



Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR ERZIEHUNGSWISSENSCHAFT

Erkennen von Lernbarrieren in Textvignetten

ErLe

**Entwicklung und Validierung eines quantitativen Verfahrens zur
Beurteilung situationsspezifischer Fähigkeiten**

Dissertation
zur Erlangung des akademischen Grades
einer Doktorin der Philosophie
am Fachbereich 2: Schul- und Grundschulpädagogik,
Sozialpädagogik sowie
Pädagogik bei Behinderung und Benachteiligung
der Universität Hamburg

Vorgelegt von
Stephanie Katharina Wenck
aus Hamburg

Hamburg, Februar 2024

Erstgutachterin: Prof. Dr. Gabriele Ricken, Universität Hamburg

Zweitgutachter: Prof. Dr. Jörg Doll (em.), Universität Hamburg

Mündliche Gutachterin: Prof. Dr. Barbara Hänel-Faulhaber, Universität Hamburg

Tag der mündlichen Prüfung: 06. Juni 2024

Dankworte

Zuerst möchten ich meinen drei Kindern Henri, Leo und Oskar danken, weil sie mir am Ende verziehen haben, dass ich in intensiven Arbeitsphasen kaum ein Ohr für einfache Fragen hatte. Meinem Mann Christopher und meinen Eltern danke ich dafür, dass sie immer an mich geglaubt haben.

Mein Dank gilt auch allen, die im Laufe der Zeit zu meiner Arbeit auf unterschiedliche Weise entscheidend beigetragen haben, mit: Codierungen, Statistikexpertise, kritischen Nachfragen, Seminarzeit für Testdurchführungen, offenen Ohren, Druckaufträgen, fachlichen Diskussionen, Motivationsreden, neuen Ideen u. v. m.....

Jörg Doll danke ich für den prüfenden Blick auf meine Daten und die statistische Beratung.

Und last but not least danke ich meiner Doktormutter Gabi Ricken für ungezählte Stunden intensiven Nachdenkens und Diskutierens, für Geduld, Genauigkeit und für die Impulse, ohne die diese Arbeit nicht in dieser Form entstanden wäre. Und für Schokolade.

Inhaltsverzeichnis

	Seite
Abkürzungsverzeichnis	8
Abbildungsverzeichnis	9
Tabellenverzeichnis	10
1 Einleitung	13
I Theoretischer Hintergrund und Stand der Forschung	17
2 Was sind Lernbarrieren?	18
3 Was bedeutet „Erkennen“?	21
3.1 Die Basis: Das Professionelle Wissen	21
3.2 Die Schlüsselkompetenz: Situationsspezifische Fähigkeiten	22
3.3 Das Diagnostische daran: Erkennen als diagnostische Kompetenz	25
4 Forschungsstand zur Erfassung der professionellen Wahrnehmung	27
4.1 Einsatz von Vignetten als Stimulus	27
4.2 Vignettenbasierte Verfahren	28
4.2.1 Beobachtungsgegenstände.....	39
4.2.2 Methodische Ansätze.....	39
4.2.3 Erfasste Kompetenzfacetten und -unterschiede	40
4.2.4 Praxisrelevanz der erfassten Kompetenzen.....	41
5 Forschungsdesiderata	42
I Entwicklung eines Verfahrens zur Erfassung der Kompetenz Erkennen von Lernbarrieren	44
6 Systematisierung und Operationalisierung von Lernbarrieren	45
6.1 Unterrichtsqualität	45
6.2 Merkmale guten Unterrichts	47
6.3 Konkretisierung und Operationalisierung von Unterrichtsmerkmalen und Lernbarrieren	54
6.3.1 Klassenführung	56
6.3.2 Lernunterstützung	58
6.3.3 Lernbarrieren	61

7	Modellierung der Kompetenz Erkennen von Lernbarrieren	66
8	Konstruktionskriterien für Vignetten und Items	70
8.1	Vignettenformat	70
8.2	Komplexität der Vignette.....	73
8.3	Sprachliche Gestaltung und Informationsgehalt der Vignette	75
8.4	Vignette Philipp	78
8.5	Erprobung einer Kurzvignette.....	79
8.6	Qualitative und quantitative Ansätze	80
8.7	Präsentation der Vignetten	81
8.8	Itemkonstruktion	82
8.9	Auswertungsverfahren	83
9	Erster Entwicklungsschritt	84
9.1	Fragestellungen und Hypothesen des ersten Entwicklungsschritts.....	84
9.2	Konstruktion der Aufgaben.....	85
9.3	Stichprobe.....	89
9.4	Design und Durchführung des ersten Entwicklungsschritts	90
9.4.1	Auswertung der Akzeptanzbefragung.....	91
9.4.2	Auswertung der ersten Teilaufgabe.....	91
9.4.3	Auswertung der offenen zweiten Teilaufgabe.....	91
9.4.4	Auswertung der geschlossenen zweiten Teilaufgabe	92
9.5	Ergebnisse und Interpretation des qualitativen Ansatzes	93
9.5.1	Ergebnisse der Akzeptanzbefragung.....	93
9.5.2	Interpretation der Akzeptanzbefragung	94
9.5.3	Ergebnisse der ersten Teilaufgabe	94
9.5.4	Interpretation der Ergebnisse der ersten Teilaufgabe.....	98
9.5.5	Ergebnisse der zweiten offenen Teilaufgabe	99
9.5.6	Interpretation der Ergebnisse der zweiten offenen Teilaufgabe	101
9.6	Ergebnisse und Interpretation des quantitativen Ansatzes.....	102
9.6.1	Ergebnisse zu mittleren Schwierigkeitsgraden, Trennschärfen und Reliabilität	103
9.6.2	Interpretation der mittleren Schwierigkeitsgrade, Trennschärfen und Reliabilität	104
9.6.3	Ergebnisse zur Validität.....	105
9.6.3.1	Ergebnisse zur Konstruktvalidität	105
9.6.3.2	Interpretation der Konstruktvalidität.	106
9.6.3.3	Ergebnisse zur Kriteriumsvalidität	107
9.6.3.4	Interpretation der Kriteriumsvalidität.	108

9.7	Zusammenfassung und Konsequenzen für den nächsten Entwicklungsschritt	108
10	Zweiter Entwicklungsschritt	110
10.1	Überarbeitung der Vignette und Items.....	110
10.2	Fragestellungen und Hypothesen	115
10.3	Stichprobe.....	116
10.4	Design und Durchführung.....	117
10.5	Ergebnisse und Interpretation	118
10.5.1	Ergebnisse zu mittleren Schwierigkeitsgraden, Trennschärfen und Reliabilität	118
10.5.2	Interpretation der mittleren Schwierigkeitsgrade, Trennschärfen und Reliabilität	121
10.5.3	Ergebnisse zur Konstruktvalidität	124
10.5.4	Interpretation der Konstruktvalidität	125
10.5.5	Ergebnisse zur Kriteriumsvalidität	125
10.5.6	Interpretation zur Kriteriumsvalidität.....	125
10.6	Diskussion und Konsequenzen für das weitere Vorgehen.....	126
11	Dritter Entwicklungsschritt	129
11.1	Konstruktion der Vignetten und Items	129
11.1.1	Konstruktion der Vignetten	130
11.1.2	Konstruktion und Zuordnung der Items.....	132
11.1.3	Überarbeitung der Instruktion und der Skala	139
11.2	Fragestellungen und Hypothesen	140
11.3	Stichprobe.....	141
11.4	Design und Durchführung.....	142
11.5	Ergebnisse und Interpretation	143
11.5.1	Ergebnisse zu mittleren Schwierigkeitsgraden	144
11.5.2	Interpretation zu mittleren Schwierigkeitsgraden	145
11.5.3	Ergebnisse zu Trennschärfen und Reliabilität	146
11.5.4	Interpretation der Trennschärfen und Reliabilität.....	148
11.5.5	Ergebnisse zur Konstruktvalidität	149
11.5.6	Interpretation der Konstruktvalidität	150
11.5.7	Ergebnisse zur Kriteriumsvalidität	152
11.5.8	Interpretation der Kriteriumsvalidität	154
11.5.9	Ergebnisse zur Kreuzvalidität	155
11.5.10	Interpretation der Kreuzvalidität	157
11.6	Diskussion und Konsequenzen für die Weiterentwicklung.....	158

12	Vierter Entwicklungsschritt	160
12.1	Überarbeitung der Vignetten und Items	160
12.1.1	Überarbeitung der Vignetten.....	161
12.1.2	Überarbeitung der Instruktion, Items und Skala.....	165
12.2	Fragestellungen und Hypothesen	167
12.3	Stichprobe.....	169
12.4	Design und Durchführung.....	170
12.4.1	Auswertung des Vignettentests	171
12.4.2	Auswertung der Markierungsaufgabe	172
12.5	Ergebnisse und Interpretation	172
12.5.1	Ergebnisse zur Authentizität und Vorstellbarkeit	172
12.5.2	Interpretation zu Authentizität und Vorstellbarkeit	173
12.5.3	Ergebnisse zu mittleren Schwierigkeitsgraden	173
12.5.4	Interpretation zu mittleren Schwierigkeitsgraden	175
12.5.5	Ergebnisse zu Trennschärfen und Reliabilität	175
12.5.6	Interpretation der Trennschärfen und Reliabilität.....	176
12.5.7	Ergebnisse zur Konstruktvalidität	176
12.5.8	Interpretation zur Konstruktvalidität.....	178
12.5.9	Mögliche Einflussfaktoren auf das Erkennen von Lernbarrieren.....	179
12.5.9.1	Ergebnisse zu Merkmalen der Lesenden: Einfluss der Expertise.	179
12.5.9.2	Interpretation zu Merkmalen der Lesenden: Einfluss der Expertise.	180
12.5.9.3	Ergebnisse zum Einfluss durch die Vignettengestaltung und die Einschätzung der Klassenatmosphäre: Das Gelingen der Situation insgesamt.	181
12.5.9.4	Interpretation zum Einfluss durch die Vignettengestaltung und die Einschätzung der Klassenatmosphäre: Das Gelingen der Situation insgesamt.....	186
12.5.9.5	Ergebnisse zur Instruktion: Einfluss der Markierungsaufgabe....	187
12.5.9.6	Interpretation zur Instruktion: Einfluss der Markierungsaufgabe.....	189
12.6	Diskussion	191
13	Kritische Reflexion und Ausblick	193
13.1	Der Barrierebegriff.....	193
13.2	Die Kompetenz des Erkennens von Lernbarrieren	194
13.3	Die Vignettenkonstruktion.....	195
13.4	Der methodische Ansatz.....	197

13.5 Die Stichprobe	199
13.6 Die zentralen Ergebnisse	200
13.7 Ausblick.....	202
Literaturverzeichnis	206
Anhang	237
Zusammenfassung	335
Summary	338
Liste der aus dieser Dissertation hervorgegangenen Veröffentlichungen	341
Lebenslauf	342

Abkürzungsverzeichnis

Abkürzung	Bedeutung
CLIPSS	Classroom Management in Primary and Secondary Schools
CME	Classroom Management Expertise
COACTIV-R	Cognitive activation in the mathematics classroom and professional competence of teachers – Referendariatsstudie
KF	Klassenführung
LU	Lernunterstützung
LUV	Lerneß aus Unterrichtsvideos
PISA	Program for International Student Assessment
PW	Professionelle Wahrnehmung
TEDS-FU	Teacher Education and Development Study - Follow Up
TIMSS	Trends in International Mathematics and Science Study
ViU	Videobasierte Unterrichtsanalyse

Abbildungsverzeichnis

	Seite
Abbildung 1	Das Kompetenzmodell der Transformation von Kompetenz in Performanz24
Abbildung 2	Übersicht über die Entwicklungsschritte von ErLe.....43
Abbildung 3	Kompetenzmodell Erkennen von Lernbarrieren.....69
Abbildung 4	Aufgaben und Aufbau des ersten Entwicklungsschritts.....88
Abbildung 5	Histogramm zur Häufigkeitsverteilung der Testscores im Erkennen von Lernbarrieren101

Tabellenverzeichnis

	Seite
Tabelle 1 Übersicht über Studien zur vignettenbasierten Erfassung professioneller Wahrnehmung	30
Tabelle 2 Übersicht über Unterrichtsmerkmale und Bündelung zu Dimensionen	50
Tabelle 3 Überblick über die in ErLe verwendeten Lernbarrieren.....	62
Tabelle 4 Häufigkeiten für die Stichprobe des ersten Entwicklungsschritts.....	90
Tabelle 5 Mittelwerte zur Einschätzung der Verständlichkeit, Schwierigkeit und Interessantheit	94
Tabelle 6 Ergebnis der Inhaltsanalyse der ersten Teilaufgabe: Von Studierenden genannte Kategorien.....	95
Tabelle 7 Häufigkeiten für die Nennungen der Kategorien je Studienganggruppe	96
Tabelle 8 Häufigkeiten für die Nennungen der zusammengefassten Kategorien nach Einflussfaktoren je Gruppe.....	98
Tabelle 9 Häufigkeiten für die dichotomisierten Variablen, erkannte Barrieren nach Studiengängen	100
Tabelle 10 Deskriptive Statistik zu geschlossenen Items zu Barrieren	103
Tabelle 11 Trennschärfen der Items zu Barrieren	104
Tabelle 12 Ergebnis der Faktorenanalyse	106
Tabelle 13 Ergebnisse der Expertendiskussion und Überarbeitungen von Vignette und Items	111
Tabelle 14 Häufigkeiten der Fachsemester für die Stichprobe der zweiten Datenerhebung	117
Tabelle 15 Deskriptive Statistik zu Items zu Barrieren.....	119
Tabelle 16 Mittelwerte, Standardabweichung und Variationskoeffizienten der Items zu Barrieren für die Gruppe der Sonderpädagogikstudierenden, 1. und 2. Datenerhebung im Vergleich	120
Tabelle 17 Trennschärfen der Items zu Barrieren	121
Tabelle 18 Rangfolge der Items zu Barrieren gemessen am Schwierigkeitsgrad, 1. und 2. Datenerhebung im Vergleich	123
Tabelle 19 Ergebnis der Faktorenanalyse	124
Tabelle 20 Veränderungen der Trennschärfe und Faktorenzugehörigkeit in Bezug zu Überarbeitungen von Kontext, Vignette, Item-Polung und Item-Formulierung..	127

Tabelle 21	Übersicht über die Verteilung der Barrieren und Itempolungen auf die vier Vignetten.....	132
Tabelle 22	Häufigkeiten für die Stichprobe der dritten Datenerhebung	142
Tabelle 23	Deskriptive Statistik zu Items zu Barrieren und Items zu gelingenden Situationen	144
Tabelle 24	Trennschärfen der Items zu Barrieren in Vignette 1	146
Tabelle 25	Trennschärfen der Items zu Barrieren in Vignette 2	147
Tabelle 26	Trennschärfen der Items zu Barrieren in Vignette 3	147
Tabelle 27	Trennschärfen der Items zu Barrieren in Vignette 4	147
Tabelle 28	Trennschärfen der Items zu Barrieren für die Gesamtskala der Vignetten 2, 3 und 4	148
Tabelle 29	Ergebnis der Faktorenanalyse	150
Tabelle 30	Vergleich der Mittelwerte von Items zu Barrieren und zu gelingenden Situationen für Teilstichproben mit unterschiedlich zutreffenden Beurteilungen von Barrieren.....	152
Tabelle 31	Ergebnisse der Faktorenanalyse.....	156
Tabelle 32	Aufteilung der Lernbarrieren auf die Vignetten 1–4 unter Zuordnung von gelingenden Situationen und Barrieren	165
Tabelle 33	Häufigkeiten für die Stichprobe des vierten Entwicklungsschritts.....	170
Tabelle 34	Mittelwerte für die Items zu Authentizität und Vorstellbarkeit der Vignetten 1–4, unterteilt nach Teilstichproben	173
Tabelle 35	Deskriptive Statistik zu Items zu Barrieren und Items zu gelingenden Situationen	174
Tabelle 36	Trennschärfen der Items zu gelingenden Situationen in den Vignetten 1-4.....	175
Tabelle 37	Trennschärfen der Items zu Barrieren in den Vignetten 1-4.....	176
Tabelle 38	Ergebnisse der Faktorenanalyse.....	177
Tabelle 39	Mittelwerte für die Einschätzungen der Klassenatmosphäre in den Vignetten 1–4, aufgeteilt nach Teilstichproben.....	182
Tabelle 40	Mittelwerte für Items zu Barrieren und Items zu gelingenden Situationen in Abhängigkeit von der Anzahl der Barrieren pro Vignette, aufgeteilt nach Teilstichproben.....	183
Tabelle 41	Anzahl markierter Stellen, Anteil richtiger Markierungen, Anteil richtiger Beurteilungen und Anteil falscher Beurteilungen in Bezug zum Mittelwert aller Items für die Vignetten 1-4	188

Tabelle 42	Mittelwerte der Items zu Barrieren und der Items zu gelingenden Situationen in Abhängigkeit von der Bearbeitung der Markierungsaufgabe	189
-------------------	---	-----

1 Einleitung

Schule spielt im Leben von Kindern, Jugendlichen und deren Familien eine zentrale Rolle. Sie prägt den Alltag, bestimmt Bildungschancen und ist ein entscheidender Faktor für die Gesundheit und das Wohlbefinden (Bilz et al., 2016; Hanisch & Hennemann, 2020; Kaman et al., 2020). Lehrkräften kommt damit eine Schlüsselrolle zu, denn sie gestalten als Akteurinnen und Akteure den Schulalltag.

Im Zuge der Inklusion werden die Beschreibung, Erfassung und Ausbildung der Kompetenzen von Lehrkräften verstärkt diskutiert (Baumert & Kunter, 2006, 2011; Blömeke et al., 2015a; Frey, 2008; Heran-Dörr et al., 2007; Lindmeier, 2011; Terhart, 2007). Mit dem Unterrichten in inklusiven Klassen ist angesichts heterogener Lerngruppen vor allem die Herausforderung verbunden, die Aktivität und Teilhabe aller Kinder und Jugendlichen zu ermöglichen. Es sind also insbesondere Kompetenzen gefragt, barrierefreie und für alle zugängliche schulische Bedingungen zu gestalten.

Die Begriffe Barriere bzw. Barrierefreiheit werden häufig im Kontext von räumlichen Bedingungen genutzt. Als barrierefrei werden beispielsweise rollstuhlgerechte Wohnungen bezeichnet oder Informationen für sehbehinderte Menschen in öffentlichen Gebäuden (vgl. Barrierefrei leben e. V., o. D.; Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderung, o. D.a; Kompetenzzentrum für ein barrierefreies Hamburg, o. D.). Barrieren können aber auch in Unterrichtsprozessen entstehen und alle Bedingungen betreffen, die für Lehr-Lernprozesse eine Rolle spielen. Sie können in Handlungen entstehen, in der Gestaltung von Unterricht, in der Interaktion zwischen Schülerinnen und Schülern und Lehrkräften. Der Kompetenz, solche Barrieren zu erkennen, kommt eine grundlegende Bedeutung für die Gestaltung barrierefreien Unterrichts zu. Eine größere Heterogenität in inklusiven Lerngruppen bringt dabei zwar eine besondere Herausforderung mit sich, allerdings können Barrieren in allen Lerngruppen und für alle Kinder und Jugendlichen entstehen, so dass das Erkennen von Lernbarrieren nicht als rein sonderpädagogische oder auf Inklusion bezogene Kompetenz verstanden werden soll.

Die vorliegende Arbeit ist im Rahmen des Projekts *Professionelles Lehrerhandeln zur Förderung fachlichen Lernens unter sich verändernden gesellschaftlichen Bedingungen* (ProfaLe) entstanden, einem Projekt der vom Bundesministerium für Bildung und Forschung finanzierten *Qualitätsoffensive Lehrerbildung*. Ziel war die Koordination fachlicher, fachdidaktischer, pädagogischer sowie schulpraktischer Aspekte sowie die Überprüfung von deren Wirksamkeit innerhalb der Lehrkräftebildung.

Ziel des Handlungsfelds *Inklusion* war unter anderem die Beschreibung von Expertisen für inklusiven Unterricht (vgl. Ricken & Degenhardt, 2016) und die Entwicklung von Seminarkonzepten, die die Auseinandersetzung mit inklusiven Lehr-Lernprozessen fördern. Im Zusammenhang damit stellt sich auch die Frage nach Verfahren, die die Wahrnehmung und Analysefähigkeit solcher Lehr-Lernprozesse und die Lernstände und -fortschritte von Studierenden erfassen.

Auch wenn die Beschreibung, Ausbildung und Erfassung von Expertisen für inklusiven Unterricht alle aktiven Lehrkräfte betreffen, liegt der Fokus dieser Arbeit auf dem Lehramtsstudium und stellt Ergebnisse überwiegend für Stichproben dar, die aus Studierenden bestehen.

Die Arbeit umfasst die folgenden leitenden Fragestellungen:

1. Was sind Lernbarrieren und welche Lernbarrieren sollten sinnvollerweise für die Entwicklung eines Verfahrens zum Erkennen von Lernbarrieren konkret beschrieben und operationalisiert werden?
2. Was bedeutet es, Lernbarrieren zu erkennen, und welche Kompetenz ist dafür erforderlich?
3. Mit welchem Ansatz lässt sich diese Kompetenz erfassen?

Es wird in dieser Arbeit konzeptuell mit dem Begriff *Lernbarriere* gearbeitet. Darunter kann grundsätzlich alles verstanden werden, was Lehr-Lernprozesse stört und alle Faktoren betreffen, die in einer komplexen Lehr-Lernsituation zum Lernen beitragen (Ricken, 2017). Im ersten Teil der Arbeit wird dies näher betrachtet und es werden Barrieren in der Unterrichtsgestaltung fokussiert. Dies erscheint relevant, um der Komplexität von Lehr-Lernprozessen gerecht zu werden und Ursachen für Lern- und Verhaltensschwierigkeiten auch über individuelle Bedingungen hinausgehend zu bestimmen.

Der Begriff des Erkennens wird im nächsten Abschnitt kompetenztheoretisch beleuchtet. Das Konzept der *professionellen Wahrnehmung* (Seidel et al., 2010; Sherin & van Es, 2009) und das *Kompetenzmodell* von Blömeke et al. (2015a) sind geeignet, die Kompetenz des Erkennens von Lernbarrieren theoretisch zu beschreiben. Das situationsspezifische Wahrnehmen unterrichtsrelevanter Ereignisse wird in dieser Arbeit in einen diagnostischen Kontext gerückt und im Sinne einer nicht beiläufigen, sondern gezielten Wahrnehmung als Entsprechung einer Beobachtung verstanden. Beobachten gilt als grundlegende diagnostische Methode zur Analyse von Lehr-Lernsituationen (Hollenweger, 2021). Aus sonderpädagogischer Sicht sind insbesondere solche Unterrichtsmomente relevant, die Annahmen über die Entstehung von Lernbarrieren zulassen (Tönnissen et al., 2023).

Studien zur Erfassung der professionellen Wahrnehmung (PW) zeigen, dass es grundsätzlich möglich ist, diese Kompetenz mittels Vignetten, also kurzen Darstellungen von Unterrichtsszenen, zu erfassen. Erfasst werden vor allem die PW fachspezifischer oder fachdidaktischer Aspekte sowie Aspekte der Unterrichtsmerkmale Klassenführung (KF) und Lernunterstützung (LU). Das Erfassen der Erkennung von Lernbarrieren als diagnostische Aufgabe stellt ein Forschungsdesiderat dar.

Hauptanliegen dieser Arbeit ist die mehrschrittige Entwicklung und Validierung eines Verfahrens zur Erfassung der Kompetenz, Lernbarrieren zu erkennen. Dabei stellt sich die Frage danach, wie Impulse durch die Vignettenkonstruktion und das Befragungsformat gestaltet werden müssen, um die Fokussierung, Wahrnehmung und Beurteilung von Lernbarrieren in der Unterrichtsgestaltung auszulösen und abzubilden. In einem pilotierenden, explorativen Vorgehen werden Texte, Instruktionen und Items schrittweise entwickelt, verändert und erprobt. Zunächst werden Lernbarrieren definiert und operationalisiert. Da Lernbarrieren in der Gestaltung von Unterricht verortet werden sollen, werden zunächst Kriterien zur Beschreibung qualitätvollen Unterrichts aufgezeigt. Diese dienen als Grundlage für die anschließende Beschreibung von Lernbarrieren im Sinne nicht-erfüllter Qualitätskriterien.

Die Kompetenz, diese Lernbarrieren zu erkennen, wird durch eine Adaption des Kompetenzmodells von Blömeke et al. (2015a) modelliert. Das Modell betont unter anderem die Fokussierung von Unterrichtsmerkmalen als Voraussetzung für (1) die Wahrnehmung von Interaktionen, die mit dem Fehlen oder dem fehlerhaften Umsetzen von Unterrichtsmerkmalen zusammenhängen, sowie (2) die Beurteilung der Auswirkung auf den Lehr-Lernprozess.

Im ersten Entwicklungsschritt ist die Frage leitend, wie Impulse durch Vignette, Instruktion und Frageformat bzw. Items gestaltet werden müssen, damit Barrieren in Unterrichtsprozessen fokussiert werden, wahrnehmbar sind und Beurteilungsprozesse hinsichtlich deren Auswirkung in Gang gesetzt werden. Es werden unterschiedliche Vignettentypen in Kombination mit offenen und geschlossenen Ansätzen erprobt. Die Ergebnisse geben die Schwerpunktsetzung für die weitere Entwicklung vor und bestimmen das grundsätzliche Vignetten- und Itemformat.

Die Gestaltung von Vignette, Instruktion und Items wird im zweiten Entwicklungsschritt durch eine Expertendiskussion qualifiziert und erneut erprobt. Ergebnisse zur Itemanalyse und Testgüte sollen Hinweise zur Tauglichkeit des Ansatzes liefern. Sie werden außerdem unter der Fragestellung diskutiert, welchen Einfluss Veränderungen bzgl. der sprachlichen, formalen und methodischen Gestaltung auf Beurteilungsprozesse haben.

