa. The Royal College of Physician and Surgeons of Canada.
- 30 Working Group under a Mandate of the Joint Commission of Swiss Medical Schools (2008). Swiss Catalogue of Learning Objectives for Undergraduate Medical Training. [http://www.sclo.smifk.ch/downloads/sclo\\_2008.de](http://www.sclo.smifk.ch/downloads/sclo_2008.de).
- 31 The General Medical Council Education Committee (2009). Tomorrow's doctors: recommendations on undergraduate medical education. London: General Medical Council. [http://www.gmc-uk.org/TomorrowsDoctors\\_2009.pdf\\_39260971](http://www.gmc-uk.org/TomorrowsDoctors_2009.pdf_39260971).
- 32 Fischer, MR., Bauer D, Mohn K (2015). "Finally finished! National competence based catalogues of learning objectives for undergraduate medical education (NKLM) and dental education (NKLZ) ready for trial." *GMS Zeitschrift für medizinische Ausbildung*, 32(3). <https://doi.org/10.3205%2Fzma000977>.

- 33 Bundesrat. Beschluss des Bundesrates (2012). Erste Verordnung zur Änderung der Approbationsordnung für Ärzte. Drucksache 238/12. (Grunddrucksache 862/11). [http://www.med.uni-frankfurt.de/stud\\_med/download/aeappo\\_aenderungen\\_2012.pdf](http://www.med.uni-frankfurt.de/stud_med/download/aeappo_aenderungen_2012.pdf).
- 34 Langewitz W (2012). Zur Erlernbarkeit der Arzt-Patienten-Kommunikation in der Medizinischen Ausbildung. *Bundesgesundheitsblatt*, 55, 1176–1182.
- 35 Stepien KA, Baernstein A (2006). Educating for empathy – a review. *Journal for General Internal Medicine*, 21(5), 524–30. <https://doi.org/10.1111/j.1525-1497.2006.00443.x>.
- 36 Veloski J, Boex JR, Grasberger MJ, Evans A, Wolfson DB (2006). Systematic review of the literature on assessment, feedback and physicians' clinical performance. *Medical Teacher*, 28, 117–128. <https://doi.org/10.1080/01421590600622665>.
- 37 Hattie J, Timperley H (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.
- 38 Ericsson, K (2003). "Deliberate practice and acquisition of expert performance: a general overview." *Academic emergency medicine*, 15, 11, 988–994. <https://doi.org/10.1111/j.1553-2712.2008.00227.x>.
- 39 Schuwirth L, Van der Vleuten C (2003). ABC of learning and teaching in medicine. Written assessment. *British Medical Journal*, 326, 643–645.
- 40 Hemmerdinger JM, Stoddart SD, Lilford RJ (2007). A systematic review of tests of empathy in medicine. *BMC Medical Education*, 7, 1–8. <https://doi.org/10.1186/1472-6920-7-24>.
- 41 Davis MH (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44 (1), 113–126. <https://doi.org/10.1037/0022-3514.44.1.113>.
- 42 Hojat M, Mangione S, Nasca TJ, Cohen MJ, Gonnella JS (2001). The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. *Educational and Psychological Measurement*, 61, 349–365. <https://doi.org/10.1007/s10459-018-9839-9>.
- 43 Mercer SW, Maxwell M, Heaney D, Watt GC (2004). The consultation and relational empathy (CARE) measure: development and preliminary validation and reliability of an empathy-based consultation process measure. *Family Practice*, 21, 699–705. <https://doi.org/10.1093/fampra/cmh621>.
- 44 Turner J, Dankoski M (2008). Objective structured clinical exams: a critical review. *Family Medicine*, 40 (8), 574–578.
- 45 Harden RM (1988). What is an OSCE. *Medical Teacher*, 10, 19–22. <https://doi.org/10.3109/01421598809019321>.
- 46 Setyonugroho W, Kennedy KM, Kropmans TJB (2015). Reliability and validity of OSCE checklists used to assess the communication skills of undergraduate medical students: A systematic review. *Patient Education and Counseling*, 98, 1482–1491. <https://doi.org/10.1016/j.pec.2015.06.004>.
- 47 Cömert M, Zill JM, Christalle E, Dirmaier J, Härter M, Scholl I (2016). Assessing Communication Skills of Medical Students in Objective Structured Clinical Examinations (OSCE) - a systematic review of rating scales. *PLoS ONE*, 11 (3), e0152717. <https://doi.org/10.1371/journal.pone.0152717>.
- 48 Mazor KM, Ockene JK, Rogers HJ, Carlin MM, Quirk ME (2005). The relationship between checklist scores on a communication OSCE and analogue patients' perceptions of communication. *Advances in Health Sciences Education*, 10, 37–51. <https://doi.org/10.1007/s10459-004-1790-2>.
- 49 Burt J, Abel G, Elmore N, Campbell J, Roland M, Benson J (2014). Assessing communication quality of consultations in primary care: initial reliability of the Global Consultation Rating Scale, based on the Calgary-Cambridge Guide to the Medical Interview. *British Medical Journal Open*, 4, e004339. <https://doi.org/10.1136/bmjopen-2013-004339>.
- 50 Edgcumbe DB, Silverman J, Benson J (2012). An examination of the validity of EPSCALE using factor analysis. *Patient Education and Counseling*, 87, 120–124. <https://doi.org/10.1016/j.pec.2011.07.01>.