Diese Erkenntnisse werden im dritten Schritt zur Konstruktion und Validierung neuer Vignetten und zur Überarbeitung von Items und Instruktion genutzt. Es werden anders als in den

ersten Entwicklungsschritten mehrere Vignetten eingesetzt, die Situationen in unterschiedlichen Unterrichtsfächern und Arbeitsformen und -phasen darstellen. Dies rückt die Anforderung der situationsspezifischen Fähigkeit des Erkennens von Lernbarrieren näher an die realen Anforderungen in der Praxis, da Barrieren über verschiedene Situationen hinweg erkannt werden müssen. Neben der Vignetten- und Itementwicklung rückt damit auch die Art und Weise von Beurteilungsprozessen in den Blick. Itemanalysen und Verfahren zur Prüfung der Testgüte werden erneut mit einer Expertendiskussion ergänzt, um das Verfahren nochmals zu überarbeiten.

Im abschließenden vierten Entwicklungsschritt soll die Testgüte nochmals verbessert werden. Gefragt wird außerdem nach Einflüssen der Lesenden auf Beurteilungsprozesse. Es sollen für das Verfahren relevante Erkenntnisse gewonnen werden, welche anderen Aspekte – außer unterschiedlichen Fähigkeitsausprägungen – eine Rolle für die Wahrnehmung und Beurteilung von Lernbarrieren spielen.

I Theoretischer Hintergrund und Stand der Forschung

2 Was sind Lernbarrieren?

*Barriere, die: Absperrung, die jemanden, etwas von etwas fernhält
(Dudenredaktion, o. D.a)*

Das Behindertengleichstellungsgesetz (BGG, 2002) soll für alle Menschen in Deutschland die Grundlage für ein barrierefreies Leben schaffen. Laut § 4 sind gestaltete Lebensbereiche dann barrierefrei, „wenn sie für Menschen mit Behinderungen in der allgemein üblichen Weise, ohne besondere Erschwernis und grundsätzlich ohne fremde Hilfe auffindbar, zugänglich und nutzbar sind. Hierbei ist die Nutzung behinderungsbedingt notwendiger Hilfsmittel zulässig“.

Schulen sind dabei nicht explizit erwähnt, können aber zu Einrichtungen wie im Folgenden beschrieben gezählt werden: Unter „auffindbar, zugänglich und nutzbar“ wird verstanden, dass „Einrichtungen und Informationen ... nicht nur (z. B. von blinden Menschen und Menschen mit Sehbehinderungen) gefunden und (z. B. stufenlos mit dem Rollstuhl) erreicht werden, sondern auch sinnvoll genutzt werden können (z. B. indem Informationen auch für sinnesbeeinträchtigte Menschen verfügbar sind)“ (Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderung, o. D.b)

Barrierefreiheit bezieht sich hier also in erster Linie auf bauliche Maßnahmen und die Aufbereitung von Materialien, also auf eine Anpassung der physischen Umwelt an individuelle Bedürfnisse. Sind Schulen und Klassenräume sowie dort genutzte Materialien „auffindbar, zugänglich und nutzbar“ (ebd.), ist dies jedoch keine hinreichende Voraussetzung dafür, dass allen Schülerinnen und Schülern ein barrierefreies Lernen gelingen kann.

Bezieht man den Begriff Barrierefreiheit auch auf Unterricht, so rücken Lehr-Lernprozesse in den Fokus, die von mehr als den individuellen Voraussetzungen und der physischen Lernumgebung abhängen. Vielmehr werden diese von einem komplexen System von Faktoren bestimmt, zu denen auch Aspekte der Unterrichtsgestaltung und der Interaktion zwischen Lehrenden und Lernenden gehören (Ricken, 2017). Demnach sind auch Hindernisse in Lehr-Lernprozessen – hier als *Lernbarrieren* verstanden – systemisch zu verorten.

Bisher finden sich kaum Ansätze, in denen sich der Begriff Lernbarriere auf Interaktionen oder Bedingungen der Lehr-Lernsituation im Unterricht bezieht. Verwendung findet der Begriff Lernbarriere z. B. im Kontext von organisatorischem Lernen und Organisationsentwicklung (z. B. Gebert & Boerner, 1997; Schüppel, 1996) in dem Sinne, dass dysfunktionale Prozesse Entwicklung und Wachstum einer Organisation hemmen. Hinsichtlich schulischen Lernens wird der Begriff Lernbarriere z. B. bei Mayer (2016) oder Seiffert (2016) im Sinne innerer, durch sprachliche

Entwicklungsstörungen verursachter Lernbarrieren genutzt, die sich z. B. auf den Erwerb mathematischer Kompetenzen auswirken.

Von Benachteiligungen ist in den Standards für die Lehrerbildung der Kultusministerkonferenz die Rede. In Kompetenzbereich 4 wird als Teil der praktischen Ausbildung beschrieben: „Die Absolventinnen und Absolventen erkennen Benachteiligungen und realisieren pädagogische Hilfen und Präventionsmaßnahmen (...)“ (Kultusministerkonferenz, 2004, S. 9). Diese Benachteiligungen beziehen sich nach dem Verständnis der Standards für die Lehrerbildung auf individuelle Voraussetzungen, nämlich auf die pädagogische, soziologische, psychologische, kulturelle und geschlechtsspezifische Entwicklung und deren Einflüsse auf Lernprozesse.

In anderen Zusammenhängen wird der Begriff Lernbarriere im Sinne von Lernblockade, Lernhemmnis, Lernstörung, Lernverweigerung oder auch Lernwiderstand genannt: „Diese Wörter beschreiben einen Zustand der ‚Minderleistung‘ beim absichtsvollen Lernen ... Er äußert sich darin, dass das gewünschte Wissen, Können und Verhalten ... nicht in ausreichender Qualität, nicht mit ausreichender Sicherheit sowie nicht in der dafür vorgesehenen Zeit erworben wird“ (Lauth et al., 2014, S. 17).

Lernbarrieren werden hier also als erschwerende Bedingungen verstanden, die in den individuellen Voraussetzungen liegen, bzw. auch als Zustand, der beschreibt, dass ein erwartetes Lernziel nicht erreicht wird. Gründe dafür können beispielsweise motivationale Zustände, Misserfolgserwartungen, soziale Probleme sowie physische oder psychische Beeinträchtigungen sein (Lauth et al., 2014).

Individuelle Lernvoraussetzungen sind ein wichtiger Faktor für den Lernerfolg, so dass Lehrkräfte darüber Kenntnis besitzen müssen und es zu ihren Aufgaben gehört, durch fortlaufende Diagnostik individuelle Voraussetzungen zu berücksichtigen (Barth, 2017). Zu den Lernvoraussetzungen gehören sowohl grundsätzliche persönliche Merkmale wie Vorkenntnisse, Intelligenz, die sprachliche Entwicklung oder soziale und affektive Voraussetzungen (Leuders & Prediger, 2016) als auch auf den konkreten Unterricht bezogene Merkmale des Kindes, die vor allem Kognition, Emotion und Motivation betreffen (Barth, 2017; Hartig & Klieme, 2006). Für den Lernerfolg maßgebliche Faktoren wie aktive Lernzeit, Anstrengung, Aufmerksamkeit oder Lernstrategien (Barth, 2017) hängen aber nicht nur von individuellen Lernvoraussetzungen ab, sondern auch von der Gestaltung von Unterrichtsprozessen (Seidel, 2003b), so dass Ursachen für Störungen in Lehr-Lernprozessen – also Lernbarrieren – nicht allein auf Seiten der Schülerinnen und Schüler gesucht werden können.

Beukelman und Mirenda (1998) unterscheiden in ihrem *Partizipationsmodell*, das vor allem in der Intervention mit unterstützter Kommunikation Anwendung findet, zwischen (1) *Zugangsbarrieren* und (2) *Gelegenheitsbarrieren*. Das Modell hat dabei bzgl. der (1) Zugangsbarrieren nicht nur das Kind mit seinen individuellen Fähigkeiten und Einstellungen im Blick, sondern beschreibt – im Einklang mit dem Partizipationsverständnis der *International Classification of Functioning, Disability and Health* (World Health Organization, 2001) – (schulische) Barrieren im Zusammenhang mit Umgebungsfaktoren wie Sitzplatz oder Tafelposition. Auch die Einstellungen der beteiligten Personen (also auch der Lehrkräfte) wie Fähigkeit und Wille, sich auf eine Intervention einzulassen, werden als Gelingensbedingungen bzw. mögliche Barrieren beschrieben. In Bezug auf den Unterricht betreffen Zugangsbarrieren also neben den Fähigkeiten des Kindes auch die *Interaktion* zwischen Lehrkraft und Kind sowie zwischen Umgebungsfaktoren und Kindern. (2) Als Gelegenheitsbarrieren werden alle Faktoren beschrieben, die unabhängig von den individuellen Voraussetzungen des Kindes hinderlich wirken. Beukelman und Mirenda (1998) nennen als Beispiel politische und soziale Gegebenheiten, eingefahrene Praktiken oder Wissen sowie Fähigkeiten und Einstellungen beteiligter Personen. Fähigkeiten und Einstellungen werden bei Beukelman und Mirenda (1998) sowohl den Zugangs- als auch den Gelegenheitsbarrieren zugeordnet. Die Zuordnung lässt sich dahingehend unterscheiden, ob es sich um Wissen, Fähigkeiten und Einstellungen bzgl. individueller Bedingungen des Kindes handelt, die somit die Interaktion betreffen, oder um Wissen, Fähigkeiten und Einstellungen bzgl. Bedingungen, die unabhängig vom Kind sind.

Potenzielle, die Partizipation einschränkende Lernbarrieren entstehen nach diesem Verständnis also innerhalb eines komplexen Gefüges, zu dem die Gestaltung von Unterricht unbedingt dazugehört. Die Beschreibung und Operationalisierung konkreter Lernbarrieren in Unterrichtsprozessen stellt ein erstes Forschungsdesiderat dar.

Der Begriff *Barrieren* bezieht sich bislang meist auf die physische Umwelt, der Begriff *Lernbarrieren* auf individuelle Bedingungen, die für das Lernen hinderlich sind. Lernbarrieren, die die Gestaltung von Lehr-Lernprozessen betreffen, sind bislang dagegen kaum beschrieben. Zur systemischen Analyse problematischer Lehr-Lernsituationen müssen Lernbarrieren systematisiert und operationalisiert werden.

3 Was bedeutet „Erkennen“?

erkennen, unregelmäßiges Verb:

(1) so deutlich sehen, dass jemand weiß, wen oder was er vor sich hat;

(2) aufgrund bestimmter Merkmale ausmachen, identifizieren;

(3) Klarheit über jemanden, etwas gewinnen; richtig einschätzen

(Dudenredaktion, o. D.b)

Das Inkrafttreten der UN-Behindertenrechtskonvention und die Umsetzung der Inklusion in Deutschland erfordern es, die damit einhergehenden Anforderungen an Lehrkräfte zu beschreiben und entsprechend in die Lehrkräfteausbildung aufzunehmen (Baumert & Kunter, 2006, 2011; Blömeke et al., 2015a; Frey, 2008; Heran-Dörr et al., 2007; Lindmeier, 2011; Lütjeklöse & Neumann, 2018; Terhart, 2007). Neben gesellschaftlichen Entwicklungen und Bildungsreformen macht eine verstärkte Zuwanderung des Unterrichten heterogener (Agarwala, 2020; von Dewitz et al., 2016; Morris-Lange et al., 2016; Otto et al., 2016; Vogt & Rogalla, 2009). Die zunehmende Beanspruchung im Lehrerberuf einhergehend mit hohen Ausstiegswerten zeigen die Dringlichkeit, angehende Lehrkräfte adäquat auf ihre vielfältigen Aufgaben vorzubereiten (Bönte et al., 2019; Dicke et al., 2015; Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD], 2016). Damit rückt verstärkt in den Blick, dass es zu den grundlegenden Kompetenzen gehört, eine große Varianz an Lernprozessen beobachten und pädagogisch begleiten zu können. Dieses Beobachten von Lernprozessen ist – in allen, also auch weniger heterogenen Klassen – eine Voraussetzung für faire Bildungschancen. Wie das Erkennen von Lernbarrieren damit zusammenhängt, soll in den folgenden Ausführungen gezeigt werden.

3.1 Die Basis: Das Professionelle Wissen

Es liegen zahlreiche Untersuchungen zur professionellen Kompetenz von Lehrkräften vor, die sich auf das *Professionelle Wissen* als eine grundlegende Komponente beziehen (Baumert & Kunter, 2011; Blömeke et al., 2008; Kleickmann et al., 2014). Es ist ein Ordnungssystem entstanden, das drei wesentliche Bereiche des professionellen Wissens unterscheidet: (1) Fachwissen, das vertieftes (Hintergrund-)Wissen über das zu unterrichtende Fach beinhaltet, (2) fachdidaktisches Wissen, das sich auf Kenntnisse über das Vermitteln der fachlichen Inhalte bezieht sowie (3) pädagogisch-psychologisches Wissen, das Wissensanteile vereint, die für die Unterrichtstätigkeit von Bedeutung sind (Baumert & Kunter, 2011; Blömeke & Suhl, 2010; Helmke, 2009; Lipowsky, 2006). Darüber hinaus werden teils unterschiedliche Wissensbereiche ergänzt wie z. B. das Beratungs- und Organisationswissen im COACTIV-Modell (Baumert & Kunter, 2011).

Während über die genannten drei Bereiche professionellen Wissens grundsätzlich Einigkeit besteht, fällt die genauere Konzeptualisierung des pädagogisch-psychologischen Wissens unterschiedlich aus. Es werden unterschiedliche Begriffe wie pädagogisches (Professions-)Wissen (König & Seifert, 2012), pädagogisches Unterrichtswissen (König & Blömeke, 2010), pädagogische Kompetenz (Seifert et al., 2009), erziehungswissenschaftliches Wissen (Blömeke et al., 2008), bildungswissenschaftliches Wissen (Kunina-Habenicht et al., 2013; Terhart et al., 2012) oder der auch in dieser Arbeit verwendete Begriff pädagogisch-psychologisches Wissen (Voss & Kunter, 2011; Voss et al., 2014) genutzt. Voss et al. (2015) geben einen Überblick über verschiedene Definitionen. Diese beziehen sich vor allem auf die Gestaltung des Lehr-Lernprozesses und eine effiziente Klassenführung, zum Teil aber auch auf die Planung von Unterricht. Shulman (1987) versteht unter pädagogischem Wissen vor allem Kenntnisse zur KF. Andere Ansätze wie beispielsweise von Baumert und Kunter (2006) vereinen verschiedene Dimensionen wie Konzeptions- und Planungswissen, Wissen über Unterrichtsführung und fachübergreifende Prinzipien des Diagnostizierens. Während beispielsweise Borko und Putnam (1996) auch *beliefs* zum pädagogisch-psychologischen Wissen zählen, werden in späteren Konzeptualisierungen *beliefs* und Wissen als unterschiedliche Aspekte identifiziert (Baumert & Kunter, 2006; Blömeke et al., 2008; Weinert, 2001a).

Die Abhängigkeit von Schülerleistungen vom professionellen Wissen ist sowohl für das pädagogisch-psychologische Wissen (Lenske et al., 2016; Voss et al., 2011) als auch für das Fachwissen und fachdidaktische Wissen belegt (Baumert et al., 2010; Kunter et al., 2011; Kunter et al., 2013). Es wird deutlich, dass das pädagogisch-psychologische Wissen dabei eine grundlegende Rolle spielt und eine fachübergreifende Voraussetzung darstellt. Jedoch führt Wissen allein nicht unmittelbar zu einem qualitätvollen und barrierefreien Unterricht, sondern muss im Zusammenhang mit weiteren Fähigkeiten als Teil eines komplexeren Kompetenzmodells verstanden werden.

3.2 Die Schlüsselkompetenz: Situationsspezifische Fähigkeiten

Der Kompetenzbegriff wurde im Zusammenhang mit Lehrkraftkompetenzen vielfach diskutiert und modelliert (Baumert & Kunter, 2006; Baumert & Kunter, 2011; Blömeke et al., 2015a; Frey, 2008; Heran-Dörr et al., 2007; Klieme & Hartig, 2007; Lindmeier, 2011; Prenzel et al., 2007). Das für die Studie COACTIV entwickelte Kompetenzmodell vereint professionsspezifische Erkenntnisse (z. B. Bransford et al., 2005; Bromme, 1992; Bromme, 1997; Bromme, 2008; Shulman, 1986; Shulman, 1987) mit Ergebnissen aus der Kompetenzforschung und Kompetenzdiagnostik (Baumert & Kunter, 2011), um konkrete Anforderungen an Lehrkräfte zu beschreiben.

Grundlegend für das COACTIV-Modell ist die Definition von Weinert (2001b), nach der Kompetenzen verstanden werden als

...die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbunden motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können. (S. 27)

Wesentliche Aspekte dieses Ansatzes sind zum einen das Verständnis von Kompetenz als etwas, das erlernt und entwickelt werden kann, und zum anderen die Zusammenführung unterschiedlicher Voraussetzungen mit der Nutzung in variablen Situationen.

Als Kern der Lehrtätigkeit wird das Vorbereiten und Durchführen von Unterricht betrachtet mit dem Anspruch, auch in heterogenen Klassen alle Kinder im Blick zu haben und individuelle Lernziele zu erreichen. Die Gestaltung des Unterrichts ist dabei – in inklusiven Klassen mehr denn je – von Unvorhersehbarkeit gekennzeichnet, die in hohem Maße Flexibilität erfordert. Im Zusammenhang mit Lehrkompetenz wird deshalb auch der Begriff der adaptiven Lehrkompetenz verwendet (Fischer et al., 2014; Kufner, 2014; Martschinke, 2015). Da Situationen variieren, kann auch die Handlungskompetenz in Abhängigkeit dieser Situationen variieren (Maag Merki & Grob, 2005), so dass sich Kompetenz nicht nur über die Zeit verändern kann (z. B. durch Lernzuwachs), sondern auch situationsabhängig variabel ist.

Deshalb sind nach der Konzeptualisierung von Fähigkeiten und Fertigkeiten, die dem Wissensbereich zuzuordnen sind, auch zunehmend *situationsspezifische Fähigkeiten* in den Blick genommen worden, die unterschiedlich benannt und konzeptualisiert werden. Zu den Konzepten gehören z. B. *ability to analyze teaching* (Santagata et al., 2007), *teacher noticing* (Jacobs et al., 2010) oder *professional vision* (Sherin & van Es, 2009) bzw. *professionelle Wahrnehmung* (Seidel et al., 2010). Unter dem ursprünglichen Begriff *professional vision* beschreibt Goodwin (1994), wie Situationen fachunabhängig professionell erschlossen werden, nämlich durch das Zuordnen von Beobachtungen zu vorhandenem professionellem Wissen (*coding*), dem Hervorheben relevanter Aspekte in komplexen Situationen (*highlighting*) sowie dem Generieren und Veranschaulichen von Repräsentationen dieser relevanten Aspekte (*producing and articulating material representations*). Später wurde dieses Konzept von Sherin (2001) auf Unterrichtssituationen angewendet.

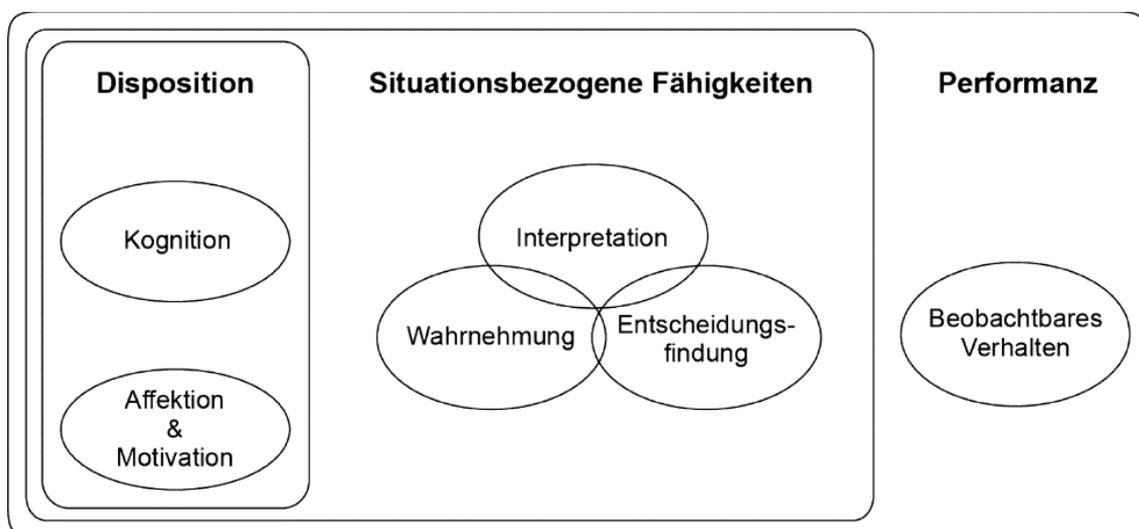
Die Verarbeitung wahrgenommener Unterrichtssituationen beschreiben Sherin und van Es (2009, S. 24) mit drei unterschiedlichen Facetten des *reasoning*: (1) Beschreiben bedeutet, dass unterrichtsrelevante Aspekte zunächst nur genannt werden. (2) Bewertungen enthalten zudem

Aussagen über die Qualität der Beobachtung. Darüber hinaus sind für (3) Interpretationen Rückschlüsse über Wirkmechanismen nötig, also eine komplexere Verarbeitung der beobachteten Aspekte (Sherin, 2007; Sherin & van Es, 2009). In weiteren Konzeptualisierungen werden auch andere Prozesse beschrieben, z. B. das Generieren von Handlungsalternativen (Santagata & Yeh, 2016). Die Arbeitsgruppe um Seidel und weiteren Forschenden unterscheidet drei miteinander verbundene Teilfähigkeiten ähnlich zu Sherin und van Es (2009): Beschreiben, Erklären und Vorhersagen (Seidel et al., 2010; Seidel & Stürmer, 2014; Stürmer & Seidel, 2017).

Das Modell von Blömeke et al. (2015a) vereint diese Aspekte und stellt Kompetenzen von Lehrkräften als Kontinuum dar: Affektiv-motivationale Merkmale und professionelles Wissen sind als Dispositionen vorhanden. Zu den Voraussetzungen gehört zum einen professionsspezifisches Wissen, das deklarative und prozedurale Anteile (Wissen und Können) enthält, sowie motivationale und affektionale Dispositionen, zu denen Werte, Überzeugungen, subjektive Theorien und normative Präferenzen und Ziele gehören sowie Fähigkeiten der professionellen Selbstregulation (Baumert & Kunter, 2011). Sie bilden die Grundlage für die situationspezifischen Fähigkeiten, unterrichtsrelevante Ereignisse wahrzunehmen und wissensgeleitet zu interpretieren und so Entscheidungen für das weitere Vorgehen zu treffen, das dann in beobachtbarem Handeln sichtbar wird (siehe Abbildung 1). Das Wahrnehmen und Interpretieren lernrelevanter Situationen wird dabei als die Schlüsselkompetenz von Lehrkräften betrachtet (Blömeke et al., 2015b; Blömeke et al., 2022; Holodynski et al., 2017; Kersting et al., 2012; Kramer et al., 2017; Meschede et al., 2015; Roth et al., 2011; Seidel & Stürmer, 2014; Stürmer & Seidel, 2015; Sherin et al., 2011a und 2011b).

Abbildung 1

Das Kompetenzmodell der Transformation von Kompetenz in Performanz (Blömeke et al., 2015a, S. 7)



Das professionelle Wahrnehmen lernrelevanter Situationen ist oft anspruchsvoll, weil sich Wirkungsmechanismen nicht immer offensichtlich an der Oberfläche zeigen, sondern auch Tiefenstrukturen des Unterrichts betreffen (Seidel & Prenzel, 2007). Im Unterricht müssen Lehrpersonen dazu eine Vielzahl zum Teil simultan ablaufender Handlungen im Blick haben, die unterschiedlich relevant für die Lehr-Lernsituation sind und dementsprechend nicht alle gleichermaßen berücksichtigt werden müssen und können (Doyle, 1986; Sabers et al., 1991; Sherin, 2007; Sherin & Star, 2011). Vielmehr kommt es darauf an, diejenigen Aspekte professionell wahrzunehmen, die im jeweiligen Moment für den Lernprozess relevant sind, also das Lernen entweder fördern oder – im Sinne einer Lernbarriere – hemmen (Meschede, 2014; Seidel & Stürmer, 2014; Teuscher et al., 2017).

Die in dieser Arbeit mit dem Begriff Erkennen bezeichnete Kompetenz lässt sich in der ursprünglichen Bedeutung des Wortes also gut in diesem Kompetenzmodell verorten. Wissen über Lernbarrieren muss genutzt werden, um diese im jeweiligen Kontext aufgrund bestimmter Merkmale auszumachen und zu identifizieren. Lernbarrieren müssen so deutlich gesehen werden, dass die Lehrkraft weiß, was sie vor sich hat (*wissensgeleitetes Wahrnehmen*), und schließlich müssen sie hinsichtlich ihrer Auswirkung auf den Lehr-Lernprozess interpretiert werden, so dass die Lehrkraft Klarheit über die Situation gewinnt und diese richtig einschätzt.

Die Fokussierung auf relevante Aspekte bedeutet im Kontext des Erkennens von Lernbarrieren, sich auf lernhemmende Unterrichtsbedingungen und -prozesse zu konzentrieren. Diese spezifische Fokussierung, Wahrnehmung und Interpretation entsprechen einer diagnostischen Aufgabe, die im Folgenden näher betrachtet werden soll.