- 51 Scheffer S, Muehlinghaus I, Froehmel A, Ortwein H (2008). Assessing students' communication skills: validation of a global rating. *Advances in Health Sciences Education*, 13, 583–592. <https://doi.org/10.1007/s10459-007-9074-2>.
- 52 Roter D, Larson S (2002). The Roter Interaction Analysis System (RIAS): utility and flexibility for analysis of medical interactions. *Patient Education and Counseling*, 46, 243–251. [https://doi.org/10.1016/s0738-3991\(02\)00012-5](https://doi.org/10.1016/s0738-3991(02)00012-5).
- 53 Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R (2003). "Shared decision making: developing the OPTION scale for measuring patient involvement." *British Medical Journal Quality & Safety*, 12(2), 93–99.
- 54 Kiessling C, Fabry G, Fischer R, Steiner C & Langewitz WA (2013). German translation and construct validation of the Patient-Provider-Orientation Scale (PPOS-D12). *Psychotherapie, Psychosomatik, Medizinische Psychologie*, 64.3-4, 122–127.
- 55 Charlin B, Brailovsky C, Roy L, Goulet F (2000). The Script Concordance Test: A Tool to Assess the Reflective Clinician. *Teaching and Learning in Medicine*, 12(4), 189–195. [https://doi.org/10.1207/S15328015TLM1204\\_5](https://doi.org/10.1207/S15328015TLM1204_5).
- 56 Hrynchak P, Glover Takahashi S, Nayer M (2014). Key-feature questions for assessment of clinical reasoning: a literature review. *Medical Education*, 48(9), 870–883. <https://doi.org/10.1111/medu.12509>.
- 57 Patterson F, Ashworth V, Zibarras L, Coan P, Kerrin M, O'Neil P (2012). Evaluations of situational judgement tests to assess non-academic attributes in selection, *Medical Education*, 46 (9), 850–868. doi:<http://dx.doi.org/10.1111/j.1365-2923.2012.04336.x>.
- 58 McDaniel MA, Nguyen NT (2001). Situational judgment tests: a review of practice and constructs assessed. *International Journal of Selection and Assessment*, 9 (1–2), 103–113. doi:<http://dx.doi.org/10.1111/1468-2389.00167>.
- 59 Nguyen NT, McDaniel MA (2003). Response instructions and racial differences in a situational judgment test. *Human Resource Management Journal*, 8 (1), 33–44.
- 60 McDaniel MA, Hartman NS, Whetzel DL, Grubb WL III (2007). Situational judgment tests, response instructions, and validity: a meta-analysis. *Journal of Personnel Psychology*, 60 (1), 63–91. doi:<http://dx.doi.org/10.1111/j.1744-6570.2007.00065.x>
- 61 Christian MS, Edwards BD, Bradley JC (2010). Situational judgment tests: constructs assessed and a meta-analysis of their criterion-related validities, *Journal of Personnel Psychology*, 63 (1), 83–117. doi:<http://dx.doi.org/10.1111/j.1744-6570.2009.01163.x>.
- 62 Motowidlo SJ, Hooper AC, Jackson HL (2006). Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. *Journal of Applied Psychology*, 91 (4), 749–761. doi:<http://dx.doi.org/10.1037/0021-9010.91.4.749>.
- 63 Motowidlo SJ, Beier ME (2010). Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal of Applied Psychology*, 95 (2), 321–33. doi:<http://dx.doi.org/10.1037/a0017975>.
- 64 Lievens F, Patterson F (2011). The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *Journal of Applied Psychology* 96 (5), 927–940. doi:<http://dx.doi.org/10.1037/a0023496>.
- 65 Patterson F, Rowett E, Hale R, Grant M, Roberts C, Cousans F, Martin S (2016). The predictive validity of a situational judgement test and multiple-mini interview for entry into post-graduate training in Australia. *BMC Medical Education*, 16 (1), 87. doi:<http://dx.doi.org/10.1186/s12909-016-0606-4>.
- 66 Chan D, Schmitt N (2002). Situational judgment and job performance, *Human Performance*, 15 (3), 233–254. doi:[http://dx.doi.org/10.1207/S15327043HUP1503\\_01](http://dx.doi.org/10.1207/S15327043HUP1503_01).
- 67 Ployhart RE, Ehrhart MG (2003). Be careful what you ask for: effects of response instructions on the construct validity and reliability of situational judgment tests. *International Journal of Selection and Assessment*, 11 (1), 1–16. doi:<http://dx.doi.org/10.1111/1468-2389.00222>.