3.3 Das Diagnostische daran: Erkennen als diagnostische Kompetenz

Galten diagnostische Ergebnisse im Schulkontext lange Zeit als Rechtfertigung für die Klassifikation von Kindern oder für schulorganisatorische Maßnahmen, besteht nun Einigkeit darüber, dass der kindbezogene diagnostische Blick um eine Analyse der gesamten Lehr-Lernsituation erweitert werden muss (Langner & Jugel, 2019; Ricken, 2017; Ricken & Schuck, 2011). Zum diagnostischen Gegenstand gehört somit auch das Unterrichtsgeschehen mit all seinen Gestaltungsmerkmalen, Handlungen und Interaktionen:

Ein wohlwollender und verantwortungsvoller diagnostischer Blick, der das Individuum wahrnimmt, erkennt und anerkennt, ist durch eine konsequente systemische Diagnostik unterrichtlicher und schulischer Lernbedingungen zu ergänzen ... Dieser Logik folgend wäre ggf. einer Schule bzw. einem unterrichtlichen Angebot ein Förder- bzw. Entwicklungsbedarf zu attestieren und nicht den Schüler*innen. (Kopmann, 2022, S. 25)

Als besonders geeignete diagnostische Methode zur Analyse von Unterrichtssituationen gilt das Beobachten. Es verlangt neben methodischen Fertigkeiten auch Wissen über den jeweiligen spezifischen, diagnostischen Gegenstand, die wissensgeleitete Interpretation und Reflexion des Beobachteten sowie eine Einschätzung der möglichen Auswirkungen und zukünftigen Entwicklungen (Ingenkamp & Lissmann, 2008; Karst, 2017; Kopmann, 2022). Die diagnostische Methode des Beobachtens entspricht damit stark der PW (Tönnissen et al., 2023). Aus sonderpädagogischer Perspektive ist insbesondere das Beobachten von problematischen Situationen relevant. Es ermöglicht eine genaue Beschreibung von Lern- und Entwicklungsprozessen als Grundlage für Hypothesenbildungen und Veränderungen. Im weiteren Verlauf dient es der Evaluation von Veränderungen und ist damit Bedingung für jede prozessorientierte Diagnostik und Förderung sowie für die Passung individueller und unterrichtlicher Bedingungen (Ricken, 2017). Das Erkennen von Lernbarrieren im Lehr-Lernprozess ist somit eine wichtige diagnostische Kompetenz für die Gestaltung barrierefreien Unterrichts.

Allerdings zeigt die Praxis, dass das Beobachten von Unterrichtsbedingungen keineswegs selbstverständlich ist (Kopmann, 2022; Tönnissen et al., 2023). Novizinnen und Novizen neigen dazu, die Ursache für Schulschwierigkeiten und Schwierigkeiten im sozial-emotionalen Bereich dem Kind zuzuschreiben und Umweltfaktoren auszublenden (Thiel, 2021). Aber auch erfahrene Lehrkräfte verstehen Diagnostik oft als Aufgabenbereich fachlich versierter Expertinnen und Experten, die sich auf die individuellen Voraussetzungen des Kindes fokussieren (Kopmann, 2022). Dies zeigt die Notwendigkeit einer Erweiterung des diagnostischen Gegenstands.

Zum einen müssen also, wie in Kapitel 1 gezeigt, Lernbarrieren systematisiert und operationalisiert werden, um sie beobachtbar zu machen. Die Modellierung eines spezifischen Kompetenzmodells zum Erkennen von Lernbarrieren stellt ein weiteres Desiderat dar. Die Notwendigkeit der Ausbildung dieser Kompetenz bringt auch mit sich, dass Verfahren entwickelt werden müssen, die diese Kompetenz erfassen, um Ausbildungsangebote entwickeln und evaluieren zu können. Dies ist das Hauptanliegen dieser Arbeit. Das folgende Kapitel gibt zunächst einen Überblick über bereits vorhandene Studien zur Erfassung der PW.

Die situationsspezifische Fähigkeit, relevante Situationen im Unterricht professionell wahrzunehmen, gilt als Schlüsselkompetenz von Lehrkräften. Das Erkennen von Lernbarrieren ist dieser Kompetenz zuzuordnen. Die spezifische Fokussierung auf lernhemmende Aspekte der Unterrichtsgestaltung betont die diagnostische Relevanz dieser Kompetenz. Dennoch fehlen bislang eine Modellierung dieser Kompetenz sowie Verfahren zur Erfassung.

4 Forschungsstand zur Erfassung der professionellen Wahrnehmung

Vor dem Hintergrund der Ergebnisse von Schülerleistungsstudien wie im *Trends in International Mathematics and Science Study* (TIMSS) oder im *Program for International Student Assessment* (PISA), aber auch angesichts struktureller Reformen der Lehrerbildung und der Herausforderung, Unterricht inklusiv zu gestalten, ist in den vergangenen Jahren verstärkt gefordert worden, professionelle Kompetenzen von Lehrkräften nicht nur zu beschreiben, sondern auch messbar zu machen (Kunter & Klusmann, 2010). Während deklaratives Wissen gut mit Wissenstests und Einstellungen mit subjektiven Maßen wie der Selbstbeurteilung erfasst werden können, sind die Anforderungen an Verfahren zur Erfassung der PW in konkreten Situationen anspruchsvoller (Maag Merki & Werner, 2011). Da situationsspezifische Fähigkeiten erfasst werden sollen, müssen Teststimuli diese Situations- und Kontextgebundenheit abbilden (Borko, 2004). Die Beobachtung realer Unterrichtssituationen (Helmke, 2009) ist als Testverfahren nicht praktikabel, so dass Unterrichtssituationen simuliert werden müssen, die die Dynamik und Komplexität von Unterricht authentisch darstellen (Ophardt & Thiel, 2017). Deshalb werden in Studien üblicherweise Fallvignetten eingesetzt.

4.1 Einsatz von Vignetten als Stimulus

Ursprünglich versteht man unter einer Vignette entweder eine Randverzierung in der Buchkunst oder eine Gebührenmarke, die an der Windschutzscheibe eines Autos anzubringen ist. In der Medizin und Psychologie wird darunter ein beispielhafter Fallverlauf verstanden, der als Anschauungs- oder Lehrmaterial dient (Schnurr, 2003). Vignettenanalysen werden seit den 1950er Jahren vor allem in der Soziologie eingesetzt, um Werthaltungen, Einstellungen und Urteile von Probandinnen und Probanden zu erheben. Dazu werden Fallbeschreibungen – oder allgemeiner ausgedrückt Merkmalsausprägungen (Schnurr, 2003) – in Form von Situations-, Interaktions- oder Personenbeschreibungen genutzt, die anhand von Interviews oder Skalen beurteilt werden sollen bzw. als Stimulus dienen, situationsangemessene Handlungsmöglichkeiten zu benennen und zu begründen (Brovelli et al., 2013; Jacobs, 2017; Rosenberger, 2016; Schnurr, 2003; Stürmer & Seidel, 2017; Zucker, 2019).

Vignetten finden unter anderem dann Verwendung, wenn Wahrnehmungsschemata sowie Interpretationen und Beurteilungen der Probandinnen und Probanden untersucht werden und implizites Wissen rekonstruiert werden sollen (Rosenberger, 2016; Schnurr, 2003), so dass sie als angemessenes Instrument für die Erhebung der PW betrachtet werden können:

Vignetten liefern die rudimentäre Version eines Falls, die erst unter dem professionellen Blick und unter dem fallkonstruierenden Zutun eines Professionals seine „geschlossene Gestalt“ erhält, indem er im Kontext der spezifischen Relevanz- und Beobachtungsmuster der Profession/eines Professionellen typisiert und in den Zustand eines bearbeitbaren Falls transformiert wird. Eines der wichtigsten Einsatzgebiete von Vignetten ist daher die Erschließung der impliziten Wahrnehmungs-, Beurteilungs- und Entscheidungskriterien von Professionals. (Schnurr, 2003, S. 394)

Insbesondere bei Forschungsdesigns, bei denen verglichen werden soll, wie Personen mit unterschiedlichem Ausbildungs- und Erfahrungshintergrund einen identischen Fall beurteilen, kommen Vignetten zum Einsatz (Schnurr, 2003). Vignetten werden dabei so konstruiert, dass sie spezifische Anforderungssituationen im Sinne der zu erfassenden Dimensionen abbilden (Kramer et al., 2017; Ollesch et al., 2018; Vogler, 2019). Die in den Vignetten dargestellten Situationen können auf unterschiedliche Weise beurteilt, zugeordnet, beschrieben oder interpretiert werden. Es kann auch nach Handlungsoptionen und ihren Begründungen gefragt werden. Die Ergebnisse müssen Rückschlüsse auf die konzeptuell zugrunde gelegten situationsspezifischen Fähigkeiten zulassen (Krammer, 2014; Weber et al., 2020). Der Einsatz realitätsnaher und handlungsorientierter Vignetten und deren Analyse kann zu einer Verzahnung von Ausbildung und Praxis beitragen und prozedurales Wissen stärker in den Fokus der Lehrkräfteausbildung rücken (Bönte et al., 2019; König et al., 2014, Rutsch et al., 2017; Terhart et al., 2012).

Möglich ist dies grundsätzlich mit jeder Form von Vignette wie Video, Text, Bild oder Comic. Die Komplexität und Simultanität von Unterricht wird meist durch den Einsatz authentischer (z. B. Barth, 2017; Keppens et al., 2019; König, 2015; Wöhlke, 2020) oder inszenierter (z. B. Bönte et al., 2019; Hoppe et al., 2020; Scholten et al., 2019; Wiedmann, 2015) Videos, seltener auch (ergänzend oder ausschließlich) durch schriftliche Unterrichtsbeschreibungen (z. B. Brovelli et al., 2013; Hoppe et al., 2020; Zucker, 2019) abgebildet.

4.2 Vignettenbasierte Verfahren

Objektive Verfahren mit proximalen Beurteilungskriterien finden sich vor allem im deutschsprachigen Raum (Voss et al., 2015). Auch im Rahmen der *Qualitätsoffensive Lehrerbildung* sind in den vergangenen Jahren Studien durchgeführt worden, die sich der objektiven Erfassung situationsspezifischer Fähigkeiten widmen. Dabei finden sich Unterschiede in den Konzeptualisierungen, untersuchten Dimensionen, Zielgruppen und methodischen Ansätze. Ein deutlicher Schwerpunkt ist bei qualitativen Studien mit Videovignetten zu naturwissenschaftlichen Fächern auszumachen.

Schwerpunkte sind neben der eigentlichen Entwicklung von Verfahren auch die Untersuchung der Struktur und Genese situationsspezifischer Fähigkeiten sowie Zusammenhänge mit anderen Konstrukten oder die Ausbildung und Förderung der PW. Eine Erfassung dieser Fähigkeit ist aber in jedem dieser Fälle erforderlich, auch wenn das *Wie* nicht immer im Vordergrund steht.

Exemplarisch zeigt Tabelle 1 einige Studien der vergangenen Jahre, die die PW mithilfe von Vignetten erfasst haben. Recherchiert wurden diese Studien mithilfe einschlägiger im Internet verfügbarer Datenbanken und anhand der Suchbegriffe *Vignetten* sowie *professionelle Wahrnehmung*. Ausgehend von frühen Arbeiten von Santagata et al. (2007) und Seidel und Prenzel (2007) wurden Studien ab 2007 berücksichtigt. Ausgewählt wurden Studien, die die PW von Lehramtsstudierenden oder Lehrkräften im Schulunterricht erfassten. Die Tabelle stellt die Studien sortiert nach dem Hauptanliegen der jeweiligen Studien dar. Einige Studien verfolgen mehrere Ziele (siehe Tabelle 1).

Tabelle 1

Übersicht über Studien zur vignettenbasierten Erfassung professioneller Wahrnehmung

	Autorinnen und Autoren (Jahr), ggf. Projekt	Gegenstand der Untersuchung	Stichprobe	erfasste Fähigkeiten	Unterrichtsfach	Vignettenformat	Methode	zentrale Ergebnisse
	Entwicklung von Verfahren zur Erfassung der PW							
1	Wöhlke (2020)	Entwicklung eines Instruments zur Erfassung der physikspezifischen professionellen Unterrichtswahrnehmung in Bezug auf Schülervorstellungen und Erkenntnisgewinn	Studierende Lehramt	Wahrnehmen, Interpretieren, Entscheiden	Physik	Video	geschlossen (Multiple-Choice) mit Option offen zu antworten	Test sensibel für unterschiedliche Studiengänge, aber nicht für unterschiedliche Semester
2	Keppens et al. (2019)	Entwicklung eines Instruments zur holistischen Erfassung des Unterrichtens in inklusiven Klassen in Bezug auf Lehrkraft-Lernenden-Interaktion und Differenzierung	Studierende Lehramt Grundschule	Wahrnehmen, Interpretieren	verschiedene	Video	<i>comparative judgement</i> , geschlossen	Holistische Erhebung der PW möglich, es lassen sich zwei Dimensionen (Interaktion und Differenzierung) abbilden
3	Zucker (2019)	Entwicklung eines Instruments zur Erfassung der PW von formativem Assessment (unterrichtsbegleitende Diagnostik von Lernständen und Rückmeldungen)	Studierende Lehramt	Erkennen, Beschreiben	Naturwissenschaften	Text und Video	offen, geschlossen	Zwei Faktoren Diagnostik und Rückmeldung bei offenen Fragen, kein passendes Modell für geschlossene Items, sensibel für Expertiseunterschiede zwischen 1. und 6. Semester

4	Barth (2017)	Entwicklung eines Instruments zur Erfassung der PW von Unterrichtsstörungen	Studierende Lehramt	Wissen, Erkennen, Beurteilen, Handlungsalternativen generieren, Entscheiden, Implementieren	Mathematik	Video	offen	Instrument ist valide und reliabel, Expertiseunterschiede zwischen Bachelor- und Masterstudierenden, Erkennen und Beurteilen korrelieren stark miteinander
5	König & Lebens (2012), König (2015), König & Kramer (2016) Classroom management Expertise (CME)	Entwicklung eines Instruments zur Erfassung von Klassenführungsexpertise (Übergänge, Instruktionszeit, Schülerverhalten regeln, Feedback), Zusammenhang zwischen Wissen und Wahrnehmen	Lehrkräfte	Wissen, Wahrnehmen; Genauigkeit der Wahrnehmung, holistische Wahrnehmung, Rechtfertigung von Handlungen	verschiedene	Video	offen, geschlossen (Multiple-Choice)	Instrument ist reliabel, Ein-Faktorenstruktur, deklaratives pädagogisches Wissen und PW sind getrennte, aber korrelierende Konstrukte, PW hängt von Expertise in Form von Praxiserfahrung ab
6	Blömeke et al. (2015b), König et al. (2014), König et al. (2015) Teacher Education and Development Study - Follow Up (TEDS-FU)	Entwicklung und Anwendung eines Instruments zur handlungsnahen Erfassung von allgemeinpädagogischen und mathematikdidaktischen Kompetenzen	Lehrkräfte	Wahrnehmen, Analysieren, Handlungskompetenz	Mathematik	Staged Video	offen, geschlossen	Fachwissen, fachdidaktisches Wissen und pädagogisches Wissen bilden drei Dimensionen; Wahrnehmen ist empirisch von Analysieren und Handlungskompetenz trennbar; professionelle Unterrichtswahrnehmung verbessert sich mit zunehmender Praxiserfahrung

7	Wiedmann (2015), Wiedmann & Leuders (2019)	Entwicklung eines Instruments zur Erfassung der PW von Schülerinnen- und Schülerinteraktion in Phasen kooperativen Lernens	Studierende Lehramt	Beschreiben, Bewerten, Vorhersagen	Mathematik	Staged Video	geschlossen	Instrument erfasst PW und ist sensibel für Ergebnisse von Trainings auf Gruppenebene
8	Wolters (2014)	Entwicklung eines Instruments zur Erfassung der PW kognitiver Aktivierung	Studierende Lehramt, Lehrkräfte	Bemerken, Interpretieren	Naturwissenschaften	Video	geschlossen	Instrument ist reliabel und valide, Lehrkräfte unterscheiden sich signifikant von Bachelorstudierenden
9	Gold et al. (2013), Hellermann et al. (2015), Holodynski et al. (2017) Videobasierte Unterrichtsanalyse (ViU)	Entwicklung und Validierung zweier Instrumente zur Erfassung der PW der KF und LU (Monitoring, prozessuale Strukturierung des Unterrichts, Etablierung eines lernförderlichen Systems von Regeln und Routinen, inhaltliche Strukturierung, kognitive Aktivierung); Förderung der PW	Studierende Lehramt	Wahrnehmen, Beschreiben, Interpretieren	Naturwissenschaften	Video, realer Unterricht	geschlossen	Gute interne Konsistenz sowie Konstruktvalidität, sensitiv für Expertiseunterschiede zwischen Bachelor- und Masterlehramtsstudierenden sowie Lehrpersonen. Die Arbeit mit Vignetten (Video oder Text) im Seminar fördert die PW von KF
10	Brovelli et al. (2013)	Entwicklung eines Instruments zur Erfassung handlungsnaher professioneller Kompetenzen für den naturwissenschaftlichen Unterricht: fachdidaktisches Wissen, pädagogisches Wissen, LU	Studierende Lehramt	Erkennen, Begründen	Naturwissenschaften	Text	offen	Signifikanter Kompetenzanstieg bei Lehramtsstudierenden der Naturwissenschaften im Verlauf des Studiums im Vergleich zu anderen Studierendengruppen (Naturwissenschaften ohne

								Lehramt, Lehramt ohne Naturwissen- schaften)
Förderung der PW (Anzahl und Qualität von Be- obachtungen)								
11	Orschulik (2021)	Förderung der professionellen Unterrichtswahrnehmung mathematikbezogener Unterrichtsereignisse	Studierende in Praxisphasen	Wahrnehmen, Interpretieren, Entscheiden	Mathematik	Video	offen (lautes Denken, leitfadengestützte Interviews)	Spezifische Lerngelegenheiten verbessern die PW hinsichtlich der Anzahl wahrgenommener Ereignisse, höherer Anteil von Interpretationen mit didaktischem Bezug, mehr Entscheidungen mit didaktischem und methodischem Bezug
12	Krosanke (2021)	Förderung der PW sprachlicher Aspekte im mathematischen Lehren und Lernen	Studierende Lehramt	Wahrnehmen, Interpretieren, Entscheiden	Mathematik	Video	offen	Bessere Wahrnehmung und elaboriertere Interpretation der sprachlichen Perspektive durch spezifische Lerngelegenheit
13	Hoppe et al. (2020)	Entwicklung und Erprobung von Methoden zur Förderung unterrichtsbegleitender Diagnostik von Schülerinnen- und Schülervorstellungen, Analyse	Studierende Lehramt	Wahrnehmen, Interpretieren	Biologie	Text und staged Video	offen	Lernen ist effektiver mit Vignetten, unabhängig vom Format, diagnostische Fähigkeiten werden inhaltspezifisch erworben

		von Schülerinnen- und Schüleräußerungen im Hinblick auf ihre Anschlussfähigkeit						
14	Bönte et al. (2019) Classroom Management in Primary and Secondary Schools (CLIPSS)	Entwicklung und Evaluation von videobasiertem Unterrichtsmaterial sowie Erfassung der Kompetenzentwicklung im Bereich des prozeduralen Wissens in KF und des speziellen Wissens und der Reflexion zu nonverbalen Klassenführungsstrategien	Studierende Lehramt, Lehrkräfte	Wissen, Wahrnehmen	verschiedene	Staged Video	geschlossen	Das anwendungsbezogene Klassenführungswissen kann mit Anwendung der Videos signifikant gesteigert werden, die staged Videos werden überwiegend als authentisch beurteilt
15	Sunder et al. (2016)	Förderung der PW von LU	Studierende Lehramt	Wahrnehmen, Interpretieren	Sachunterricht	Video	offen	Förderung durch Intervention ist möglich, besserer Lernzuwachs durch Einsatz von Videos, Verbesserung themenabhängig
16	Sherin & van Es (2009)	Beobachtungsanalyse nach Bezugsperson, Gegenstand (Klassenmanagement, Klassenklima, didaktische Fähigkeit und mathematisches Denkvermögen) und Analyseebene	Lehrkräfte	Beschreiben, Bewerten Interpretieren	Mathematik	Video	offen (Interviews)	Teilnahme am Videoclub führt dazu, dass zunehmend mehr relevante Merkmale erkannt und häufiger interpretierende Analyseebenen eingenommen werden, die selektive Wahrnehmung und das wissensgestützte Schlussfolgern verbessern sich

17	Santagata et al. (2007)	Förderung der Analysefähigkeit bzgl. Elaboration, Präzisierung mittels Situationsbezügen aus dem Videofall, Bezugnahme zum mathematischen Inhalt der Unterrichtsstunde, Lernprozessen sowie kritischer Beurteilung	Studierende Lehramt	Beschreiben, Erklären, Handlungsalternativen generieren	Mathematik	Video	offen, <i>Lesson Analysis Framework</i> (schriftliche Kommentare)	Steigerung der Antwortqualität durch Videotraining, Verbesserung bezüglich des Erkennens von relevanten Merkmalen, des Beurteilens und Findens von Alternativen
Struktur, Genese und Expertiseunterschiede								
18	Scholten et al. (2018, 2019)	Expertiseunterschiede in der geographiespezifischen PW in Bezug auf Kartenkompetenz und geographisches Denken	Lehrkräfte	Wahrnehmen, Interpretieren, Entscheiden	Geografie	Staged Video	offen (lautes Denken)	Unterschied zwischen Expertinnen und Experten und Novizinnen und Novizen in Anzahl der wahrgenommenen Ereignisse, Interpretation und Entscheidungsfindung
19	Meschede et al. (2015)	Erfassung der PW von LU (kognitive Aktivierung, inhaltliche Strukturierung)	Studierende Lehramt	Beschreiben, Interpretieren	Sachunterricht	Video	geschlossen	Eher eindimensionales Modell, Expertiseunterschied zwischen Bachelor- und Masterstudierenden
20	Schäfer & Seidel (2015)	Erfassung des Zusammenhangs von <i>Noticing</i> und <i>Reasoning</i> in Bezug auf Zielklarheit und Unterrichtsklima, Expertiseunterschiede, Vergleich zweier methodischer Ansätze	Studierende Lehramt	Wahrnehmen, Beschreiben, Erklären, Vorhersagen	Physik	Video	geschlossen, offen	Studierende nehmen relevante Ereignisse wahr, haben im Vergleich zum Rating der Expertinnen und Experten aber Probleme

								beim Erklären und Vorhersagen, bessere Ergebnisse für Zielklarheit, positive Zusammenhänge zwischen den beiden methodischen Ansätzen
21	Steffensky et al. (2015) Videobasierte Unterrichts-analyse (ViU)	Zusammenhang zwischen PW von KF und LU	Studierende Lehramt, Lehrkräfte	Beschreiben, Interpretieren	Sachunter-richt	Video	geschlossen	KF und LU sind ver-schiedene aber mitei-ner korrelierende Konstrukte, keine ein-deutig belegbaren Subfacetten für LU, Abhängigkeit der LU von Unterrichtsinhal-ten
22	Voss et al. (2014) Cognitive activa-tion in the math-ematics class-room and professional competence of teachers – Refer-endarisstudie (COACTIV-R)	Erfassung der Entwicklung pro-zeduralen Wissens über die Ef-fizienz von KF und des Zusam-menhangs mit Unterrichtsqualität	Referenda-rinnen und Referendare	Wahrnehmen, Maßnahmen nennen	Mathe	Staged Vi-deo	offen	Signifikanter Zuwachs im Wissen über KF während des Referen-dariats, das pädago-gisch-psychologische Wissen angehender Lehrkräfte sagt statis-tisch signifikant die spätere Effizienz der KF sowie die konstruk-tive LU vorher, keine Bedeutung für das Po-tenzial zur kogniti-ven Aktivierung

23	Seidel et al. (2010), Stürmer & Seidel (2015), Seidel & Stürmer (2014) Observe	Erfassung der Entwicklung und Struktur der professionellen Unterrichtswahrnehmung. Erster Durchlauf: Lernatmosphäre, Lernbegleitung, Zielorientierung in Phasen der Erarbeitung und Übung; Erweiterung des Instruments: Lernzielfestlegung (<i>goal setting</i>), Zielorientierung (<i>goal orientation</i>), Durchführung von Lernaktivitäten (<i>execution of learning activities</i>), Monitoring und Regulation des Lernprozesses (<i>regulation</i>) sowie die dessen Evaluierung (<i>evaluation</i>)	Studierende Lehramt, Fachkräfte in der Lehrkraftbildungsforschung	Beschreiben, Erklären, Vorhersagen	Englisch, Französisch, Geschichte, Mathe, Physik	Video, realer Unterricht	geschlossen	Daten unterstützen die dreidimensionale Struktur Beschreiben, Erklären, Vorhersagen, erweiterte Version sensitiv für Expertiseunterschiede
24	Oser et al. (2010)	Erprobung der Erfassung professioneller Kompetenzen unter Berücksichtigung der Situation, Authentizität und Komplexität von Unterricht durch den advokatorischen Ansatz, Kompetenzprofil Gruppenunterricht	Lehrkräfte und Nicht-Lehrkräfte	Wahrnehmen, Beurteilen	Berufsfachschule	Video	geschlossen	Lehrkräfte waren näher am Masterrating als Nicht-Lehrkräfte, theoretisch modellierte Qualitätsdimensionen lassen sich faktorenanalytisch abbilden
25	Star & Strickland (2008)	Expertiseunterschiede in der Erfassung von Klassenzimmer, Klassenmanagement, mathematischem Unterrichtsgegenstand, Aufgaben und Kommunikation	Studierende Lehramt, Lehrkräfte	Wahrnehmen	Mathe	Video	offen, geschlossen (Multiple-Choice)	Expertinnen und Experten nehmen stärker als Novizinnen und Novizen dynamische Aspekte wahr, Verbesserung des <i>Noticing</i> durch Methodenkurs

26	Kersting (2008)	Expertise-Unterschiede und Zusammenhang mit fachspezifischem Wissen und Analysekompetenz bzgl. LU und Schülerinnen- und Schülerfehlern	Lehrkräfte	Beschreiben, Erklären, Bewerten	Mathe	Video, realer Unterricht	offen	Novizinnen und Novizen fokussieren eher Lehrkraft handeln und haben weniger Lernaktivitäten der Schülerinnen und Schüler im Blick, fachspezifisches Wissen korreliert positiv mit der PW
27	Seidel & Prenzel (2007), Schwindt (2008) Lernen aus Unterrichtsvideos (LUV)	Erfassung der Struktur der PW und Expertise-Unterschiede bzgl. fachdidaktischer Aspekte, Experimente, Fehlerkultur, Lernbegleitung und Zielorientierung	Studierende Lehramt, Lehrkräfte, Schulinspi-zierende	Beschreiben, Erklären, Bewerten	Physik	Video, realer Unterricht	geschlossen	Die Faktoren Beschreiben, Erklären und Bewerten lassen sich abbilden, Expertiseunterschiede werden deutlich

4.2.1 Beobachtungsgegenstände

In allen vorgestellten Studien konnte die PW erfolgreich durch den Einsatz von Vignetten erfasst werden, auch wenn nur bei einem Teil der Studien die eigentliche Entwicklung des Instruments zur Erfassung der PW im Vordergrund stand (siehe Tabelle 1, Studien 1–10). Gegenstand der PW waren z. B. fachspezifische Aspekte (siehe Tabelle 1, Studien 1, 6, 10, 11, 18 und 26), sprachliche Aspekte im Fachunterricht (siehe Tabelle 1, Studie 12), die Interaktion (siehe Tabelle 1, Studie 2 und 7), formatives Assessment (siehe Tabelle 1, Studie 3), KF (siehe Tabelle 1, Studie 9, 14 und 15) oder die LU (siehe Tabelle 1, Studie 8, 15, 19 und 21). Überwiegend bildet dafür naturwissenschaftlicher Unterricht den Kontext. Dabei wird in Studien mit unterschiedlichen Beobachtungsfokussen deutlich, dass die PW eine für den Gegenstand spezifische Fähigkeit ist. Erfasste Dimensionen wie KF und LU (siehe Tabelle 1, Studie 9 und 21) oder Diagnostik und Rückmeldung (siehe Tabelle 1, Studie 3) korrelieren stark miteinander, lassen sich empirisch aber dennoch trennen.

Barth (2017) bezieht in ihre Studie misslingende Unterrichtssequenzen ein. In einem achtminütigen inszenierten Unterrichtsvideo werden Störungen dargestellt, die primär von vor allem einem Kind ausgehen (z. B. Paul ruft rein, Paul provoziert die Lehrkraft) und auf die die Lehrkraft jeweils (unpassend) reagiert, sowie Störungen durch ungünstige Handlungen der Lehrkraft (z. B. unklares Tafelbild), aber auch affektiv-motivationale Informationen zur Lehrkraft (z. B. die Lehrkraft wirkt genervt). Es wird qualitativ erhoben, auf welche dieser Aspekte Studierende Bezug nehmen, und welche pädagogischen Aspekte erwähnt werden (Wahrnehmung). Zudem wird erfasst, ob und wie diese theoretisch beurteilt werden (Beurteilen). Als korrekt werden dabei Ausführungen gewertet, die in Bezug auf die KF relevant sind und sich auf den Klassen- oder Individualfokus beziehen (vgl. Ophardt & Thiel, 2013). Als Theoriebezug werden Begründungen auf drei Niveaustufen bewertet, die sich auf das Wissen über die Etablierung von Verhaltenserwartungen bzw. die Modifikation von Verhalten, den Unterrichtsfluss sowie das Klassenklima beziehen (Barth, 2017, S. 141ff.). Mit diesem qualitativen Zugang gelingt es Barth, die PW von Unterrichtsstörungen valide und reliabel zu erfassen und Unterschiede zwischen Bachelor- und Masterstudierenden aufzuzeigen.

4.2.2 Methodische Ansätze

Die vorgestellten Studien bedienen sich unterschiedlicher Methoden. Es überwiegen *mixed-methods*-Ansätze sowie qualitative Methoden wie offene Fragen, Interviews, lautes Denken oder der advokatorische Ansatz. Bei der Auswertung wird überwiegend auf die qualitative Inhaltsanalyse zurückgegriffen, wodurch zahlreiche Informationen berücksichtigt werden können (siehe Tabelle 1, Studie 1, 3, 4, 5, 6, 10, 11, 12, 13, 15, 16, 17, 18, 20, 22, 25 und 26). Eine

quantitative Erfassung, z. B. durch Likert-Skalen oder Multiple-Choice-Items, ist dagegen ökonomischer (siehe Tabelle 1, Studie 7, 8, 9, 14, 19, 21, 23, 24 und 27). Als Auswertungsmethoden werden verschiedene Ansätze genutzt. Star und Strickland (2008) legen beispielsweise theoriegeleitet fest, welche Antworten als richtig oder falsch zu bewerten sind. In anderen Studien (siehe Tabelle 1, Studie 9, 23 und 24) kommen Ratings von Expertinnen und Experten zum Einsatz und die Auswertung erfolgt anhand der Übereinstimmung mit diesen Ratings.

Im Hinblick auf die Förderung der PW mit Vignetten zeigt sich, dass der Einsatz von Vignetten effektiver ist als Lehrformate ohne Vignettenaufgaben, und zwar unabhängig vom verwendeten Vignettenformat. Zudem wird die Arbeit mit Vignetten gut akzeptiert (siehe Tabelle 1, Studie 9, 13, 14 und 15).

4.2.3 Erfasste Kompetenzfacetten und -unterschiede

Unterschiede zeigen sich bei den vorliegenden Ansätzen auch bei der Modellierung der erfassten Kompetenzfacetten, wobei häufig die Unterschiede in der Benennung liegen. Einige Studien erfassen neben der PW auch deklaratives Wissen (siehe Tabelle 1, Studie 4, 5 und 14) sowie Vorschläge zu Handlungsalternativen oder -entscheidungen (siehe Tabelle 1, Studie 1, 4, 11, 12, 17 und 18). Bezogen auf das dieser Arbeit zugrundeliegende Kompetenzmodell wird in der Mehrzahl der Arbeiten eine wesentliche Unterscheidung zwischen den Facetten Wahrnehmen und Interpretieren getroffen. Empirische Befunde zeigen sowohl eine mögliche Trennung von Facetten (siehe Tabelle 1, Studie 5, 6, 21, 23, 24 und 27) als auch Belege für eine eher einfaktorielle Struktur (siehe Tabelle 1, Studie 19) bzw. eine sehr starke Korrelation der beiden Subfacetten (siehe Tabelle 1, Studie 4). Expertiseunterschiede durch Praxiserfahrung und/oder spezifische Lerngelegenheiten hinsichtlich dieser Kompetenzfacetten lassen sich in allen Studien nachweisen. Sie äußern sich hinsichtlich der Wahrnehmung in der Anzahl der wahrgenommenen Ereignisse sowie in elaborierteren Interpretationen mit mehr Theoriebezug. Während Novizinnen und Novizen Situationen oft nicht verknüpft, sondern isoliert zueinander betrachten, das Handeln der Lehrkraft stark fokussieren und dabei weniger das Lernen der Schülerinnen und Schüler im Blick haben (Kersting, 2008; Sherin & van Es, 2009; Star & Strickland, 2008), können Expertinnen und Experten Wissensstrukturen vernetzen und beobachtete Situationen mit theoretischem Wissen verknüpfen. Durch Unterrichtserfahrungen können sie Schemata aufbauen und bestehende Schemata anpassen, was dazu führt, dass anwendbares Wissen in komplexen Situationen schnell zur Verfügung steht (König & Kramer, 2016). Dies hat insbesondere Einfluss auf die Fähigkeit, Situationen zu erklären und vorherzusagen. Da Novizinnen und Novizen Erfahrungen mit der Gestaltung von Lehr-Lernprozessen noch fehlen, konzentrieren sie sich weniger

auf die Beobachtung von Unterrichtsprozessen und mehr auf Einzelsituationen, die an der Oberfläche erkennbar sind und sich auf Lerninhalte beziehen (Kim & Klassen, 2018; Star & Strickland, 2008; Wolff et al., 2017). Sie benutzen eher subjektive Theorien zum Lehren und Lernen und beschreiben Situationen, statt sie wissensgeleitet zu interpretieren (van Es & Sherin, 2008; Gold et al., 2016b; Kim & Klassen, 2018; Schäfer & Seidel, 2015; Schwindt, 2008).

4.2.4 Praxisrelevanz der erfassten Kompetenzen

Die Relevanz der Expertiseunterschiede in vignettenbasierten Verfahren zeigt sich in Studien, die den Zusammenhang zwischen erfassten Fähigkeiten und tatsächlichen Kompetenzen im Unterricht, also die prognostische Qualität untersuchen (Kunter & Klusmann, 2010; Seidel & Prenzel, 2007). So nehmen Rutsch et al. (2018b, S. 18) an: „Ebenso ist zu erwarten, dass die Testwerte, die bei der Bearbeitung von Unterrichtsvignetten erzielt werden, valide Indikatoren für Unterrichtsqualität, Unterrichtshandeln und Schüler*innenvariablen darstellen.“ Ein weiterer, für die Praxis wichtiger Befund betrifft Studien, die unterschiedliche Impulse durch mehrere, verschiedene Vignetten nutzen. Sie legen nahe, dass Fähigkeiten *inhaltsspezifisch* erworben werden. Dies zeigt sich darin, dass Testergebnisse nicht nur von der erfassten Fähigkeit, sondern auch von der jeweiligen Vignette abhängen bzw. dass sich Kompetenzzuwächse nur kontextspezifisch zeigen (siehe Tabelle 1, Studie 13, 15 und 21).

Die professionelle Wahrnehmung lässt sich mithilfe von Vignetten reliabel und valide erfassen. Bisher fehlt ein Ansatz, der aus sonderpädagogischer diagnostischer Perspektive die PW von Lernbarrieren in der Unterrichtsgestaltung zum Gegenstand hat. Eine Erfassung der Kompetenz Erkennen scheint angesichts bisheriger Befunde möglich und sinnvoll, wobei sich Stichproben unterschiedlicher Expertise hinsichtlich des Studiengangs, des Fachsemesters, der Berufserfahrung und Lerngelegenheiten anbieten. Ein fachübergreifender Ansatz unter Verwendung mehrerer Vignetten scheint dafür sinnvoll zu sein, wobei zwischen unterschiedlichen Vignettenformaten sowie quantitativen und qualitativen Ansätzen abgewogen werden muss.

5 Forschungsdesiderata

Im vorherigen Kapitel wurde einerseits die Möglichkeit PW zu erfassen sowie die Vielfältigkeit der Ansätze gezeigt. Andererseits wird deutlich, dass Verfahren, die aus diagnostischer Perspektive explizit Störungen in Lehr-Lernprozessen fokussieren, fehlen. Lernbarrieren in Unterrichtsprozessen werden als spezifischer Beobachtungsgegenstand kaum thematisiert. Daraus ergeben sich die folgenden Forschungsdesiderata:

Systematisierung und Operationalisierung von Lernbarrieren

Als Lernbarrieren können grundsätzlich sämtliche Hindernisse in Lehr-Lernprozessen betrachtet werden. Sie können, wie in Kapitel 2 gezeigt, z. B. die Lernumgebung, Material und Aufgaben, Unterrichtsprozesse oder individuelle Bedingungen betreffen.

Im weiteren Vorgehen werden Barrieren in der Gestaltung von Unterricht in den Mittelpunkt gesetzt. Diese müssen in einem ersten Schritt definiert und operationalisiert werden.

Modellierung der Kompetenz Erkennen von Lernbarrieren

Nachdem Lernbarrieren konkretisiert wurden, ist die Kompetenz, diese zu erkennen, auf der Basis des vorgestellten Kompetenzmodells von Blömeke et al. (2015a) zu spezifizieren. Dabei findet der diagnostische Aspekt besondere Berücksichtigung.

Vignetten- und Itemkonstruktion

Wie bisherige Studien zeigen, lässt sich die PW mit dem Einsatz von Vignetten erfassen. Somit wird auch für diese Studie ein Vignettenformat gewählt und es werden Vignetten konstruiert, die die entworfenen Lernbarrieren enthalten.

Unabhängig vom Vignettenformat kommen qualitative und quantitative Verfahren in Frage, um das Erkennen der Barrieren in Lehr-Lernsituation zu erfassen. Vor- und Nachteile verschiedener Itemformen werden diskutiert. Auf dieser Grundlage werden offene Fragen und geschlossene Items entworfen.

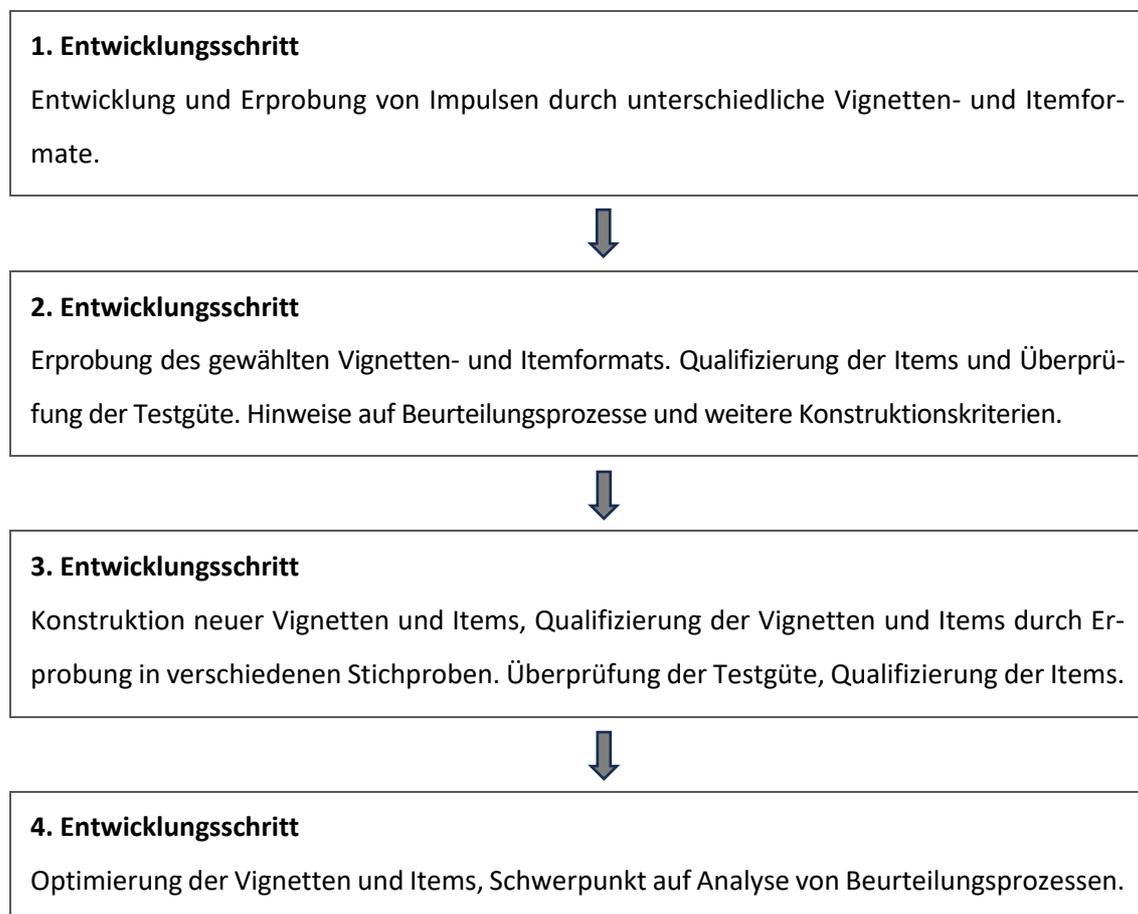
Entwicklung und Erprobung des Verfahrens zur Erfassung der Kompetenz Erkennen von Lernbarrieren

Das Verfahren wird in mehreren Schritten entwickelt. Im ersten Entwicklungsschritt sollen Informationen zu den eingesetzten Vignetten und Items gewonnen werden. Der Schwerpunkt liegt auf der Frage, wie Impulse durch Vignetten und Fragen bzw. Items beschaffen sein müssen, um Urteile zu Lehr-Lernsituationen so hervorzurufen, dass eine Beurteilung der Fähigkeit zum Erkennen von Lernbarrieren möglich wird. Die gewonnenen Erkenntnisse werden genutzt, um

Vignetten- und Itemformat festzulegen und im zweiten Entwicklungsschritt zu erproben. Neben der Prüfung der Testgütekriterien und der Itemqualifizierung soll der zweite Entwicklungsschritt auch Hinweise zu den Urteilsprozessen geben, die durch die entworfenen Impulse angeregt werden. Im dritten Schritt folgen daraus Konstruktionskriterien für weitere Vignetten und Items, die in einer größeren Stichprobe eingesetzt werden und in einem vierten Entwicklungsschritt nochmals optimiert werden. Im letzten Entwicklungsschritt werden vor allem Beurteilungsprozesse und Einflüsse auf diese in den Blick genommen (siehe Abbildung 2).

Abbildung 2

Übersicht über die Entwicklungsschritte von ErLe



I Entwicklung eines Verfahrens zur Erfassung der Kompetenz Erkennen von Lernbarrieren

6 Systematisierung und Operationalisierung von Lernbarrieren

Es gibt eine Vielzahl von Faktoren, die Einfluss auf Lehr-Lernprozesse haben und darüber bestimmen, ob diese barrierefrei verlaufen oder nicht. Die Suche nach Merkmalen zur Beschreibung von Unterricht hat eine lange Tradition und lässt sich aus unterschiedlichen Blickwinkeln betrachten. Im Fokus können z. B. effektives Lernen, didaktische Fragen, die Beziehungen zwischen Schülerinnen bzw. Schülern und Lehrkräften, die Förderung des Selbstkonzepts, soziales Lernen, Aufgabenkonstruktionen oder die Raumgestaltung stehen (Clausen, 2002; Einsiedler, 2001; Gruehn, 1995; Gruschka, 2007; Schrader et al., 1997). Grundsätzlich lassen sich Merkmale ausmachen, die eher kognitiv oder eher affektiv-motivational ausgerichtet sind. Allen Ansätzen gemeinsam ist das Bestreben, Bedingungen für guten Unterricht zu identifizieren und zu beschreiben.

Zum einen kann dafür der jeweilige Lernerfolg der Schülerinnen und Schüler betrachtet werden und als Maß dienen, ob der Unterricht – unabhängig von Kriterien oder Prinzipien, nach denen er vorher gestaltet wurde – gut war. Diese Vorgehensweise entspricht einem funktionalen Verständnis von Unterricht. Es werden besonders erfolgreiche Lehrkräfte ausgewählt, um zu beschreiben, welche Faktoren deren Unterricht bestimmen (*best practice*) (Helmke, 2011).

Zum anderen gibt es Bestrebungen, empirisch belegbare Qualitätsmerkmale von Unterricht zu bestimmen und theoretisch zu beschreiben. Dies entspricht dann einer normativen Betrachtungsweise (z. B. Berliner, 1987; Ditton, 2006; Helmke, 2009; Helmke, 2010; Helmke, 2011; Oser et al., 1992). Dieser Ansatz wird für die vorliegende Studie verwendet. Ziel für ErLe soll es sein, Lernbarrieren zu definieren, die auch unabhängig von der *fachlichen* Situierung beschrieben werden können. Es werden zunächst Merkmale erörtert, die dazu geeignet sind, Unterricht hinsichtlich seiner Qualität zu beobachten, zu beschreiben und zu beurteilen. In einem weiteren Schritt können von diesen Qualitätsmerkmalen Lernbarrieren abgeleitet werden, die definiert werden als nicht oder mangelhaft umgesetzte Merkmale guten Unterrichts. Die Kompetenz, Lernbarrieren zu erkennen, führt also über ein *spezifisches Beobachten* relevanter Unterrichtsmerkmale.

6.1 Unterrichtsqualität

Einsiedler (2002) beschreibt Unterrichtsqualität als

Bündel von Unterrichtsmerkmalen, die sich als „Bedingungsseite“ (oder Prozessqualität) auf Unterrichts- und Erziehungsziele („Kriterienseite“ oder Produktqualität) positiv auswirken, wobei die Kriterienseite überwiegend von normativen Festlegungen bestimmt ist und der Zusammenhang von Unterrichtsmerkmalen und Zielerreichung von empirischen Aussagen geleitet ist. (S. 195)

Qualitätsmerkmale beschreiben also Facetten der Unterrichtsgestaltung, die bestimmte Effekte auf Unterrichtsprozesse, Schülerinnen und Schüler und ihre Lernerfolge haben. Dabei soll die Wirkungsrichtung nicht als unidirektional verstanden werden, sondern als komplexes Wirkungsgefüge. Helmke (2006, 2007, 2010) beschreibt in seinem *Angebots-Nutzungsmodell* Interaktionsprozesse im Lehr-Lerngeschehen (Prozessebene) sowie Aspekte der kognitiven und motivational-affektiven Ziel- und Leistungs-Erreichung (Produktebene), wobei die Wirkungsrichtung der Komponenten als reziprok zu verstehen ist.

Um Merkmale der Unterrichtsqualität zu betrachten, muss also auch die Reaktion der Lernenden auf das Lernangebot in den Blick genommen werden (Helmke, 2010; Vieluf et al., 2020). Scriven (1991) beschreibt, dass Lernende individuelle Merkmale in den Lernprozess einbringen und Merkmale guten Unterrichts deshalb unterschiedlich effektiv sein können. Er unterscheidet zwei Dimensionen: Die Güte (*merit*) gibt an, ob Unterricht fachwissenschaftlich gestützt konzipiert ist und relativ allgemeingültigen Kriterien guten Unterrichts entspricht. Tauglichkeit (*worth*) dagegen beschreibt, wie förderlich Unterricht in der ganz konkreten Situation, Interaktion und Konstellation ist. Hier zeigt sich die oben beschriebene Herausforderung der PW von Unterricht: Beobachtbare Maßnahmen müssen nicht nur beschrieben werden, sondern es muss jeweils auch situationsspezifisch deren (mögliche) Wirkung beurteilt werden.

Auch Berliner (2005) unterscheidet guten Unterricht (*good teaching*), der normativen Prinzipien und aktuellen Standards im Sinne gesellschaftlich geteilter Vorstellungen und pädagogischer Überzeugungen folgt, und effektiven Unterricht (*effective teaching*), der die Erreichung unterschiedlicher Lernziele beschreibt. Als dritten Begriff führt er zusätzlich das qualitätvolle Unterrichten (*quality teaching*) ein, welcher die beiden vorherigen Aspekte guten Unterrichts vereint. Dabei bezieht sich die Effektivität nicht nur auf kognitiven, sondern auch auf emotional-motivationalen Lernerfolg sowie überfachliche Kompetenzen (Helmke, 2009). Die Verwendung des Begriffs Unterrichts*qualität* impliziert also, dass belegte Qualitätsmerkmale effektiv eingesetzt werden.

Dieser effektive Einsatz ist abhängig von einer Vielzahl von Faktoren wie Schulform, Klassengröße und -zusammensetzung, Altersstufe, Unterrichtsfach oder Arbeitsform (Gruschka, 2007). Zur Beurteilung von Unterrichtsqualität sind normative Qualitätsmerkmale also einerseits nötig, andererseits nicht allgemeingültig, sondern „nur unterschiedlich gut legitimierbar oder nützlich“ (Klieme, 2013, S. 437). Zudem sind Merkmale nicht so zu verstehen, dass eine maximale Ausprägung aller Merkmale gleichzeitig zu optimalem Unterricht führt, sondern immer die richtige Auswahl, Dosierung und Passung ausschlaggebend sind (Helmke, 2006; Helmke,

2009; Helmke & Schrader, 2008; Lipowsky, 2009; Weinert et al., 1989). In Bezug auf die Kompetenz des Erkennens von Lernbarrieren können Qualitätsmerkmale demnach als spezifischer Beobachtungsfokus auf die Unterrichtsgestaltung dienen, deren Effektivität einer weiteren Beurteilung bedarf.

6.2 Merkmale guten Unterrichts

Im Zusammenhang mit Konzepten der Unterrichtsqualität wird häufig auf Carroll (1973) und Bloom (1976) verwiesen. Carroll (1973) nennt im *Modell schulischen Lernens* fünf Variablen, die maßgeblich für den Lernerfolg sind: (1) die aufgabenspezifische Begabung der Lernenden, (2) die Fähigkeit dem Unterricht zu folgen, (3) die Ausdauer der Lernenden, (4) die zugestandene Lernzeit und als grundlegende Variable (5) die Unterrichtsqualität. Als Merkmale von Unterrichtsqualität nennt Carroll die klare Darstellung der Aufgabenanforderung, die adäquate Darbietung der Aufgabe, eine angemessene Lernschrittfolge sowie die Berücksichtigung der Schülerbedürfnisse und -charakteristika. Diese Faktoren haben wiederum einen wesentlichen Einfluss auf die effektiv zur Verfügung stehende Lernzeit (Gruehn, 2000, S. 6f.).

Das Modell von Bloom (1976) orientiert sich stark am Modell von Carroll (1973), betont jedoch weniger die Lernzeit und mehr den aufeinander aufbauenden Charakter von Lernprozessen. Lernergebnisse vergangener Aufgaben, gerade auch affektiver Natur, bestimmen den Erfolg beim Bearbeiten neuer Aufgaben maßgeblich mit, z. B. durch das Vertrauen in die eigenen Fähigkeiten. Bloom erweitert damit das auf kognitiven Lernerfolg ausgerichtete Modell um affektiv-motivationale Aspekte des Lernens.

Wesentliche Faktoren der Unterrichtsqualität sind nach Bloom (1976) Hinweise auf eine inhaltlich sequenzierte Darbietung des Lernstoffs, Bekräftigung, die zu Hinwendung zur Aufgabe, positiver Einstimmung zur Schule und zum Fach sowie zu einer positiven Fähigkeitseinschätzung führt, Beteiligung der Lernenden am Unterricht durch Erhöhung der effektiven Lernzeit, insbesondere durch Beobachten und Managen des Klassengeschehens sowie Rückmeldungen zu Leistungen und Anregungen zu Lernverbesserungen, auch durch alternative Inhalte und/oder mehr Zeit (Gruehn, 2000, S. 8).

Als gemeinsame Qualitätsmerkmale von Carroll und Bloom, auf denen viele weitere Modelle aufbauen (vgl. Gruehn, 2000), lassen sich also die folgenden definieren: die Verständlichkeit und inhaltliche Klarheit der Instruktion, die inhaltliche Sequenzierung von Aufgaben, die Adaptivität des Unterrichts und des Lernstoffs, die Unterstützung der Lernenden durch Feedback sowie eine Optimierung der effektiven Lernzeit durch Management des Klassengeschehens.