- 68 Koczwara A, Patterson F, Zibarras L, Kerrin M, Irish B, Wilkinson M (2012). Evaluating cognitive ability, knowledge tests and situational judgement tests for postgraduate selection. *Medical Education*, 46 (4), 399–408. doi:http://dx.doi.org/10.1111/j.1362923.2011.04195.x.
- 69 Husbands A, Rodgerson MJ, Dowell J, Patterson F (2015). Evaluating the validity of an integrity-based situational judgement test for medical school admissions, *BMC Medical Education*, 15 (1), 144. doi:http://dx.doi.org/10.1186/s12909-015-0424-0.
- 70 Patterson F, Zibarras L, Ashworth V (2016). Situational judgement tests in medical education and training: research, theory and practice: AMEE Guide No. 10. *Medical Teacher*, 38 (1), 3–17. doi:http://dx.doi.org/10.3109/0142159X.2015.1072619.
- 71 Lievens F, Sackett PR (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, 97 (2), 460–468. doi:http://dx.doi.org/10.1037/a0025741.
- 72 Whetzel DL, McDaniel MA, Nguyen NT (2008). Subgroup differences in situational judgment test performance: a meta-analysis. *Human Performance*, 21 (3), 291–309. doi:http://dx.doi.org/10.1080/08959280802137820.
- 73 Luschin-Ebengreuth M, Dimai HP, Ithaler D, Neges HM, Reibnegger G (2015). Situational judgment test as an additional tool in a medical admission test: an observational investigation. *BMC Research Notes*, 8 (1), 81. doi:http://dx.doi.org/10.1186/s13104-015-1033-z.
- 74 Roberts C, Clark T, Burgess A, Frommer M, Grant M, Mossman K (2014). The validity of a behavioural multiple-mini-interview within an assessment centre for selection into specialty training. *BMC Medical Education*, 14 (1), 169. doi:http://dx.doi.org/10.1186/1472-6920-14-169.
- 75 Goss BD, Ryan AT, Waring J, Judd T, Chiavaroli NG, O'Brien RC, McColl GJ (2017). Beyond selection: the use of situational judgement tests in the teaching and assessment of professionalism. *Academic Medicine*, 92 (6), 780–784. doi:http://dx.doi.org/10.1097/ACM.0000000000001591.
- 76 Graupe T, Fischer MR, Strijbos JW, Kiessling C (2020). Development and piloting of a Situational Judgement Test for emotion-handling skills using the Verona Coding Definitions of Emotional Sequences (VR-CoDES). *Patient Education and Counseling*, 103, 1839–1845. https://doi.org/10.1016/j.pec.2020.04.001.
- 77 Giemsa P, Wübbolding C, Fischer MR, Graupe T, Härtl A, Lenz C, Sanftenberg L, Schelling J, Schüttpeitz-Brauns K & Kiessling C (2020). What works best in a general practice specific OSCE for medical students: Mini-CEX or content-related checklists? *Medical Teacher*, 42(5), 578–584. doi: 10.1080/0142159X.2020.1721449.

## Danksagung

Mein besonderer Dank gilt meiner Betreuerin Frau Professor Dr. Claudia Kiessling für ihre intensive Betreuung, für die vielen konstruktiven Ideen, für ihre Begeisterung für statistische Fragestellungen, für ihr Netzwerk, für ihre wertvolle Zeit sowie für die gemeinsame kritische Auseinandersetzung mit der Thematik meiner Dissertation, durch die ich sehr viel gelernt habe.

Ich danke meinem Doktorvater Herrn Professor Dr. Martin Fischer für seine konstruktive Begleitung meiner Arbeit, seine Geduld und die herzliche Anbindung an das Institut für Didaktik und Ausbildungsforschung in der Medizin (DAM) am LMU Klinikum. Ich danke meinem Doktorvater Professor Dr. Jan-Willem Strijbos für seine klugen Einschätzungen, seine klare Linie und sein stets schnelles und effektives Feedback.

Ich danke allen Studierenden der Medizinischen Fakultät der LMU München, die durch ihre Teilnahme an unseren Studientagen an der Entstehung dieser Arbeit mitgewirkt haben.

Meiner lieben Mama sowie all meinen wundervollen Freunden, danke ich für ihre großartige emotionale Unterstützung, ihren unerschütterlichen Glauben an mich und für viele geleistete Babysitter-Dienste. Mein allerliebster Dank gilt meinem Ehemann Florian, der mich über die Jahre hinweg stets mit Rat und Tat unterstützt hat. Er war immer für mich da, hat mit seinem ästhetischen Blick und seinen großartigen EDV-Fähigkeiten die Optik all meiner Abbildungen maßgeblich verbessert, hörte sich zum wiederholten Male Vorträge zu meinem Thema an, hielt mir zu jeder Zeit den Rücken frei und war meine allergrößte Stütze. Danke!