Im Einklang mit diesen Erkenntnissen leiten Brophy und Good (1986) aus der Forschung zur prozessbezogenen Unterrichtsforschung Variablen für die Unterrichtsqualität ab: (1) die Informationsdarbietung mit den Punkten Strukturierung, Sequenzierung, Klarheit, Enthusiasmus sowie Tempo; (2) die Gestaltung von Fragen der Lehrkraft mit den Aspekten Schwierigkeitsgrad, kognitive Anforderung, Klarheit, Wartezeit bis zum Aufrufen, Auswahl der antwortenden Schülerinnen und Schüler und Wartezeit nach Aufrufen der Schülerrinnen und Schüler sowie (3) die Reaktion auf unterschiedliche Antworten, Fragen und Kommentare der Lernenden.

Die unterschiedlichen Qualitätsmerkmale machen deutlich, dass auf Unterrichtsseite vor allem fünf miteinander zusammenhängende Faktoren eine Rolle spielen, nämlich der Lernstoff, die Gestaltung von Aufgaben, Instruktionen, effektive Lernzeit und das Verhalten der Lehrkraft (bzgl. Fragen, Reaktionen, Feedback, Sprache und Management des Unterrichtsgeschehens). Die Anforderungen bzgl. der Gestaltung dieser Faktoren bestehen in Strukturierung, Sequenzierung, Klarheit, Tempo, Adaptivität sowie kognitiver Anforderung. Diese Merkmale tauchen in unterschiedlichen Formulierungen und Zusammenstellungen immer wieder auf. Exemplarisch sollen die Arbeiten von drei viel rezipierten Autoren vorgestellt werden: Hilbert Meyer (z. B. 2004), Andreas Helmke (z. B. 2004) und John Hattie (z. B. 2008).

Meyer (2004) fasst die Ergebnisse empirischer Unterrichtsforschung in 10 Merkmalen guten Unterrichts zusammen: (1) klare Strukturierung, (2) hoher Anteil echter Lernzeit, (3) lernförderliches Klima, (4) inhaltliche Klarheit, (5) sinnstiftendes Kommunizieren, (6) Methodenvielfalt, (7) individuelles Fördern, (8) intelligentes Üben, (9) transparente Leistungserwartungen und (10) vorbereitete Umgebung. Er nennt damit Merkmale und Maßnahmen, die sich den bereits genannten Faktoren zuordnen lassen, und fasst zum Teil verschiedene inhaltliche Punkte zusammen. Zum Beispiel ordnet er der klaren Strukturierung des Unterrichts die Unterpunkte Prozess-, Ziel- und Inhaltsklarheit, Rollenklarheit sowie Absprache von Regeln, Ritualen und Freiräumen zu. Er mischt dabei in einem Merkmal Aspekte, die zum Teil die Beziehungen betreffen können (z. B. Rollenklarheit), zum Teil die Unterrichtsorganisation (z. B. Prozessklarheit oder Regeln und Rituale), aber auch kognitive Prozesse adressieren können (z. B. Inhaltsklarheit). Inhaltliche Klarheit wird zudem auch als separater Punkt aufgeführt (Gruschka, 2007).

Auch in den 10 fachübergreifenden Merkmalen, die Helmke (2004) als Kern guten Unterrichts zusammenstellt, tauchen diese Aspekte auf. Zudem nennt er im Zusammenhang mit der Maximierung echter Lernzeit die KF als wichtige Variable für Unterrichtsqualität und meint damit „Konzepte, Strategien und Techniken, die dem Ziel dienen, einen störungsfreien und reibungslosen Unterrichtsverlauf zu ermöglichen und damit aktive Lernzeit zu maximieren“ (Helmke & Helmke, 2014, S. 9).

Die 10 Merkmale von Helmke (2004) umfassen (1) effiziente KF und Zeitnutzung (Störungsprävention und -behebung, Regeln), (2) lernförderliches Unterrichtsklima (u. a. konstruktiver Umgang mit Fehlern), (3) vielfältige Motivierung, (4) Strukturiertheit und Klarheit (u. a. klare Sprache, strukturierende Hinweise), (5) Wirkungs- und Kompetenzorientierung, (6) Schülerinnen- und Schülerorientierung, Unterstützung, (7) Förderung aktiven, selbständigen Lernens, (8) angemessene Variation von Methoden und Sozialformen, (9) Konsolidierung, Sicherung, und intelligentes Üben sowie (10) Passung.

Wie bei Meyer (2004) werden in der Merkmalsliste Aspekte, die eher die Unterrichtsgestaltung und den Ablauf betreffen, wie die KF und in Teilen die Strukturiertheit und Klarheit, hier in einem Punkt zusammengefasst. Dazu kommen die Lernatmosphäre sowie mehrere lernunterstützende Maßnahmen, die vor allem kognitive Prozesse betreffen. Helmke und Schrader (2008, S. 22) kommen wie Meyer (2004) zu der Schlussfolgerung, dass die Punkte *Klarheit* und *Strukturiertheit* wesentliche Merkmale darstellen, und fügen diesen die KF hinzu. Letztere ist im Sinne von Steuerungswissen, das der Reflexion und Analyse des Unterrichts dient, zu verstehen. Zudem werden in unterschiedlicher Formulierung die Lernatmosphäre übereinstimmend genannt sowie diverse Maßnahmen, die man zusammengefasst als kognitiv lernunterstützend beschreiben könnte.

Die vielleicht meistzitierte Studie zu Einflussfaktoren auf den Lernerfolg ist die Metastudie von John Hattie (2008), in der er Befunde aus über 50 000 weltweit veröffentlichten empirischen Studien zusammenträgt. Unter den untersuchten Merkmalen finden sich in Bezug auf Unterrichtsbedingungen auch hier die *Klarheit der Lehrperson* und *Feedback* als Merkmale mit sehr starken Effekten sowie *effektive KF* mit deutlicher Effektstärke. Die Studie von Hattie berücksichtigt allerdings nur wenige Metaanalysen zur KF. Ferner finden sich Aspekte der LU und Lehrkraft-Lernenden-Beziehung unter den effektstärksten Merkmalen. Sie können als zentrale Bedingungen für qualitätvollen Unterricht angesehen werden (Hattie, 2008).

Die folgende Tabelle zeigt einen Versuch, die gesammelten Merkmale zu ordnen und aufeinander zu beziehen sowie sie zu Dimensionen zu bündeln (siehe Tabelle 2):

Tabelle 2

Übersicht über Unterrichtsmerkmale und Bündelung zu Dimensionen

Dimension	Carroll (1973) & Bloom (1976)	Brophy & Good (1986)	Meyer (2004)	Helmke (2004)	Hattie (2008) (in Anlehnung an Beywl & Schmidt, 2014, S. 12; Lipowsky & Bleck, 2019)
Klassenführung	Hoher Anteil echter Lernzeit durch Management des Klassengeschehens (bei Carroll (1973) als Ergebnis von Unterrichtsqualität)		Klare Strukturierung (Unterpunkt Regeln und Rituale), hoher Anteil echter Lernzeit	Effiziente KF und Zeitznutzung (Regeln, Störungsprävention und Störungsbehebung)	KF, Beeinflussen von Verhalten in der Klasse, aktive Lernzeit, Reduzieren von Unterrichtsstörungen
Inhaltliche Klarheit	Verständlichkeit der Instruktion durch die Lehrkraft	Informationsdarbietung: Klarheit, Gestaltung von Fragen: Klarheit	Inhaltliche Klarheit	Strukturiertheit und Klarheit (u. a. klare Sprache, strukturierende Hinweise)	Klarheit der Lehrperson (in Bezug auf Erläuterungen, klare Sprache)
Zielklarheit		(keine Unterscheidung zur inhaltlichen Klarheit)	Transparente Leistungserwartungen, klare Strukturierung	Strukturiertheit und Klarheit (auch bzgl. der Ziele)	Klarheit der Lehrperson (in Bezug auf Ziele und Kriterien)
Strukturierung (inhaltlich)	inhaltliche Sequenzierung von Aufgaben	Informationsdarbietung: Strukturierung und Sequenzierung	Klare Strukturierung	(wie inhaltliche Klarheit)	Klarheit der Lehrperson (in Bezug auf Inhalte)
Strukturierung (zeitlich)		Informationsdarbietung: Tempo,			

		Gestaltung von Fragen: Wartezeit bis zum Aufrufen, Wartezeit nach Aufruf eines Schülers oder einer Schülerin			
An die Lernenden angepasste Lernunterstützung	Adaptivität des Unterrichts	Gestaltung von Fragen: Schwierigkeitsgrad, kognitive Anforderung, Auswahl des antwortenden Schülers oder der Schülerin	Methodenvielfalt, individuelles Fördern, intelligentes Üben, sinnstiftendes Kommunizieren	Wirkungs- und Kompetenzorientierung, angemessene Variation von Methoden und Sozialformen Konsolidierung, Sicherung, intelligentes Üben Schülerinnen- und Schülerorientierung, Unterstützung, Förderung aktiven, selbständigen Lernens, Passung	Lerntechniken, rhythmisiertes vs. geballtes Üben, direkte Instruktion, metakognitive Strategien, Problemlösen
Lernförderliches Klassenklima	Reaktion auf unterschiedliche Antworten, Fragen und Kommentare der Schülerinnen und Schüler	Informationsdarbietung: Enthusiasmus, Gestaltung von Fragen: Wartezeit bis zum Aufrufen, Wartezeit nach Aufruf	Lernförderliches Klima	Lernförderliches Unterrichtsklima (u. a. konstruktiver Umgang mit Fehlern), Schülerinnen- und Schülerorientierung (Unterpunkt affektiver Aspekt der Beziehung zwischen Lehrkraft und Schülerinnen und Schülern), vielfältige Motivierung	Angstarmut, Beziehung zwischen Lehrkraft und Schülerinnen und Schülern, Klassenzusammenhalt, Nicht-Etikettieren der Lernenden

Lernumgebung

Feedback

Feedback

Vorbereitete Umgebung

Sinnstiftendes Kommunizieren (Unterpunkt Schülerinnen- und Schüler-Feedback)

Schülerinnen und Schüler-Orientierung (Unterpunkt Schülerinnen- und Schüler-Feedback)

Feedback

Hier wird deutlich, dass viele zusammenwirkende Merkmale letztlich die Qualität von Unterricht ausmachen. KF, Klarheit, Strukturierung, LU und die Lernatmosphäre sind wiederkehrende zentrale Elemente¹. Die Lernumgebung und das Feedback der Lehrkraft tauchen zudem punktuell auf.

Als Folge großer Schulleistungsstudien wie PISA oder TIMSS gab es in der empirischen Unterrichtsqualitätsforschung zunehmend Bestrebungen, solche Merkmalsammlungen zu fachübergreifenden Dimensionen zusammenzufassen. Die Unterrichtsbeobachtungen zur empirischen Wirksamkeit von Unterricht bei TIMSS (Hiebert et al., 2003) ergeben drei wesentliche Faktoren für Unterrichtsqualität: (1) Unterrichts- und Klassenführung wird hier verstanden als die Steuerung des Unterrichts, der Lernorganisation und der Störungsprävention zur optimalen Nutzung von Lernzeit (König, 2015; Kunter & Ewald, 2016). Die zweite Basisdimension ist (2) Schülerinnen- und Schülerorientierung und konstruktive Unterstützung der Lernenden durch Differenzierung, adaptive Unterstützung und Rückmeldung (Klieme & Rakoczy, 2008). (3) Kognitive Aktivierung als dritte Basisdimension meint das Potenzial, Lernende zu vertieftem Nachdenken über die Unterrichtsinhalte anzuregen. Der positive Zusammenhang dieser Dimensionen mit dem Lernerfolg konnte unter anderem auch in der COACTIV-Studie belegt werden. Für die Dimension der KF zeigt sich, dass diese sich sowohl auf die kognitiven Lernerfolge als auch auf die emotional-motivationale Ebene auswirkt (Baumert et al., 2010; Kunter et al., 2013). Über diese Basisdimensionen von Unterrichtsqualität ist sich die aktuelle Forschung grundsätzlich einig, auch wenn es Unterschiede in der Benennung und Operationalisierung gibt (Fauth et al. 2014; Klieme et al., 2006; Kunter & Ewald, 2016; Kunter et al., 2013; Kunter & Voss, 2011; Praetorius et al., 2014; Praetorius et al., 2018).

Kleickmann et al. (2020) überprüften diese drei Basisdimensionen anhand von Ratings von Schülerinnen und Schüler zur Unterrichtsqualität. Sie kommen zu dem Schluss, dass eine vierte und bislang nicht gesondert aufgeführte Dimension (*kognitive Unterstützung*) den Lernerfolg von Schülerinnen und Schülern vorhersagt und als Ergänzung zu einem differenzierteren und treffenderen Modell der Unterrichtsqualität beiträgt. *Kognitive Aktivierung* zielt demnach darauf ab, die Schülerinnen und Schüler in Denkprozesse höherer Ordnung und in die Wissenskonstruktion und -revision einzubeziehen (z. B. durch herausfordernde Aufgaben und die Erforschung von Vorwissen). *Konstruktive Unterstützung* bedeutet in erster Linie, die Motivation der Schülerinnen und Schüler zu fördern (z. B. durch die Unterstützung von Erfahrungen mit Auto-

¹ Je nach Konzeptualisierung wird Klarheit auch als ein Merkmal der Lernunterstützung verstanden und Struktur entweder als ein Merkmal der Klassenführung (prozessuale Strukturierung) oder der Lernunterstützung (inhaltliche Strukturierung).

nomie und sozialer Verbundenheit). Kognitive Unterstützung dient dazu, die kognitiven Anforderungen in herausfordernden Lernumgebungen zu reduzieren, damit die Lernenden sie beherrschen können. Mit Blick auf die Merkmalstabelle finden sich hier auch Items, die sich der Strukturierung, Zielklarheit und sprachlichen Klarheit zuordnen lassen, z. B. *I know exactly what to do when I am working* oder *Our teacher often uses foreign words we do not understand* (Kleickmann et al., 2020, S. 54).

Insgesamt finden sich die Merkmale der obigen Tabelle in diesen vier Basisdimensionen gut wieder: Das Merkmal KF wird explizit aufgeführt, zur kognitiven Unterstützung lassen sich die Merkmale der Klarheit und Strukturierung zuordnen, die Merkmale kognitiver Aktivierung finden sich im Bereich der LU und die konstruktive Unterstützung beinhaltet die Aspekte zur Förderung der Lernatmosphäre.

Das Feedback wird als Merkmal unterschiedlich verstanden und verortet. Es kann sowohl Feedback der Lehrkraft betreffen (z. B. Bloom, 1976) als auch das Einfordern von Feedback durch Schülerinnen und Schüler an die Lehrkraft (z. B. Helmke, 2004). Zudem wird es teilweise eher als kognitiv unterstützende Maßnahme verstanden (z. B. Kleickmann et al., 2018; Lipowsky & Bleck, 2019), hat aber auch Einfluss auf die konstruktive Unterstützung und die Lernatmosphäre (z. B. Brophy & Good, 1986).

Die KF nimmt unter den Basisdimensionen eine exponierte Stellung ein, da ihr besonderer Wert als ein grundlegender Faktor in einem komplexen Wirkungsgeflecht für den Lernerfolg sowie die Lernatmosphäre mehrfach belegt worden ist (Helmke et al., 2010; Helmke & Helmke, 2014; Klieme & Rakoczy, 2008; Marzano, 2000; Oliver et al., 2011; Seidel & Shavelson, 2007; Wang et al., 1993). Wie Helmke und Helmke (2014) formulieren: „Effiziente Klassenführung ist nicht alles, aber ohne sie geht alles andere gar nicht“ (S. 9).

6.3 Konkretisierung und Operationalisierung von Unterrichtsmerkmalen und Lernbarrieren

Die Beobachtung von Lernbarrieren ist im Hinblick auf verschiedene Unterrichtsfächer und Fachdidaktiken möglich, soll aber zunächst aus überfachlicher Sicht erfolgen. Für die Operationalisierung von Unterrichtsmerkmalen und eine Ableitung von Lernbarrieren aus diesen Merkmalen sind also zunächst Aspekte zu wählen, die sich überfachlich beschreiben lassen, so dass ein Erkennen von Lernbarrieren auch ohne fachspezifisches Wissen möglich ist.

Unter Einbezug der Ausführungen im vorangegangenen Kapitel zeigen sich KF und lernunterstützende Merkmale als zentrale Elemente. Der Begriff des *Classroom Management* wird

schon seit vielen Jahrzehnten diskutiert und unterschiedlich konzeptualisiert (Brophy, 2006; Cipani, 2008; Taubert, 2007). Im deutschsprachigen Raum werden dafür die Begriffe Klassenmanagement (Merkens, 2010; Ophardt & Thiel, 2013; Rogers & Kühn, 2013) oder KF (Baumert & Kunter, 2006; Gold & Holodyski, 2011; Haag & Streber, 2012; Seidel, 2015) genutzt. Damit verbunden sind unterschiedliche Sichtweisen: Die behavioristische Sichtweise von KF fokussiert vor allem den Umgang mit Unterrichtsstörungen und Verhaltensbeobachtungen und legt Verstärkemechanismen zugrunde. Die Erziehungsstilforschung nimmt diese Bedingungen und Auswirkungen auf die Schülerinnen und Schüler näher in den Blick (Haag, 2018). Die kontextualistisch-ökologische Sichtweise bezieht sich auf die Schaffung und Aufrechterhaltung einer Lernumgebung mit einer bestimmten Aktivitätsstruktur, die auf Lernziele ausgerichtet ist (Doyle, 2006; Kounin, 2006; Piwowar, 2013). Planung und Regeln sollen dabei Abläufe präventiv optimieren und in der konkreten Unterrichtssituation etabliert, geschützt und gestärkt werden, so dass dieses Verständnis über den reaktiven Umgang mit Störungen hinausgeht (Ophardt & Thiel, 2017). Zudem wurde der Begriff des *Classroom Managements* erweitert und umfasst nicht mehr nur Aspekte der Disziplin, sondern allgemeiner das Schaffen einer lernförderlichen Umgebung (Bönte et al., 2019; Haag, 2018).

KF ist also essentiell für die aktiv genutzte Lernzeit (Lipowsky, 2009; Lipowsky & Bleck, 2019) und Voraussetzung für einen kognitiv aktivierenden und strukturierten Unterricht sowie für die Lernmotivation (Fauth et al., 2014; Hattie, 2008; Holodyski et al., 2017; Kunter et al., 2013). Es besteht Einigkeit darüber, dass diese als generische Dimension von Unterrichtsqualität gelten kann (vgl. Klieme et al., 2006; Kunter & Voss, 2011; Kunter et al., 2013; Praetorius et al., 2014).

Ergänzt wird diese durch kognitiv aktivierende und lernunterstützende Maßnahmen, die meist fachspezifisch verstanden werden oder fachspezifische Anteile enthalten (Brunner, 2018; Heinitz & Nehring, 2020; Jentsch et al., 2020; Praetorius & Charalambous, 2018). Kognitive Aktivierung, kognitive Unterstützung und konstruktive Unterstützung sind insofern generisch, als dass sie für jeden Unterricht gelten. Fachspezifische Anteile werden jedoch dann deutlich, wenn es um Wissen und Kompetenzen zur Umsetzung dieser Qualitätsmerkmale sowie zur Beurteilung konkreten Unterrichts geht. Lipowsky et al. (2018), die sich mit der Korrelation fachspezifischer und generischer Qualitätsmerkmale befassen, kommen zu dem Schluss, dass fachspezifische Qualitätsmerkmale und generische Qualitätsmerkmale verschiedene Dimensionen von Unterrichtsqualität darstellen. Sie beziehen sich dabei auf die Analyse konkreten Fachunterrichts, dessen Erfolg durch generische und fachspezifische Qualitätsmerkmale gemeinsam am besten erklärt werden kann (Brunner, 2018; Heinitz & Nehring, 2020; Jentsch et al., 2020; Praetorius & Charalambous, 2018). Einige Merkmale können dabei als „Konglomerat von fachspezifischen und generischen Anteilen“ (Heinitz & Nehring, 2020, S. 324) beschrieben werden, „so

z. B. das Kriterium der ‚Klarheit‘ (Brunner, 2018) oder der ‚kognitiven Aktivierung‘ (Praetorius & Charalambous, 2018), da zu einer Beurteilung dieser Kriterien sowohl fachliches oder fachdidaktisches Wissen notwendig ist als auch ein grundsätzliches pädagogisches Wissen“ (Heinitz & Nehring, 2020, S. 324).

Im Projekt *Videobasierte Unterrichtsanalyse: Early Science* – Theoretische Modellierung und empirische Erfassung der Kompetenzen zur Analyse der Lernwirksamkeit von naturwissenschaftlichem Grundschulunterricht (ViU; Holodynski et al., 2017) werden KF und LU operationalisiert und genutzt, um die PW anhand videografierter Unterrichtssituationen zu erfassen. Dazu wurden Szenen aus dem naturwissenschaftlichen Unterricht gewählt. Das Instrument enthält kurze Unterrichtsszenen aus dem Sachunterricht zu den Themen *Schwimmen und Sinken* sowie *Aggregatzustände*, in denen Merkmale der KF und LU erkennbar sind. Zu diesen Merkmalen sind geschlossene Items anhand einer Skala zu beurteilen (Gold et al., 2013). Die LU bezieht sich dabei auf eine fachdidaktische Perspektive und vereint Merkmale, die naturwissenschaftliche Lernprozesse unterstützen. Die KF bezieht sich auf die pädagogisch-psychologische Perspektive und umfasst generische Merkmale.

Da die Bedeutung von KF und LU für den Erfolg von Lehr-Lernprozessen fach- und jahrgangübergreifend empirisch gut belegt ist (Janeczko et al., 2022) und die PW von Maßnahmen der KF und LU im Projekt ViU erfolgreich erfasst werden konnte, soll die Operationalisierung für ErLe in Anlehnung an das Projekt ViU vorgenommen werden. Merkmale der LU sollen dabei so weit wie möglich aus dem fachlichen Kontext gelöst werden und generische Aspekte repräsentieren. Im Folgenden werden die Unterrichtsmerkmale KF und LU vorgestellt.

6.3.1 Klassenführung

Grundsätzlich lassen sich drei Dimensionen der KF unterscheiden, nämlich (1) das Monitoring, das vor allem beinhaltet, dass die Lehrkraft das Klassengeschehen im Blick hat, (2) die Maximierung von Lernzeit durch einen reibungslosen Verlauf und flüssige Übergänge sowie eine klare Unterrichtsstruktur und schließlich (3) der effektive Umgang mit Unterrichtsstörungen, auch präventiv durch das Etablieren von Regeln und Routinen sowie eine klare Kommunikation von Zielen und Erwartungen und Feedback (Gold & Holodynski, 2011; Ophardt & Thiel, 2013; Seidel, 2015). Im Folgenden werden diese drei Facetten genauer dargestellt und Merkmale für die Nutzung in der vorliegenden Studie festgelegt.

Monitoring

Lehrkräfte können einen reibungslosen Ablauf und die Maximierung echter Lernzeit ermöglichen, indem sie jederzeit über das Geschehen in der Klasse informiert sind und die Klasse dies auch

wissen lassen. Das betrifft unter anderem die *Reaktion auf Störungen* und andere Unterrichtsunterbrechungen, die so gestaltet ist, dass sie möglichst wenig negativen Einfluss auf die Lernzeit hat. Sie muss also kurz und bündig (ohne den Unterrichtsfluss zu stören) und zum richtigen Zeitpunkt (zeitnah, bevor Störungen sich ausbreiten) erfolgen und an das richtige Kind gerichtet sein (Holodynski et al., 2017; Kounin, 2006; Thiel et al., 2012). Dabei laufen mehrere Unterrichtsprozesse gleichzeitig ab, die die Lehrkraft parallel im Auge behalten und steuern muss (*Allgegenwärtigkeit/Überlappung*). Während einer Intervention darf sie beispielsweise andere Prozesse nicht außer Acht lassen oder unterbrechen (Gold & Holodynski, 2011; Holodynski et al., 2017).

Zu den Aufgaben der Lehrkraft gehört es auch, ungünstige Einflüsse auf das Lernen einzelner Schülerinnen oder Schüler zu erkennen, z. B. in Form eines ungünstigen Sitzplatzes. Neben der Gestaltung von Unterrichtsprozessen müssen also auch *räumliche Bedingungen* der Lernumgebung beachtet werden und sollen deshalb als Merkmal Verwendung finden.

In ViU wird auch das *Feedback* zum Monitoring gezählt und meint vor allem eine positive Präsenz, zu der beispielsweise auch bestärkende Mimik und Gestik gehören. Wie oben bereits erwähnt umfasst Feedback in anderen Konzeptualisierungen aber auch andere Aspekte wie z. B. einen wertschätzenden Umgang, Empathie oder ein ehrliches Interesse an den Lernenden (vgl. Holodynski et al., 2017; Kunter & Trautwein, 2013; Kunter & Voss, 2011; Lipowsky, 2009), geduldiges und nicht abwertendes Verhalten und die Reaktion auf Schwierigkeiten und Fehler (Gabriel, 2014; Kunter & Voss, 2011) oder auch lernunterstützende Hinweise. Feedback hat somit auch Einfluss auf die Lernatmosphäre und bezieht sich gleichermaßen auf die KF, die kognitive Unterstützung und die konstruktive Unterstützung. Die Dimension der konstruktiven Unterstützung wird unterschiedlich verstanden und operationalisiert (Gruehn, 2000), z. B. werden sowohl Aspekte der subjektiv wahrgenommenen Lernatmosphäre dazu gezählt als auch die Qualität der Beziehungen zwischen Lernenden sowie zwischen Lernenden und Lehrkräften (Gruehn, 2000) oder auch Aspekte der Schul- und Unterrichtsgestaltung (Lipowsky & Bleck, 2019). Wegen der großen Bedeutung des Feedbacks für den Lernerfolg (Hattie & Timperley, 2007) soll dieses als eigenständiges Merkmal aufgenommen werden.

Prozessuale Strukturierung

Die richtige Auswahl von Unterrichtsaktivitäten ist immer auch eine fachdidaktische Entscheidung und kann nicht losgelöst von den jeweiligen Inhalten und Zielen erfolgen. In Bezug auf KF ist für eine erfolgreiche Umsetzung die Ablaufsteuerung von Unterrichtsaktivitäten (z. B. Gespräch, Instruktion, Stillarbeit oder Gruppenarbeit) maßgeblich, so dass reibungslose Übergänge ermöglicht werden (Holodynski et al., 2017).

Dafür muss für alle Beteiligten definiert sein, was zu tun ist, und alle momentan wichtigen Komponenten müssen im Blick bleiben. Innerhalb einzelner Aktivitäten sollen Lehrkräfte für einen schwungvollen Verlauf sorgen, indem das Lerntempo für jedes einzelne Kind angemessen bleibt. Zum Beispiel können Meldungen, die nicht zur Unterrichtsaktivität beitragen, abgekürzt oder eingeordnet werden, zu lang dauernde Aktivitäten unterstützt werden oder zusätzliche Zeit gewährt werden. Eine ständige Justierung des Unterrichtstempos soll sowohl Langeweile als auch Überforderung vermeiden und unterschiedliche Lerntempi auffangen (Holodynski et al., 2017; Kounin, 2006). Maßnahmen zur Gruppenmobilisierung sorgen dafür, dass alle Lernenden durchgehend gleichermaßen beschäftigt sind (Holodynski et al., 2017; Kounin, 2006).

Regeln und Routinen

Ein störungsfreier Unterricht, der reibungslos und schwungvoll verläuft, wird auch durch das Einführen und Etablieren von unterrichtsförderlichen Regeln und Routinen ermöglicht. Gerade die Durchführung von sich wiederholenden Aktivitäten kann erleichtert werden, wenn deren Ablauf eingeübt wird und allen Beteiligten klar ist, was sie wann und wie zu tun haben. Dies beugt wiederkehrenden Anweisungen seitens der Lehrkraft vor. Hier sind geübte Übergänge, Rituale und Aktivitäten denkbar, wie z. B. Begrüßung, Bilden eines Sitzkreises, Bilden von und Verhalten in Arbeitsgruppen, Organisation von und Umgang mit Arbeitsmaterialien). So werden Störungen und Zeitverlust vermieden und die Lehrkraft wird entlastet, so dass sie mehr Kapazitäten für andere unterstützende Maßnahmen hat (Holodynski et al., 2017). Neben dem Einführen und Üben der Regeln und Routinen gehört die Erinnerung an die Einhaltung der etablierten Abläufe zu den wichtigsten Maßnahmen (Emmer & Evertson, 2013; Kounin, 2006; Ophardt & Thiel, 2017).

6.3.2 Lernunterstützung

Nach lernpsychologischen Erkenntnissen ist Lernen nicht nur die bloße Aufnahme von Informationen, sondern es handelt sich um einen Prozess der Wissenskonstruktion, bei dem Kinder und Jugendliche Unterstützung brauchen (Lipowsky, 2006; Lipowsky & Bleck, 2019). Lehrkräfte müssen über die Vorstellung der Lernenden informiert sein und an diese anknüpfen, um zum Nach- und Weiterdenken anzuregen und so die Veränderung von Vorstellungen einzuleiten. Kognitive Aktivierung muss im richtigen Maß herausfordernd sein, um Denkprozesse anzuregen, darf gleichzeitig aber nicht überfordern, sondern muss Verstehensprozessen Raum geben. Insofern sind lernunterstützende Maßnahmen anders als Klassenführungsmaßnahmen in hohem Maße von den kognitiven Voraussetzungen der Lernenden abhängig.

In Bezug auf die Beobachtung ergibt sich dabei die Schwierigkeit, dass Denk- und Lernprozesse nicht unmittelbar beobachtbar sind. Kognitive Aktivierung selbst ist nicht sichtbar und es können allenfalls Maßnahmen beobachtbar sein, die ein Potenzial zur kognitiven Aktivierung haben (Lipowsky, 2009; Lipowsky & Bleck, 2019; Renkl, 2011; Wolters, 2014). Zu diesen Maßnahmen gehört die Exploration vorhandener Vorstellungen, indem die Lehrkraft diese erfragt und begründen lässt. Durch Nachfragen, Gegenüberstellen unterschiedlicher Thesen oder die Anregung zur Überprüfung der vorhandenen Vorstellungen löst die Lehrkraft kognitive Konflikte aus und schafft so die Grundlage, Unzulänglichkeiten in den eigenen Vorstellungen zu entdecken und diese dann weiterzuentwickeln. Wesentlich ist dafür das Anforderungsniveau der Aufgaben für jedes einzelne Kind. Vorwissen und Können müssen der Lehrkraft bekannt sein, damit Anforderungen in der *Zone der nächsten Entwicklung* nach Wygotski (1987) liegen und Unter- oder Überforderung vermieden werden. Dazu kann die Lehrkraft dabei unterstützen, vorhandenes Wissen mit neuen Erkenntnissen zu verknüpfen. Die Anwendung der neuen Konzepte auf neue Fragen und Aufgaben festigt das Wissen und ermöglicht den Wissenstransfer. Auch die Anregung zum Austausch über verschiedene Konzepte und Vorstellungen dient der Überprüfung, Festigung und Flexibilisierung neu erworbenen Wissens. Hierfür sind gut angeleitete Kommunikationsanlässe grundlegend, die nicht nur das Nachdenken über fachliche Inhalte fördern sollen, sondern auch die Reflexion des eigenen Lernens (Ewerhardy et al., 2012; Holodynski et al., 2017; Kunter & Voss, 2011). Indikatoren für kognitive Aktivierung auf Schülerseite können das Erläutern von Lösungswegen, das Austauschen von Argumenten, die Bezugnahme auf andere Beiträge, der Wissenstransfer auf andere Bereiche, das Formulieren von Hypothesen oder inhaltsbezogene Fragen sein (Lipowsky, 2009).

Maßnahmen der Lehrkraft und Verhaltensindikatoren auf der Seite der Lernenden sind bzgl. ihres Potenzials zur kognitiven Aktivierung allerdings kaum ohne fachspezifisches Wissen beurteilbar. Es hat sich gezeigt, dass die Fähigkeit zur kognitiven Aktivierung seitens der Lehrkraft über verschiedene Unterrichtsstunden hinweg sehr variabel ist und insofern stark von den Inhalten und Zielen der jeweiligen Stunde abhängig zu sein scheint (Lipowsky & Bleck, 2019; Praetorius et al., 2014). Zudem ergibt sich im Hinblick auf die Entwicklung eines vignettenbasierten Verfahrens die Problematik, dass sich ohne umfangreiche Zusatzinformationen kaum Informationen zu den kognitiven Voraussetzungen der Lernenden und damit zur Passung der Anforderungen transportieren lassen. Auch größere thematische Zusammenhänge lassen sich durch kurze Ausschnitte von Unterrichtssituationen nur schwer abbilden. Merkmale der kognitiven Aktivierung werden deshalb in ErLe zunächst keine Berücksichtigung finden.

Für die LU durch eine inhaltliche Strukturierung des Lerngegenstands und die Reduktion komplexer Lernsituationen (Holodynski et al., 2017; Meschede et al., 2015; Möller, 2016) ist

zwar auch fachspezifisches Wissen notwendig. Es finden sich jedoch auch Aspekte, für deren Beurteilung weniger Fachwissen bzw. fachdidaktisches Wissen nötig ist und die deshalb in ErLe aufgenommen werden sollen.

Inhaltliche Strukturierung

Gerade bei inhaltlich komplexen Lerngegenständen ist es erforderlich, diese so aufzubereiten, dass Schülerinnen und Schüler nicht überfordert werden und einen guten Überblick über einzelne Teilschritte haben. Dies betrifft die inhaltliche Strukturierung des Lerngegenstands, indem Inhalte sequenziert werden, also in sinnvolle, aufeinanderfolgende Teile gegliedert werden, um so kumulative Verstehensprozesse zu ermöglichen (Holodynski et al., 2017; Meschede, 2014). Zur Strukturierung des Interaktionsgeschehens gehört auch, dass den Lernenden stets klar ist, wie das übergeordnete Ziel der jeweiligen Lerneinheit lautet, indem Lehrkräfte *Zielklarheit* schaffen und den roten Faden im Unterricht erkennbar machen (Rakoczy & Pauli, 2006).

Gerade bei kommunikativen Anlässen wie Unterrichtsgesprächen und beim Erklären muss gewährleistet sein, dass Äußerungen von Lehrkräften inhaltlich klar werden. Der Sprache kommt hier eine besondere Bedeutung zu. Klare Formulierungen in angemessener Länge, eine deutliche Aussprache in der richtigen Lautstärke und im richtigen Tempo sowie das Vermeiden von unnötigen Fachbegriffen oder Erklären von nötigen Fachbegriffen gehören dazu. Auch Schülerinnen und Schüler sollen angeleitet werden, sich klar auszudrücken und Äußerungen sollten, wenn es nötig ist, aufgegriffen und präzisiert werden (Clausen et al., 2003). Eine solche *inhaltliche Klarheit* wird unterstützt durch veranschaulichende Maßnahmen wie das Hervorheben und Zusammenfassen wichtiger Aspekte (z. B. Hogan & Pressley, 1997; Pea, 2004). In einzelnen Teilschritten des Unterrichts soll die Lehrkraft dazu Inhalte und Ergebnisse auf den Punkt bringen und auf Begriffe, Beobachtungen oder Aussagen aufmerksam machen, die für den Verstehensprozess wichtig sind. Zusätzlich sollten Erklärungen durch enaktive oder ikonische Repräsentationen veranschaulicht werden (z. B. Drollinger-Vetter & Lipowsky, 2006). Gerade jüngere Kinder, kognitiv beeinträchtigte Kinder oder sprachlich beeinträchtigte Kinder sind mit rein sprachlichem Input schnell überfordert und ihr Denken und Verstehen ist in einem hohen Maß von einer solchen Anschaulichkeit abhängig. Zusätzlich zur Veranschaulichung kann es den Verstehensprozess unterstützen, wenn Lerninhalte zunächst modelliert werden, indem Lösungswege beispielhaft aufgezeigt werden und die Lernenden so Schritt für Schritt an eigene Lösungswege herangeführt werden (Holodynski et al., 2017).

6.3.3 Lernbarrieren

Die oben beschriebene Auswahl von Merkmalen der KF und LU repräsentiert empirisch belegte generische Qualitätsmerkmale zur Beschreibung von Unterricht. Es wird angenommen, dass eine mangelhafte oder fehlende Umsetzung dieser Maßnahmen als potenzielle Barriere verstanden werden kann.

Die im Projekt ViU verwendeten Manuale (Adamina et al., 2017; Gippert et al., 2019) liefern Beispielindikatoren für die gelungene und weniger gelungene Umsetzung von Maßnahmen der KF und LU und sollen als Grundlage für die Systematisierung von Lernbarrieren dienen. Merkmale der KF sind dabei etwas stärker vertreten, was zum einen mit der besonderen Bedeutung der KF und zum anderen mit der generischen Natur der Merkmale und der besseren Beobachtbarkeit zu rechtfertigen ist (Kunter et al., 2007; Kunter et al., 2013; Lipowsky & Bleck, 2019; Rakoczy, 2008). Die folgende Tabelle gibt einen Überblick über die in ErLe verwendeten Lernbarrieren (siehe Tabelle 3).

Tabelle 3

Überblick über die in ErLe verwendeten Lernbarrieren

Dimension	Facette	Merkmal	Operationalisierung des Merkmals	Operationalisierung der Lernbarriere
Klassenführung	Monitoring	Allgegenwärtigkeit und Reaktion auf Störungen	Die Lehrkraft hat die ganze Klasse im Blick, bekommt mit, was vor sich geht und lässt die Klasse das wissen. Zum Beispiel sieht sie, wenn Schülerinnen und Schüler mit etwas anderem beschäftigt sind, nicht mitarbeiten, bereits mit der Arbeit fertig sind oder Materialien fehlen. Sie reagiert darauf prompt, kurz und bündig und an die richtige(n) Person(en) gerichtet. Je nach Ausmaß der Störung reagiert sie angemessen mit nonverbalen Reaktionen (z. B. Blickkontakt, Handzeichen, Änderung der Position im Raum) oder verbal. Sie kann mehrere Unterrichtsprozesse parallel steuern, z. B. hört sie zu und gibt anderen gleichzeitig zu verstehen, dass sie weitere Meldungen wahrgenommen hat, oder sie kann eine Gruppenarbeit unterstützen, ohne andere Gruppen zu übersehen.	<p>Die Lehrperson schenkt nur einem Prozess Beachtung und versäumt derweil, was ansonsten in der Klasse vor sich geht, z. B. hilft sie bei einer Gruppenarbeit nur einer Gruppe und bemerkt nicht, dass andere auch Probleme haben oder schon fertig sind, oder sie verliert die Klasse während eines Tafelanschriebs aus dem Blick oder sie übersieht Meldungen oder dass Kinder nicht mitarbeiten.</p> <p>Die Lehrkraft unterbindet Störungen (sowohl im Sinne aktiver Störungen als auch im Sinne fehlender Mitarbeit) nicht (unbeachtete Störung) oder zu spät (Unruhe breitet sich aus, Zeitfehler) oder ermahnt Personen, die nicht Verursachende der Störung sind (Objektfehler). Die Lehrkraft reagiert übertrieben auf leichtes Fehlverhalten und unterbricht damit den Unterrichtsfluss oder reagiert zu schwach auf größere Störungen ohne den gewünschten Effekt (Relationsfehler).</p>

	Sitzordnung und Raum	Die Lehrkraft hat die räumlichen Arbeitsbedingungen im Blick und bemerkt z. B., wenn Lernende einen ungünstigen Sitzplatz haben, die Tafel nicht gut erkennen können oder geblendet werden. Sie reagiert entsprechend darauf und setzt z. B. Kinder um oder ändert die Lichtbedingungen.	Die Lehrkraft übersieht räumliche Bedingungen und bemerkt nicht, wenn ein Kind z. B. einen ungünstigen Sitzplatz hat (bzgl. Sitznachbarinnen bzw. Sitznachbarn, Sicht zur Tafel, Lautstärke, Lichtbedingungen oder Ablenkung durch räumliche Bedingungen wie volle Regale). Sie versäumt es, darauf zu reagieren und Bedingungen zu ändern, so dass es zu Störungen kommt, z. B. überträgt ein Kind nicht den Tafelanschrieb und es kommt zu Unruhe.
Prozessuale Strukturierung	Reibungslosigkeit und Schwung/Übergänge	Die Lehrkraft sorgt für ein angepasstes Lerntempo und organisiert nachvollziehbare Übergänge ohne Stocken. Dazu macht sie Unterrichtsaktivitäten transparent und macht deren Beginn und Ende deutlich, z. B. durch eindeutige Instruktionen oder Zeichen wie Klangschale oder Handzeichen. Der Unterricht verläuft reibungslos durch eine gute Vorbereitung (vorbereitete Instruktion, Bereithalten des Lernmaterials) und es werden unnötige Verzögerungen (z. B. Holen fehlenden Materials) vermieden.	Es kommt zu einem stockenden Unterrichtsverlauf und Verzögerungen, weil den Lernenden nicht klar ist, wann eine Unterrichtsaktivität aufhört oder beginnt oder das Tempo unangemessen ist. Sie fangen nicht an zu arbeiten, weil die Instruktion und das Material unzureichend sind oder nicht bereitstehen. Die Lehrkraft verliert Zeit, weil sie während des Unterrichts Versäumnisse nachholen muss und zum Beispiel Material suchen muss. Es entstehen unnötige Wartezeiten oder Unruhe, z. B. weil die Lehrkraft abwartet, statt klar zu kommunizieren, was im Ablauf als nächstes kommt.
Regeln und Routinen	Regeln und Routinen	Die Lehrkraft führt Regeln, Routinen und Rituale ein, etabliert und übt diese und achtet auf deren Einhaltung, z. B. beim Ablauf von Unterrichtsaktivitäten wie Gruppenarbeit (Rollenverteilung in der Gruppe, Verhalten bei Fragen oder für Gruppen die früher fertig	Es kommt zu Störungen oder Verzögerungen, weil Regeln und Routinen fehlen oder nicht eingehalten werden, z. B. langes, lautes Bilden eines Sitzkreises, Streit um Aufgaben bei

			sind), beim Wechsel der Sozialform (tischweise leise in den Sitzkreis gehen) oder bei der Organisation von Materialien (nur Unterrichtsmaterial liegt auf dem Tisch, Ordnungssysteme werden genutzt). Die Schülerinnen und Schüler halten sich auch ohne direkte Anweisung an die vereinbarten Regeln, die Lehrkraft erinnert aber ggf. an die Regeln.	der Gruppenarbeit, Reinrufen oder Ablenkung durch Materialien, die nicht zum Unterricht gehören.
Lernunterstützung	Inhaltliche Strukturierung	Inhaltliche Klarheit	Die Lehrkraft schafft inhaltliche Klarheit, indem sie z. B. Teilschritte und deren Abfolge klarmacht. Sie spricht klar und deutlich zur Klasse gewandt und vermeidet Fremd- und Fachwörter bzw. erklärt diese. Sie achtet auch bei Äußerungen von Lernenden darauf, dass diese verstanden werden.	Die Lehrkraft erklärt Aufgaben zu lang oder komplex, verwendet Fach- oder Fremdwörter oder setzt deren Verständnis voraus. Sie unterteilt Aufgaben nicht in Arbeitsschritte, sondern erklärt mehrere Teilschritte auf einmal. Sie bemerkt Verständnisprobleme oder falsche Vorstellungen nicht und fragt nicht, ob es Rückfragen gibt. Sie lässt unklare Äußerungen von Lernenden unkommentiert.
		Zielklarheit	Die Lehrkraft schafft Zielklarheit, indem sie Zielsetzung und Vorgehen verdeutlicht und das fachliche Lernziel deutlich macht. Den Schülerinnen und Schülern ist der Sinn einzelner Teilschritte und der Bezug zum Ergebnis klar. Die Lehrkraft macht Erwartungen bzgl. der Arbeitsergebnisse deutlich und gibt klare Instruktion, wie Ergebnisse festgehalten oder präsentiert werden sollen.	Die Lehrkraft erklärt eine Aufgabe und vergisst dabei zu verdeutlichen, wie oder wo Ergebnisse aufgeschrieben oder präsentiert werden sollen. Sie vergisst zu verdeutlichen, was das Ziel einzelner Arbeitsschritte ist oder welcher Arbeitsschritt als nächster daraus folgt.
		Veranschaulichende Maßnahmen	Die Lehrkraft veranschaulicht Erklärungen, Aufgaben, Instruktionen oder Beiträge von Lernenden durch Hervorhebungen oder Visualisierung oder hebt wesentliche Informatio-	Die Lehrkraft erklärt nur mündlich oder präsentiert Aufgaben nur als Text. Hervorhebungen im Text (farblich, fett, unterstrichen, Absätze etc.) fehlen. Sie wiederholt auf

Feedback**Feedback****Feedback durch die Lehrkraft**

nen hervor. Sie wiederholt z. B. verbal mit besonderer Betonung, schreibt Wichtiges an, macht Unterstreichungen oder hält Materialien zur Veranschaulichung bereit.

Die Lehrkraft vermittelt positive Präsenz durch ihre zugewandte Position im Raum sowie durch Mimik und Gestik wie zustimmendem Nicken. Sie gibt wertschätzende und konstruktive Rückmeldungen – auch bei Fehlern. Sie ordnet Beiträge der Kinder ein und gibt konstruktive Rückmeldungen, die die Weiterarbeit fördern. Kommentare zum Verhalten oder zu Arbeitsergebnissen sind sachlich und nicht persönlich.

Nachfrage Erklärungen, nimmt zur Veranschaulichung aber keine weiteren Materialien zur Hand. Sie zeigt keine oder unpassende Beispiellösungen.

Die Lehrkraft lässt Äußerungen unkommentiert stehen und gibt gar kein Feedback oder sie gibt unpassendes Feedback, das z. B. persönlich ist, nur Kritik äußert ohne konstruktive Anregungen zu enthalten, Unterstellungen enthält, verletzend ist, Personen bloßstellt oder einzelne Verhaltensweisen pauschalisierend beurteilt.

7 Modellierung der Kompetenz Erkennen von Lernbarrieren

Auf der Basis des Kompetenzmodells (siehe Kapitel 3.2) lässt sich das Erkennen von Lernbarrieren als situationsspezifische Fähigkeit verstehen. Diese setzt Wissen über Maßnahmen der KF und LU und deren Bedeutung für den Lernerfolg voraus. Auch deren Operationalisierung, also die Kenntnis von konkreten Handlungen, die mit Maßnahmen der KF und LU verknüpft sein können, muss vorhanden sein. Da Lernbarrieren verstanden werden als vernachlässigte oder fehlende Maßnahmen der KF und LU, muss die Funktion der Maßnahmen der KF und LU verstanden werden, um schlussfolgern zu können, welche Konsequenzen deren Fehlen oder Misslingen haben.

Wie zuvor ausgeführt, wird das Erkennen von Lernbarrieren als diagnostische Kompetenz verstanden. Diese setzt voraus, dass bei der Analyse einer Lehr-Lernsituation zum einen Wissen darüber vorhanden ist, welche Faktoren zum Gelingen oder Misslingen beitragen können, so dass Unterrichtsmerkmale überhaupt als Gelingensbedingung in Betracht gezogen werden. Zum notwendigen Beobachtungswissen zählt, tatsächlich beobachtbare Handlungen von Eindrücken zu trennen. Des Weiteren muss auch auf affektiv-motivationaler Ebene die Bereitschaft vorhanden sein, Problemursachen nicht nur beim Kind, sondern auch in Unterrichtsbedingungen zu suchen, die wesentlich durch das Handeln der Lehrkraft bestimmt werden. Diese spezifische Fokussierung auf und Beobachtung von Unterrichtsmerkmalen ist eine notwendige Ergänzung zum bestehenden Kompetenzmodell und stellt eine mediiierende Teilkompetenz dar.

Erst diese auf Dispositionen fußende Fokussierung ermöglicht die Wahrnehmung relevanter Unterrichtssituationen, in denen sich Indikatoren für KF und LU zeigen, und die Nutzung dieser als Basis für das Erkennen von Lernbarrieren. Wahrgenommene Indikatoren müssen anschließend hinsichtlich ihrer potenziellen Wirkung beurteilt werden, um diese letztlich als potenzielle Lernbarriere zu erkennen. Dafür müssen verschiedene Wissensanteile verknüpft und Handlungen der Lehrkraft, Umgebungsbedingungen sowie Reaktionen der Lernenden in Zusammenhang gebracht werden. Der Grad der Verknüpfung und die Abrufbarkeit von verschiedenen Wissensanteilen sind dabei ein Zeichen von Expertise (Berliner, 1994; Bromme, 1992; Bromme, 2008). Hinsichtlich der PW verlangt es mehr Expertise, mehrere Perspektiven gleichzeitig zu berücksichtigen wie den Blick auf Schülerinnen und Schüler sowie auf die Lehrperson und die Unterrichtsprozesse (Hiebert & Grouws, 2007; Jacobs, 2017; Sherin & van Es, 2009; Stockero & Rupnow, 2017). Insofern kann die Kompetenz, Lernbarrieren zu erkennen, als anspruchsvoll bezeichnet werden.

Die Herausforderung, Lernbarrieren zu erkennen, ergibt sich auch daraus, dass diese nicht immer unmittelbar sichtbar sind, sondern anhand wahrnehmbarer Handlungen, Verhaltensweisen oder Anzeichen – also sichtbarer *Indikatoren* (Decristan et al., 2020) – auf diese geschlossen

werden muss. Aus Sicht der pädagogischen Psychologie lassen sich in Bezug auf die Unterrichtsqualität Sichtstrukturen (Kunter & Trautwein, 2013; Reyer, 2004) oder auch Oberflächenmerkmale (Decristan et al., 2020) und Tiefenmerkmale unterscheiden. Zu den Oberflächenmerkmalen gehören das konkrete und beobachtbare Handeln von Lehrkräften und ihre Unterrichtspraktiken, z. B. Unterrichtsmethoden wie bestimmte Sozial- und Arbeitsformen (z. B. die Kinder machen Gruppenarbeit; Kunter & Trautwein, 2013, S. 13). Als Tiefenmerkmale werden nach Lipowsky (2006) „eher verdichtete und komplexere Kennzeichen des unterrichtlichen Interaktionsgeschehens“ (S. 55) oder nach Kunter und Ewald (2016) die „Art und Weise, wie sich Lernende mit dem Lernstoff auseinandersetzen“ (S. 14) verstanden (Oser & Baeriswyl, 2001; Pauli, 2012; Seidel, 2003a).

Teilweise wird angenommen, dass Oberflächenmerkmale dabei den Rahmen für die Tiefenmerkmale abstecken. Sie führen aber nicht automatisch zu qualitativem Unterricht. So führen kleinere Klassen beispielsweise nicht automatisch dazu, dass dort besser interagiert wird. Erst im Zusammenspiel mit der qualitativollen Umsetzung dazugehöriger Tiefenmerkmale wird Unterricht wirkungsvoll (Kunter & Ewald, 2016; Kunter & Voss, 2011; Seidel & Shavelson, 2007; Wang et al., 1993). Insofern wird den Tiefenmerkmalen oft die größere Bedeutung für den Lernerfolg zugesprochen (Lipowsky & Bleck, 2019). Andernorts ist die Beziehung zwischen Oberflächen- und Tiefenmerkmalen wechselseitig dargestellt und betrifft nicht nur die Unterrichtsdurchführung, sondern auch die Unterrichtsplanung.

Auch wenn Oberflächenmerkmale nicht automatisch lernwirksam in der Tiefe sind, haben einige Methoden, Materialien oder Sozialformen ein spezifisches Anregungspotenzial, so dass diese, auch wenn sie „zwar primär auf der Oberflächenebene verortet sind,.. durch ihre theoretischen Wirkmechanismen eng mit der Unterstützung kognitiver Prozesse verbunden sind“ (Decristan et al., 2020, S. 112). Tiefenmerkmale können demnach verstanden werden als „vermittelnde oder moderierende Variablen mit Bezug auf die Wirksamkeit von Oberflächenmerkmalen“ (Decristan et al., 2020, S. 112).

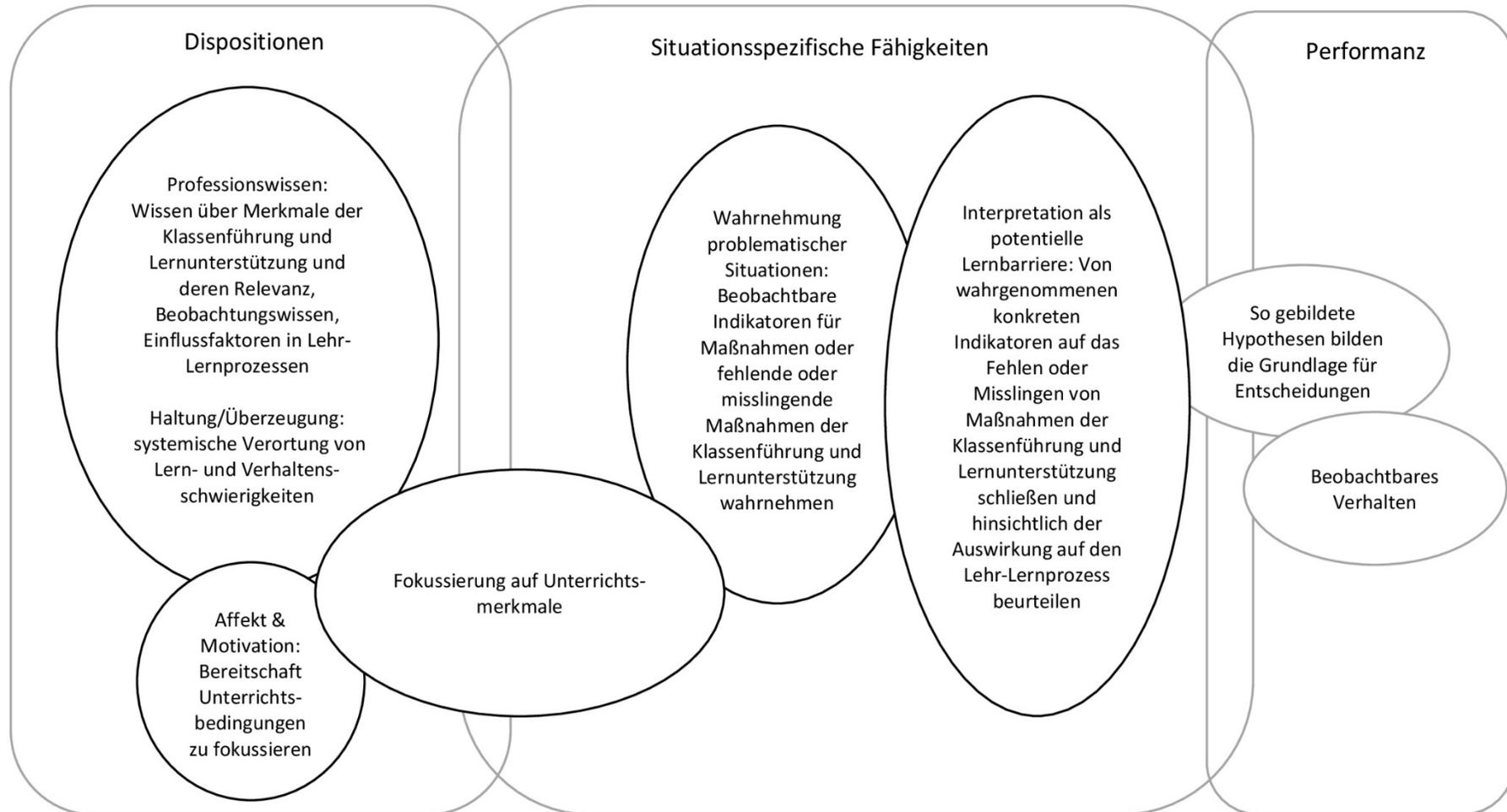
Für das Erkennen von Lernbarrieren heißt das, dass anhand beobachtbarer Indikatoren auf Lernprozesse und auf sie einwirkende Variablen – auch in der Tiefe – geschlossen werden muss. Dabei soll die Verknüpfung von beobachtbaren Indikatoren mit Merkmalen der KF und LU nicht suggerieren, dass auch in der Theorie bestimmte Oberflächen- und Tiefenmerkmale einander fest zuzuordnen sind. Vielmehr sollen im Rahmen der hier entworfenen Konzeptualisierung von Lernbarrieren aufgrund beobachtbarer Indikatoren Annahmen über *potenzielle* Lernbarrieren getroffen werden (vgl. Decristan et al., 2020). Die im vorangegangenen Kapitel dargestellte Operationalisierung bietet dafür den Zugang:

So liegt in der (Unterrichts-)Forschung stets eine zentrale Herausforderung darin, nicht direkt sichtbare (latente) Merkmale, wie psychologische Prozesse des Lehrens und Lernens, zu erfassen. Diese müssen letztlich für eine empirische Prüfung stets mit Hilfe von Operationalisierungen über (Verhaltens-) Indikatoren zugänglich gemacht und somit immer auf eine sicht- und interpretierbare Ebene gehoben werden – etwa um die „rezeptive Funktion“ von Tiefenmerkmalen abzubilden. (Decristan et al., 2020, S. 111)

Das Erkennen von Lernbarrieren erfordert also das Wahrnehmen problematischer Situationen anhand konkreter Interaktionen, deren Interpretation und die Vorhersage bzgl. der Auswirkung als potenzielle Lernbarriere (vgl. Seidel et al., 2010; Seidel & Stürmer, 2014; Stürmer & Seidel, 2017). Die Wahrnehmung von Störungen ist dafür nur ein erster Schritt. Wichtig sind dann die Interpretationen der Beobachtungen, um tatsächlich von einem *Erkennen* sprechen zu können. Die Identifizierung potenziell problematischer Interaktionen beinhaltet mehr als einen holistischen Eindruck (z. B. die Aussage: *Die Lehr-Lernsituation verläuft insgesamt problematisch*). Stattdessen wird eine konkrete Hypothesenbildung zu relevanten Handlungen erforderlich. Die folgende Abbildung modifiziert das grundlegende Kompetenzmodell von Blömeke et al. (2015a) hinsichtlich dieser Aspekte (siehe Abbildung 3).

Abbildung 3

Kompetenzmodell Erkennen von Lernbarrieren in Anlehnung an Blömeke et al. (2015a, S. 7)



8 Konstruktionskriterien für Vignetten und Items

Um das Erkennen von Lernbarrieren anregen und erfassen zu können, müssen Kriterien für die Vignetten- und Itemkonstruktion festgelegt werden. Es gibt vielfältige Varianten der Konstruktion von und Arbeit mit Vignetten, die im Folgenden erörtert werden sollen, um den geeigneten Zugang für das vorliegende Forschungsvorhaben zu wählen (vgl. Friesen et al., 2018; Zucker, 2019).

8.1 Vignettenformat

Zur Erfassung der PW wurden bislang vor allem authentische oder konstruierte Videovignetten eingesetzt (Colestock & Sherin, 2009; Kersting et al., 2012; Meschede et al., 2015; Sherin, 2007; Stürmer & Seidel, 2017). Möglich sind aber auch Textvignetten (Brovelli et al., 2013; Dreher & Kuntze, 2015; Moje & Wade, 1997) sowie Animationen, Comics oder Bildvignetten (vgl. Herbst & Kosko, 2013). Im Unterschied zu Unterrichtsbeobachtungen im Klassenraum haben Vignetten aller Formate den Vorteil, dass sie wiederholt und von zahlreichen Personen gleichzeitig angeschaut werden können und sich so für standardisierte Erhebungen eignen (Jacobs, 2017; Seifried & Wuttke, 2017; Zucker, 2019). Die Vignettenformate haben verschiedene Vor- und Nachteile:

Herbst und Kosko (2013) schlagen zur Unterscheidung von Formaten unterschiedliche Kategorien vor. Eine davon ist die *Temporalität*. Während Videos reale zeitliche Abläufe wiedergeben, stellen Texte und andere Abbildungen wie Comics Situationen ohne zeitlichen Rahmen dar und die Rezeption hängt von den Probandinnen und Probanden ab. So können sie Bilder kurz oder lang betrachten oder Texte langsam oder schnell, einmal oder wiederholt gelesen werden. Dabei entsteht eine individuelle Vorstellung des zeitlichen Geschehens.

Mit *Individualität* bezeichnen Herbst und Kosko (2013) die Konkretheit der Darstellung. Im Video ist diese am höchsten, da Personen, Handlungen, Räume etc. detailliert dargestellt sind. In Texten dagegen können die gleichen Situationen geschildert werden, ohne dass Informationen zu z. B. räumlichen Gegebenheiten oder Personen genannt werden. Äußere Gegebenheiten und nonverbale Aspekte sind dann der Vorstellung der Lesenden überlassen. Bildliche Darstellungen wie Comics können nach Friesen und Kuntze (2016) als dazwischenliegend betrachtet werden, da z. B. Gesichtsausdrücke oder äußere Gegebenheiten unterschiedlich detailliert dargestellt werden können.

Eine weitere Kategorie zur Unterscheidung von Vignettenformaten ist die *Modalität* visuell vs. audiovisuell und stehendes vs. bewegtes Bild. Dies beeinflusst unter anderem den Transport relevanter Informationen. Während in Videos und Comics bildhafte Informationen wie z. B.

räumliche Gegebenheiten in andere Informationen eingebettet sind, müssen diese in Texten benannt werden und treten so stärker hervor (vgl. Herbst & Kosko, 2013). Barth (2017, S. 82) weist allerdings darauf hin, dass auch die Dauer und Perspektive bestimmter Einstellungen im Video einen Einfluss darauf nehmen können, wie sehr die Aufmerksamkeit auf relevante Ereignisse gelenkt wird. Außerdem werden Informationen in Videos im Gegensatz zu anderen Formaten gleichzeitig und nicht nacheinander und nur einmalig präsentiert. In Videos werden zudem sehr viel mehr Informationen transportiert, die zwar Urteile beeinflussen können, in Bezug auf die zu erkennenden Merkmale aber irrelevant sind, wie z. B. das Aussehen von Personen (Bönte et al., 2019). Bei Texten und Comics ist das Maß an Kontextinformationen dagegen leichter zu kontrollieren (Friesen & Kuntze, 2016; Friesen et al., 2018; Herbst & Kosko, 2013).

Für die Beurteilung von Vignetten könnte dies bedeuten, dass Videos aufgrund der zahlreichen, komplexen und schnell aufeinanderfolgenden Informationen entweder deutlich anspruchsvoller zu analysieren sind als Texte oder aber auch besser zu analysieren sind, da die Situationen besonders gut vorstellbar sind (Friesen & Kuntze, 2016; Friesen et al., 2018). Die überwiegende Anzahl der Studien zur PW argumentiert hierbei für Videovignetten, weil sie Unterrichtssituationen besonders authentisch und realitätsnah darstellen und die Komplexität von Unterricht am besten wiedergeben können (Blömeke et al., 2015b; Bönte et al., 2019; Kaiser et al., 2015; Krammer et al., 2012; Lindmeier, 2013; Lipowsky, 2010; Seidel et al., 2011; Seifried & Wuttke, 2017).

Friesen et al. (2018) führen die Förderung von *Immersion* (wie gut kann man sich in die Unterrichtssituation hineinversetzen), *Motivation* und *Resonanz* (wie gut kann man die Unterrichtssituation mit eigenen Erfahrungen verknüpfen) als besondere Vorteile der Arbeit mit Vignetten auf (S. 159). Allerdings kann nicht eindeutig belegt werden, dass Videos dabei einen Vorteil gegenüber Textvignetten oder Comics haben. Vielmehr stellt sich in mehreren Studien die Arbeit mit unterschiedlichen Formaten als gleichwertig heraus (Friesen et al., 2018; Gold et al., 2016a; Herbst & Kosko, 2013; Herbst et al., 2013; Hoppe et al., 2020; Kramer et al., 2017; Schneider et al., 2016). Teilweise zeigen sich auch Hinweise darauf, dass Videos aufgrund der hohen Komplexität und Synchronität und der damit einhergehenden kognitiven Belastung schlechter zu bearbeiten sind als andere Formate (z. B. Sweller & Chandler, 1994; Syring et al., 2015). Schneider et al. (2016) kommen in ihrer Studie zur Auswirkung unterschiedlicher Medien beim fallbasierten Lernen so auch zu anderen Ergebnissen. Studierende, die mit Textvignetten gearbeitet hatten, wiesen hier bessere Ergebnisse in der Qualität der Interpretation wahrgenommener Unterrichtsszenen auf als Teilnehmende der Videogruppe. Auch Schneider et al. (2016) führen dies auf eine erhöhte kognitive Belastung durch Videos zurück. Sie bemerken einschränkend aber auch, dass dies möglicherweise damit zusammenhängt, dass an der Studie Studienanfänger

und -anfängerinnen teilnahmen, während in Studien mit anderen Ergebnissen (z. B. Stürmer et al., 2012) Studierende höherer Semester teilnahmen. Bezüglich der kognitiven Belastung bei der Bearbeitung von Textvignetten existieren weniger Befunde. Gold et al. (2016a) stellen in ihrer Studie eine erhöhte kognitive Belastung durch Textvignetten fest, wobei sie sehr lange Texte von etwa sechs DIN A4-Seiten einsetzten, in denen relativ detailliert nonverbale Informationen beschrieben waren und zusätzlich Zeichnungen zu Sitzordnung und Tafelposition präsentiert wurden, um den Informationsgehalt dem Video anzugleichen.

Für Textvignetten spricht, dass diese optimal auf das jeweilige Zielkonstrukt abgestimmt werden können und zugleich Störfaktoren besser kontrolliert werden können. Zudem ist die Komplexität anpassbar. Auch längere Unterrichtssequenzen oder Zeiträume lassen sich kompakt und zusammenhängend darstellen (Brovelli et al., 2013; Riese & Reinhold, 2012). Dass simultan stattfindende Unterrichtsereignisse gezwungenermaßen nacheinander geschildert werden müssen, kann die kognitive Belastung verringern. Dies stellt einen Vorteil dar, da die Vignetten damit auch von Novizinnen und Novizen besser be- und verarbeitet werden können (Schneider et al., 2016).

Weiterhin spricht der ökonomische Aspekt für den Einsatz von Textvignetten. Diese sind einfacher und schneller herzustellen und können im Entwicklungsprozess leichter angepasst werden (Rutsch et al., 2017). Gerade in der Entwicklung eines Verfahrens muss geprüft werden, ob die Darstellung der gewünschten Merkmale in der Vignette gelungen ist, so dass mehrfache Änderungen oder Neukonstruktionen nötig werden können. Dies ist mit Textvignetten unkompliziert möglich.

Bei der Entscheidung für Videos oder Textvignetten sollten auch rechtliche und ethische Aspekte beachtet werden. Für Videos mit realen Unterrichtsszenen sind Einverständniserklärungen einzuholen und die gezeigten Sequenzen müssen für alle Beteiligten unbedenklich und vertretbar sein. Da es in der vorliegenden Arbeit aber um das Erfassen von Lernbarrieren geht, also das Erkennen problematischer Lehr-Lernsituationen, erscheint die Verwendung von realen Unterrichtsvideos ethisch problematisch. Inszenierte Videos sind in ihrer Vielschichtigkeit und Komplexität mit mehreren Akteurinnen und Akteuren aufwändig und schwierig umzusetzen. Zudem ist die Authentizität der Sequenzen auch abhängig von den schauspielerischen Darbietungen. In Textvignetten können dagegen auch barrierehaltige Unterrichtssituationen ohne Bloßstellung der Akteurinnen und Akteure dargestellt werden.

Textvignetten können anhand von Transkripten realer Unterrichtsszenen sowie modifizierten Transkripten oder Transkripten von fiktiven Unterrichtssequenzen konstruiert werden (von

Aufschnaiter et al., 2017; Gold et al., 2013; Spitzer et al., 2011; Yadav et al., 2011). Die Verwendung realer Unterrichtsbeispiele kann dabei für die nötige Authentizität sorgen. Dabei können in Textvignetten auch mehrere Unterrichtsbeispiele kombiniert oder modifiziert werden, so dass die dargestellten Szenen dichter werden, ohne den Realitätsbezug zu verlieren. Auf diese Weise konstruierte Fälle können auf das Zielkonstrukt abgestimmte, typisierte Situierungen darstellen (von Aufschnaiter et al., 2017; Brovelli et al., 2014; Friesen et al., 2018; Keller et al., 2018; Meschede & Steffensky, 2018).

Solche Unterrichtsbeschreibungen enthalten auch Beschreibungen nonverbaler Informationen und weitere Beschreibungen der Lernsituation und Umgebung (z. B. Sitzordnung, räumliche Besonderheiten, Geräusche). Über solche Informationen lässt sich steuern, wie komplex die Vignette wird und welche Merkmale in die Beschreibung Eingang finden. Aspekte, deren Wahrnehmung es zu überprüfen gilt, sind somit vorselektiert und der Schwierigkeitsgrad einer Vignette kann angepasst werden (Schneider et al., 2016).

8.2 Komplexität der Vignette

Eine besondere Herausforderung bei der Konstruktion der Vignette besteht darin, diese nicht zu komplex zu gestalten und mit Merkmalen zu überladen (vgl. Brovelli et al., 2013), sie aber auch nicht zu stark zu reduzieren und zu vereinfachen, so dass sie an Authentizität verliert. Denn reale Unterrichtssituationen zeichnen sich ja gerade durch Komplexität und Uneindeutigkeit aus (Brovelli et al., 2013; Spiro et al., 1991). Soll das Erkennen voneinander *abgegrenzter* Informationen gut handhabbar sein, muss die Komplexität reduziert werden, so dass Vignetten kurze und wenig komplexe Situationen darstellen und weniger gut die PW realer Unterrichtssituationen repräsentieren (Blömeke et al., 2010; Brovelli et al., 2013). Enthält eine Vignette mehrere Merkmale, wird die ökologische Validität erhöht, wobei die Abgrenzung der einzelnen Merkmale aber schwieriger wird (Brovelli et al., 2013).

Zu berücksichtigen ist auch die Verarbeitungskapazität derjenigen, die solche komplexen Beschreibungen beurteilen sollen. Expertinnen und Experten sind dabei im Vorteil, weil sie auf mehr und besser organisiertes Wissen zurückgreifen können und deshalb relevante Aspekte genauer wahrnehmen können. Sie verarbeiten Informationen eher wissensgesteuert und fokussieren relevantere Informationen (König, 2015). Das besser vernetzte Wissen ermöglicht gleichzeitig eine holistischere Perzeption der Situation. So ist zu erwarten, dass Expertinnen und Experten die Vignette schon aufgrund der besseren Rekonstruktion der Situation besser bearbeiten können (König, 2015). Sollen auch Novizinnen und Novizen die Situation gut rekonstruieren können, ist eine angemessene Komplexität umso wichtiger. Ist die Vignette zu lang und enthält sie zu viele zu beurteilende Merkmale, besteht die Gefahr, dass die Bearbeitung

abgebrochen wird oder Urteile möglicherweise nur noch Artefakte darstellen (Auspurg et al., 2009). Ein möglicher Fehler bei zu komplexen Vignetten könnte sein, dass unauffälliger operationalisierte Merkmale ausgeblendet werden und eher heuristisch beurteilt wird (Auspurg et al., 2009). Zu wenige Merkmale könnten dagegen dazu führen, dass urteilende Personen fehlende, urteilsrelevante Informationen selber konstruieren, so dass Urteile ungenauer und schwieriger zu kontrollieren sind (Auspurg et al., 2009).

Vignetten sollten also möglichst nur die vorher festgelegten Merkmale in nicht zu großer Zahl enthalten (Auspurg et al., 2009; Jasso, 2006). Als ungefähre Richtwert wird zum Beispiel für *Factorielle Surveys* in den Sozialwissenschaften aus den Informations- und Kognitionswissenschaften die Anzahl von fünf bis neun Informationen pro Vignette herangezogen (Auspurg et al., 2009; Zimbardo, 1988). Brovelli et al. (2013) ermitteln als sinnvolle Zahl drei bis neun zu erkennende Merkmale pro Vignette.

Eine Alternative stellen Vignetten mit geringer Komplexität dar, die Merkmale nicht explizit enthalten, sondern eine eher allgemeine Situation beschreiben wie einen Klassenstatus oder Personenkonstellationen (vgl. Przibilla et al., 2016). Werden zu dieser Art Impuls Fragen gestellt, können die Antworten dahingehend analysiert werden, welches Wissen im Zusammenhang damit genutzt wird, welche Einstellungen deutlich werden oder welche subjektiven Theorien herangezogen werden, ohne dass auf diese explizit verwiesen wird. Wird in Bezug auf beschriebene Personen nach Urteilen oder Entscheidungs- und Handlungsoptionen gefragt (z. B. *Wie beurteilen Sie das Handeln von xy* oder *Wie würden Sie an dieser Stelle entscheiden/handeln?*) wird davon ausgegangen, dass eine *Projektion* stattfindet. Die Lesenden schreiben den Personen eigenes Wissen, Gefühle, Motive und Verhaltensweisen zu und machen auf diese Weise stellvertretend deutlich, wie sie urteilen und handeln würden (Stiehler et al., 2012). Die *kognitionspsychologische Skripttheorie* beschreibt, dass Vignetten-Fälle Teil kognitiver Skripte sind, die von individuellen Erfahrungen geprägt sind. Das Bearbeiten von Fall-Vignetten soll individuelle Wahrnehmungs- und Handlungsschemata sichtbar machen, die mit diesen kognitiven Skripten zusammenhängen (Stiehler et al., 2012). Den verschiedenen Konzepten ist gemeinsam, dass Vignetten als Anlass dienen, sich mit einer geschilderten Situation auseinanderzusetzen, was immer vor dem Hintergrund individueller Voraussetzungen passiert. Die Vignette muss also geeignet sein, die gewünschten Schemata anzusprechen und Urteilsprozesse in Gang zu setzen.

Für detailliertere Vignetten, die Handlungen zu den zu erkennenden Merkmalen enthalten, sollen im folgenden Abschnitt einige vor allem sprachliche Gestaltungsfragen erörtert werden.

8.3 Sprachliche Gestaltung und Informationsgehalt der Vignette

Um einen großartigen Film zu machen, brauchst du drei Dinge – das Drehbuch, das Drehbuch und das Drehbuch.
(Alfred Hitchcock)

Die Entwicklung von Textvignetten soll unter Berücksichtigung der oben beschriebenen theoretisch und empirisch begründeten sowie praxisrelevanten Maßgaben erfolgen und relevante inhaltliche Merkmale enthalten (vgl. Keller et al., 2018). Neben der Formulierung geeigneter Aufgaben (Bühner, 2006; Tepner & Dollny, 2014) werden Vignettentexte geschrieben. Unterrichtssituationen sollen möglichst authentisch und gut vorstellbar beschrieben werden, weil die Lesenden zur Bearbeitung der Vignette die Situation rekonstruieren müssen. Insofern lässt sich das Schreiben einer Textvignette mit dem Schreiben eines Drehbuchs vergleichen. Auch wenn mit Hilfe der Textvignette kein Film gedreht wird, soll sie dazu führen, dass sich die Rezipientinnen und Rezipienten den Inhalt möglichst gut vorstellen können – wie einen Film vor dem inneren Auge. All das, was nicht der Improvisation überlassen bleiben soll, muss beschrieben werden.

Der Biograf Donald Spoto (2010) beschreibt, dass Alfred Hitchcock in seinen Drehbüchern sehr präzise und detailreiche Anweisungen gab. Lesende werden dadurch stark in der Rekonstruktion der Situation unterstützt. Für die Beurteilung von Vignetten bedeutet dies, dass relevante Unterrichtssituationen mehr oder weniger deutlich werden. Das betrifft zum Beispiel die Direktheit der Information. Umgebungsbedingungen können konkret benannt werden, z. B. *Im Klassenraum ist es laut* oder *Paul sitzt mit dem Rücken zur Tafel* oder es können Handlungen beschrieben werden, die auf solche Bedingungen schließen lassen, z. B. *Mehrere Kinder halten sich die Ohren zu* oder *Der Lehrer sagt: „Paul, guck‘ mal nicht auf dein Heft, sondern dreh‘ dich zur Tafel um!“*. Bei der indirekten Beschreibung anhand von Reaktionen der Akteurinnen und Akteure auf die Umgebungsbedingungen wird gleichzeitig deutlich, wie sich diese Bedingungen auswirken. In Bezug auf das Erkennen von Lernbarrieren werden hier also Störungen durch Umgebungsbedingungen in der Interaktion erkennbar, während die konkrete Benennung von Störfaktoren leichter anhand von Wissen (z. B. Lärm kann beeinträchtigend wirken) erfasst werden kann. Beinhaltene Vignetten Handlungsbeschreibungen entspricht dies eher der Beobachtung realer Unterrichtssituationen. Lernbarrieren müssen dann in der Vignette so wie im realen Unterricht anhand beobachtbarer Indikatoren auf Handlungsebene wahrgenommen und hinsichtlich ihrer Auswirkungen interpretiert werden. Es erfolgt kein expliziter Hinweis auf potenzielle Störfaktoren.

Des Weiteren ist zu entscheiden, ob direkte Rede verwendet wird. Vogt (2013) beschreibt als wichtigste Wirkung, dass direkte Rede unmittelbar wirkt und Lesende die Figur selbst vernehmen. Einem erzählten Text gibt die direkte Rede so einen dramatischen Akzent und zieht die Lesenden stärker in das Geschehen hinein. Die erzählende Person tritt hinter die Figuren zurück, so dass direkt und ohne Wertung oder andere Beeinflussung das Geschehen unmittelbarer transportiert werden kann. Handlungen werden durch den Wortwechsel transportiert bzw. der Wortwechsel entspricht der Handlung. Die Verwendung der indirekten Rede würde dagegen nicht nur weniger authentisch wirken, sondern auch umständliche Satzkonstruktionen im Konjunktiv mit sich bringen.

Der Einsatz wörtlicher Rede in der Vignette kann unmittelbar wirken und charakterisiert durch Ausdruck und Wortwahl die handelnden Figuren, ohne dies direkt zu benennen. In vielen Textsorten wie wissenschaftlichen oder journalistischen Texten wird wörtliche Rede üblicherweise in Anführungszeichen gesetzt, was die wörtliche und authentische Wiedergabe des Gesagten markiert. Direkte Rede vermittelt so den Eindruck der Authentizität. Zusätzliche Informationen zur direkten Rede beeinflussen dagegen die Rekonstruktion der Situation. Dazu gehören Verben, Adjektive oder andere Beschreibungen in Begleitsätzen, z. B. *Die Lehrkraft sagt, Die Lehrkraft schreit, Die Lehrkraft sagt freundlich* oder *Die Lehrkraft erklärt langsam* (vgl. Katelhoen & Moroni, 2018). Die Verwendung der wörtlichen Rede und die gezielte Formulierung der Begleitsätze erzeugt also Authentizität und transportiert gezielt Informationen. Die Schwierigkeit der Wahrnehmung wird davon beeinflusst, ob relevante Merkmale in einer Umschreibung benannt werden (z. B. *Die Lehrkraft erklärt die Aufgabe stockend und mit vielen Fremdwörtern*) oder in der Handlung – hier also in der wörtlichen Rede – erkennbar werden (z. B. *Die Lehrkraft sagt: Also, ähm, ihr bildet jetzt Hypothesen und dann – also, äh, vor dem Experiment – und dann...*).

Wie anspruchsvoll die Beurteilung der Vignette ist, hängt auch davon ab, welcher Subtext durch die direkte Rede transportiert wird, d. h., was indirekt gesagt, aber nicht ausgesprochen wird. Zum Beispiel kann die Lehrkraft sagen: *Jetzt hab' ich mich so lange um Paul gekümmert, dass ich gar nicht mitbekommen habe, ob ihr schon alle fertig seid* oder diese Aussagen im Subtext transportieren, z. B. *Wie weit seid ihr denn eigentlich?*

Wie gut sich ein Merkmal direkt oder indirekt darstellen lässt, ist auch abhängig vom jeweiligen Merkmal. Feedback lässt sich beispielsweise besser in wörtlicher Rede vermitteln, z. B. *Du gibst dir einfach keine Mühe*, weil eine Beschreibung wie *Die Lehrkraft wirft dem Kind mangelnde Motivation* vor die Beurteilung bereits vorwegnehmen würde. Regeln und Routinen können dagegen beispielsweise gut aus beschriebenen Handlungen erschlossen werden, z. B. *Die Kinder gehen tischeweise in den Sitzkreis*. Dabei sollen die geschilderten Fallbeispiele der Vignetten die

Beobachtung eines realen Unterrichts ersetzen. Bei der sprachlichen Gestaltung ist also auch darauf zu achten, dass Handlungsbeschreibungen beobachtbare Informationen beinhalten. Während *Paul legt den Stift hin, lehnt sich zurück und verschränkt die Arme* beobachtbar ist, würde *Paul verweigert die Bearbeitung des Arbeitsblatts* bereits einer Interpretation entsprechen und über die bloße Beobachtungsebene hinausgehen.

Abzuwägen ist auch die Nennung von Kontextinformationen. Beschreibungen wie *Im Klassenraum riecht es nach Essen, Die Vorhänge sind blau, Draußen scheint die Sonne* oder *Es ist Freitag, 6. Stunde, die Klasse hat gerade eine Klassenarbeit geschrieben* erhöhen zum einen die Komplexität und beeinflussen zum anderen Beurteilungsprozesse allein dadurch, dass sie genannt werden. Die bloße Erwähnung kann suggerieren, dass sie von Bedeutung sein könnten. Umfassende Kontextinformationen, z. B. zu Voraussetzungen der Lernenden oder sozialen Beziehungen innerhalb der Klasse, sind in einer Vignette nicht möglich, weil diese dann zu umfangreich werden würde. Zur Bearbeitung einer Vignette ist es also nötig, dass sich die Lesenden auf die Beurteilung eines kurzen Ausschnitts mit weniger Informationen einlassen. Die Vignetten müssen dafür so konstruiert sein, dass eine Beurteilung anhand der gegebenen Informationen auch unabhängig vom Kontext möglich ist, zumindest hinsichtlich einer *potenziell* positiven oder negativen Auswirkung auf die beschriebene Lehr-Lernsituation.

Für die vorliegende Studie eignen sich am besten Vignetten in Textform, weil sich diese ökonomisch produzieren, variieren und anpassen lassen. Sie können Lernbarrieren in Unterrichtsbedingungen gezielt und ohne Bloßstellung der Akteurinnen und Akteure darstellen. Als Fallbeispiel können allgemeine Situationsbeschreibungen dienen (bereits vorhanden, z. B. von Przibilla et al., 2016) oder detailliertere Unterrichtssituationen, die zu konstruieren sind. Als Vorlage für eine möglichst große Authentizität sollen reale Unterrichtsbeispiele dienen, die aber den dazustellenden Merkmalen entsprechend modifiziert und verdichtet werden, um eine angemessene Komplexität zu erzeugen. Bei der Konstruktion solcher Vignetten sind die Anzahl der Informationen und deren sprachliche Gestaltung relevant für die Komplexität und Authentizität der Vignette. Zudem beeinflussen sie den Anspruch, der an die PW der fokussierten Merkmale gestellt wird. Die Anforderung an die beurteilenden Personen lässt sich über die unterschiedlich deutliche Darstellung der Barrieren in den Vignetten variieren. Es ist weiter zu klären, welche Vor- und Nachteile kurze Situationsbeschreibungen im Vergleich zu stärker strukturierten Impulsen haben.

8.4 Vignette Philipp

Basierend auf den vorangehenden Überlegungen wurde die folgende Vignette konstruiert. Sie enthält acht Hinweise auf misslingende Maßnahmen der KF und LU (Masterlösung siehe Anhang A):

Philipp sitzt während des Matheunterrichts in seiner Klasse an seinem Gruppentisch. Vor ihm liegt ein Arbeitsblatt mit einer Textaufgabe. Die anderen Kinder an seinem Tisch haben bereits in Gruppenarbeit mit der Bearbeitung der Aufgabe begonnen und diskutieren lebhaft. Philipp soll die Aufgabe mit Hilfe seiner Lehrerin lösen.

Lehrerin: „Wir beide gucken jetzt mal zusammen, wie du die Textaufgabe lösen kannst. Hier steht:

„Für die Musik-Aufführung in der Schulturnhalle stellt der Hausmeister 280 Stühle auf. Außerdem benötigt der Elternbeirat noch 20 Stühle für den Verkauf. Im Orchester spielen 24 Kinder, die während der Aufführung stehen. Die Stühle sind in 6er Stapeln im Material-Lager gestapelt.“

Okay?“ (wartet kurz, Philipp nickt) „Dann wollen wir jetzt mal überlegen, wie du das angehen kannst. Ich würde ganz gut finden, wenn du laut denkst, damit ich schauen kann, welche Rechenstrategien du dir überlegst und wie wir die nutzen können. Und dann kann ich dir immer dann helfen, wenn du allein nicht weiterkommst. Was haben wir denn da jetzt alles für Informationen? Vielleicht fängst du mal mit der Frage an?“

Die Lehrerin bleibt neben Philipp stehen und wartet, Philipp schaut auf den Text und bleibt still.

Dann rempelt Philipp seinen Sitznachbarn an und schimpft: „Mann, jetzt rück doch mal ein Stück weg!“

Lehrerin: „Na, na, Philipp, du hast so viel Platz wie immer. Nun fang mal an!“

Philipp: „Und wie?“

Lehrerin: „Wie gesagt, vielleicht mit der Frage. Versuch mal.“

Philipp: „Wie geht die denn?“

Lehrerin: „Das wollen wir ja jetzt rausfinden. Wie könnten wir vorgehen?“

Philipp schaut seine Lehrerin erwartungsvoll an, diese wartet geduldig.

Ein Kind vom Nachbartisch ruft derweil: „Wir sind schon fertig!“

Philipp knallt sauer den Stift auf den Tisch: „Oh Mann, ich hab' keine Lust! Ich mach das nicht!“

Lehrerin: „Na, komm. Wenn du dich schon wieder von vornherein weigerst, dann kommen wir nie weiter...“

8.5 Erprobung einer Kurzvignette

Neben Beschreibungen von Unterrichtsszenen sind wie oben beschrieben als Impuls auch Kurzvignetten denkbar, die keine Informationen zu konkreten Handlungen enthalten, sondern einen Klassenstatus oder bestimmte Konstellationen beschreiben. Przibilla et al. (2016) setzen in ihrem *mixed-method* Online-Survey solche Fallvignetten als Impuls ein, um Wissen und Einstellungen zur Inklusion von Lehrkräften zu erheben. Da mit ihrem Online-Survey ein reliables und valides Instrument vorliegt, das ebenfalls kurze Textvignetten als Impuls einsetzt, wurde eine Kurzvignette aus diesem Instrument gewählt. Die eingesetzten Fallvignetten enthalten unterrichtsbezogene Situationsbeschreibungen und verweisen auf Schülerinnen und Schüler mit unterschiedlichem Förderbedarf. Dazu sind zwei Items zu bearbeiten, die Strategien zur Gestaltung der Lernsituation abfragen. Als Prompts sind diese Fallvignetten also geeignet, eine potenziell problematische Lehr-Lernsituation vorstellbar zu machen und vorhandene Wissens- und Einstellungsaspekte zu inklusivem Unterricht zu aktivieren. Eine dieser Beschreibungen lautet:

Stellen Sie sich vor, Sie sind ein/e Lehrer/-in in einer inklusiven Klasse, die aus 20 Schülerinnen und Schülern ohne Behinderungen und 4 Schülerinnen und Schülern mit Behinderungen besteht. Ein Schüler hat eine spezifische Lernstörung und hat Schwierigkeiten beim Entziffern, in der Leseflüssigkeit und im Leseverständnis. Durchschnittlich liegt der Schüler im Vergleich zu seinen Mitschülern im Lesen etwa drei Jahre zurück und kann keinen der Klassentexte eigenständig lesen oder verstehen. (Przibilla et al., 2016, S. 42)

Diese Fallvignette wurde geringfügig geändert. Zum einen wurde das beschriebene Kind namentlich benannt, damit es in anschließenden Testteilen (s. u.) wiederzuerkennen ist. Zweitens wurde auf die konkrete Zeitangabe „drei Jahre“ verzichtet, um die Vignette von einer bestimmten Klassenstufe zu lösen. Die geänderte Version lautete:

Stellen Sie sich vor, Sie sind ein/-e Lehrer/-in in einer inklusiven Klasse, die aus 20 Schülerinnen und Schülern ohne Behinderungen und 4 Schülerinnen und Schülern mit Behinderungen besteht. Einer Ihrer Schüler heißt Philipp. Philipp hat eine spezifische Lernstörung. Er hat Schwierigkeiten beim Entziffern, in der Leseflüssigkeit und im Leseverständnis. Durchschnittlich liegt Philipp im Vergleich zu seinen Mitschülern im Lesen deutlich zurück und kann keinen der Klassentexte eigenständig lesen oder verstehen.

8.6 Qualitative und quantitative Ansätze

Vignetten aller Formate eignen sich sowohl für qualitative Zugänge wie offene Antwortformate und Interviews als auch für quantitative Zugänge. Häufig werden offene Antwortformate gewählt, um subjektive Wahrnehmungen und Deutungen erfassen zu können (Baer & Buholzer, 2005; Brovelli et al., 2013; van Es & Sherin, 2008; van Es & Sherin, 2010; Pissarek & Schilcher, 2015; Santaga et al., 2007). Sie weisen nicht auf bestimmte Aspekte hin und geben insofern keine Interpretationsrichtung vor, sondern bieten Raum für individuelle Beurteilungen der geschilderten Situation innerhalb eines eigenen Referenzsystems (Schnell et al., 2018). Das kann z. B. gewinnbringend zur Exploration von Deutungen eingesetzt werden. Gleichzeitig muss aber auch bedacht werden, dass die Antworten nicht nur von diesen zur Verfügung stehenden Referenzsystemen abhängen, sondern auch von der Artikulations- und Schreibfähigkeit der Probandinnen und Probanden (Schnell et al., 2018). Zudem bedeutet die qualitative Analyse von Daten einen großen Aufwand (Barth, 2017; Stürmer & Seidel, 2015). Für eine objektive und zeitökonomische Datenerfassung und -auswertung eignen sich quantitative Formate mit geschlossenem Antwortformat besser (Kersting, 2008; Rutsch et al., 2018a; Seidel & Stürmer, 2014; Tepner & Dollny, 2014).

Quantitative Herangehensweisen werden meist mit Aussagen, die anhand mehrstufiger Ratingskalen zu beurteilen sind, umgesetzt, z. B. bei Meschede (2014): „Die LP [Lehrperson] überprüft, ob die SuS [Schülerinnen und Schüler] die Zusammenfassung am Ende verstanden haben“ (S. 108; vgl. auch z. B. Oser et al., 2010; Seidel & Stürmer, 2014; Wolters, 2014). Bei der Formulierung der Items muss bedacht werden, dass Probandinnen und Probanden durch das Item auf das zu beurteilende Merkmal aufmerksam gemacht werden. Dadurch bleibt unklar, ob das Merkmale und die entsprechende Textstelle auch ohne das Item wahrgenommen worden wären. Möglicherweise kann ein Item oder ein Test so zu leicht werden oder es kann zu einer Überschätzung der Kompetenzen in der PW führen (Barth, 2017; Brovelli et al., 2013).

Müller und Gold (2023) zeigen, dass offene und geschlossene Befragungen unterschiedliche Ergebnisse liefern und sie somit nicht unbedingt das gleiche Konstrukt erfassen. Dies gibt einen Hinweis darauf, dass erfasste Fähigkeiten auch vom Frageformat abhängen. Beide Ansätze sind also nicht als äquivalent zu betrachten, so dass zwischen beiden nicht nur hinsichtlich der Praktikabilität abgewogen werden muss, sondern auch hinsichtlich der Frage, was genau mit den jeweiligen Frageformaten erfasst werden kann. Der Vorteil eines qualitativen Zugangs liegt darin, dass er eine Exploration der Zugangsweisen zu Vignetten erlaubt und Beobachtungs- und Beurteilungsprozesse zunächst ohne Einflussnahme sichtbar machen kann. Ein quantitativer Zugang ist nicht nur ökonomisch und mit einer größeren Anzahl von Probandinnen und Probanden durchführbar sowie objektiv auswertbar, sondern es werden zudem konkrete Handlungsanforderungen an die Lesenden gestellt. Beide Vorteile werden in der Pilotierungsphase mit einem *mixed-methods*-Ansatz genutzt.

Bezüglich der jeweiligen Instruktion stellt sich die Frage, wie deutlich diese auf die zu erkennenden Lernbarrieren verweisen sollte. Im offenen Ansatz soll erprobt werden, worauf Studierende spontan Bezug nehmen, wenn sie die Unterrichtssituationen beurteilen. Eine sehr konkrete Instruktion wie z. B. *Beurteilen Sie, wie gut Klassenführung und Lernunterstützung gelingen* würde die Fokussierung auf Unterrichtsmerkmale vorwegnehmen. Impulse zu den qualitativen Befragungen sollen deshalb möglichst offen formuliert werden und weder auf Lernbarrieren noch auf Unterrichtsmerkmale verweisen. Gleiches gilt für die Instruktion zur Beurteilung der Items, welche möglichst neutral formuliert werden soll.

8.7 Präsentation der Vignetten

Befunde zu den Einsatzmodalitäten von Vignetten variieren. Während Zucker (2019) darauf hinweist, dass der Art und Weise der Präsentation von Vignetten in der spezifischen Erhebungssituation eine wichtige Rolle zukommt (z. B. in Papierform oder online), stellen Templer und Lange (2008) sowie Jahn et al. (2011) keine bedeutsamen Effekte unterschiedlicher Erhebungssituationen fest.

Unterschiede ergeben sich aber z. B. hinsichtlich der Einsatzbreite und Durchführungsökonomie. Während Online-Verfahren für große Kohorten an verschiedenen Universitäten leicht zugänglich sind, ist das bei Papier- und Bleistift-Verfahren aufwändiger. Zudem müssen Daten auf Papier zunächst von Hand zur weiteren Verarbeitung eingegeben werden, während Daten aus Online-Umfragen mit wenig Aufwand in unterschiedliche Programme zur Auswertung überführt werden können. Andererseits ist die Bearbeitung eines Tests auf Papier z. B. innerhalb der Seminarzeit verbindlicher und lässt auf höhere Rücklaufzahlen hoffen. Die Befragung ist zudem unabhängig von Technik bzw. technischen Problemen oder dem Besitz digitaler Endgeräte.

Die Durchführungsobjektivität ist bei Online-Verfahren leichter zu bewahren, da Einleitungstexte und Erklärungen standardisiert dargeboten werden und nicht von einer oder mehreren versuchsleitenden Personen abhängig sind (Gold & Holodynski, 2017; Meschede, 2014). Vorgaben zur Bearbeitung lassen sich ebenfalls bei der Online-Bearbeitung stärker kontrollieren. So muss z. B. entschieden werden, ob die zu bearbeitende Textvignette einmalig gelesen werden soll, um die Flüchtigkeit von Unterrichtsmomenten realitätsnäher darzustellen, oder ob sie zugunsten einer genaueren Wahrnehmung wiederholt gelesen werden kann (vgl. Barth, 2017; Brovelli et al., 2013; König & Kramer, 2016; Knievel et al., 2015).

Beim Einsatz von Textvignetten kommt dem Lesen in beiden Präsentationsformaten eine wichtige Rolle zu. Im Rahmen der europäischen Forschungsinitiative *Evolution of Reading in the Age of Digitisation* (E-READ; 2019), die sich mit der Entwicklung des Lesens im Zeitalter der Digitalisierung befasst, wurde untersucht, wie Kinder und junge Erwachsene geschriebenen Text verstehen oder sich daran erinnern, wenn sie gedruckte oder digitale Materialien verwenden. Die wichtigsten Ergebnisse wurden im norwegischen Stavanger als *Stavanger Declaration* zusammengefasst. Im Hinblick auf die Bearbeitung von Textvignetten durch Erwachsene ergeben sich demnach keine Vorteile einer gedruckten Version gegenüber einer digitalen Darbietung. Häufige Arbeit am Computer führt bei kürzeren sowie narrativen Texten zu einer Gewöhnung, am Bildschirm zu lesen. Zudem können Vorteile durch die individuelle Darstellung der Texte auf den persönlichen Endgeräten genutzt werden (Clinton, 2019). Für die Durchführung der Datenerhebungen sind deshalb jeweils die Praktikabilität und die Generierung eines möglichst hohen Rücklaufs sowie die Zusammensetzung der Stichprobe entscheidend, so dass das Verfahren sowohl in Papierform als auch online eingesetzt wird.

8.8 Itemkonstruktion

Bei der Konstruktion von Items ist wie bei der Vignettenkonstruktion die sprachliche Gestaltung zu bedenken. Lehrkräfte werden als Akteurinnen und Akteure, je nachdem wie die Items formuliert sind, unterschiedlich stark in den Mittelpunkt gerückt, z. B. *Die Lehrkraft hat Regeln und Routinen etabliert* oder *Der Ablauf lässt etablierte Regeln und Routinen erkennen*. Im ersten Fall wird eher die Person, im zweiten Fall eher die Situation beurteilt. Zwischen beiden Möglichkeiten soll variiert werden, damit der Fokus der Beurteilungen nicht ausschließlich auf die Lehrperson oder die Situation gelenkt wird.

Auch Länge und Komplexität der Item-Formulierung, z. B. durch doppelte Verneinung, haben Einfluss auf die Schwierigkeit der Items und müssen bedacht werden (Moosbrugger & Kelava, 2007). Die gleichen Überlegungen gelten auch für die offen formulierten Fragen. Bei der

Konstruktion der Fragen und Items werden diese Überlegungen im Entwicklungsprozess berücksichtigt und ggf. für Veränderungen genutzt, um eine gute Verständlichkeit bei mittleren Schwierigkeitsgraden zu erreichen (Tepner & Dollny, 2014).

8.9 Auswertungsverfahren

In einem weiteren Schritt müssen Beurteilungsindikatoren für die erfassten Bearbeitungen der Vignetten festgelegt werden, die eine distale oder proximale Einschätzung von Kompetenzen ermöglichen (Brovelli et al., 2013; Kunter & Klusmann, 2010; Vogelsang, 2014). *Distale Einschätzung* meint das Erfassen von Indikatoren wie beispielsweise Schulnoten, die in größerem Abstand zur eigentlich erfassten Fähigkeit stehen und die Kompetenz nicht direkt erfassen. Sie erfassen Anhaltspunkte, die mit der zu untersuchenden Kompetenz in Zusammenhang stehen. *Proximale Indikatoren* beziehen sich dagegen direkt auf die erfasste Fähigkeit. Das bedeutet, die Kompetenz Lernbarrieren zu erkennen, wird direkt erfasst (Keller et al., 2018; Kunter & Klusmann, 2010).

Zur Auswertung muss die Qualitätsnorm für die Beurteilungen festgelegt werden (Oser et al., 2010). Während bei offenen Antwortformaten Codierverfahren entwickelt werden müssen (Santagata & Angelici, 2010; Santagata & Guarino, 2011; Santagata et al., 2007; Santagata & Yeh, 2013; Sherin & van Es, 2009), werden bei der Konstruktion von Ratingskalen richtige Antworten entweder festgelegt (Star & Strickland, 2008) oder durch Expertenratings ermittelt (Gold & Holodynski, 2017; Gold et al., 2017; Holodynski et al., 2017; Seidel et al., 2010; Stürmer & Seidel, 2015). Die Punktevergabe für die richtige Antwort bzw. für die Antwort, die mit dem Urteil von Expertinnen und Experten übereinstimmt, ermöglicht eine schnelle, objektive Auswertung und Vergleichbarkeit von Urteilsergebnissen (Zucker, 2019). Für die vorliegende Arbeit werden ein Codiermanual für die qualitative Erhebung als auch ein Masterrating für die Auswertung des geschlossenen Formats verwendet.

Die unterschiedlichen Vorteile qualitativer und quantitativer Ansätze werden im ersten Entwicklungsschritt in einem *mixed-methods*-Ansatz genutzt. Die sprachliche Gestaltung der Items wird im Entwicklungsprozess erprobt und ggf. variiert. Zur Auswertung werden proximale Kriterien festgelegt. Für die qualitative Auswertung wird ein Codiermanual erstellt und für das quantitative Format werden richtige Antworten durch die Autorin festgelegt. Eingesetzt wird das Verfahren als Papier-und-Bleistift-Variante und als Online-Variante. Maßgeblich dafür sind die Reichweite und die Generierung eines guten Rücklaufs.

9 Erster Entwicklungsschritt

Zur Entwicklung eines Verfahrens, das die Kompetenz des Erkennens von Lernbarrieren erfasst, stellt sich die Frage, mit welchen Impulsen solche spezifischen Beobachtungs- und Beurteilungsprozesse so in Gang gesetzt werden können, dass ein Erkennen oder Nicht-Erkennen von Lernbarrieren messbar wird.

Zu prüfen sind damit zum einen die Frage nach der Tauglichkeit der entwickelten Vignettenkonstruktion und zum anderen die Fragen nach den unterschiedlichen Frageformaten, die die Beurteilung auslösen sollen. Zur Entwicklung eines textvignettenbasierten Verfahrens soll außerdem erhoben werden, wie die Arbeit mit Vignetten von den Probandinnen und Probanden akzeptiert wird. Die Akzeptanz wird durch Fragen zur Verständlichkeit und Interessantheit sowie zum subjektiven Schwierigkeitsgrad der Vignettenbeurteilungen operationalisiert.

9.1 Fragestellungen und Hypothesen des ersten Entwicklungsschritts

Es ergeben sich für den ersten Entwicklungsschritt die folgenden Fragen und Hypothesen:

Akzeptanz

Frage 1: Methodische Ansätze wie die Arbeit mit Vignetten können nur erfolgreich umgesetzt werden, wenn Studierende diese akzeptieren. Bei der Konstruktion wurden deshalb Aspekte wie die Dauer und Schwierigkeit der Bearbeitung, die Praxisrelevanz und die Formulierung der Vignetten, Items und Instruktionen bedacht. Führt die Umsetzung dazu, dass Studierende diese gut akzeptieren?

Hypothese 1: Studierende akzeptieren die Arbeit mit Vignetten. Die unterschiedlichen Fragestellungen (offene Frage oder zu beurteilende Items) spielen dafür keine Rolle.

Variation der Impulse

Frage 2: Die Modellierung der Kompetenz *Erkennen von Lernbarrieren* betont die Bedeutung der Fokussierung auf Unterrichtsmerkmale. Damit stellt sich die Frage, wie vorstrukturiert und gezielt Impulse durch die Vignette und durch die Fragen bzw. Items sein müssen, damit Studierende diese Fokussierung vornehmen. Ziehen Studierende den Fokus auf Unterrichtsmerkmale für ihre Beobachtung spontan in Betracht? Oder tun sie dies erst, wenn Handlungen zur Klassenführung und Lernunterstützung konkret in der Vignette beschrieben werden? Wie hängt das außerdem vom Frageformat ab?

Hypothese 2: Es hängt von der Art der Vignette und der Frage bzw. den Items ab, ob und wie stark Studierende Unterrichtsbedingungen fokussieren und Barrieren in Unterrichtsbedingungen erkennen. Je strukturierter und gezielter Barrieren dargestellt und abgefragt werden, umso deutlicher beziehen sich Urteile darauf.

Mittlere Schwierigkeitsgrade, Trennschärfen und Reliabilität

Frage 3: Bei der Konstruktion der Vignette wurde mit den Unterrichtsmerkmalen ein differenziertes theoretisches Konstrukt hinterlegt. Durch die Items werden die operationalisierten Merkmale gezielt abgefragt. Weisen die konstruierten Items akzeptable Werte hinsichtlich der Schwierigkeit, Trennschärfen und Reliabilität auf?

Hypothese 3: Die konstruierten Items sind angemessen schwierig, trennscharf und reliabel.

Validität

Frage 4: Die Kompetenz *Erkennen von Lernbarrieren* erfordert, dass neben der notwendigen Fokussierung auf Unterrichtsmerkmale die Interpretation wahrgenommener Interaktionen hinsichtlich potenzieller Barrieren in der Klassenführung oder Lernunterstützung vorgenommen wird. Lässt sich das Erkennen von Barrieren in der Klassenführung vom Erkennen von Barrieren in der Lernunterstützung unterscheiden?

Hypothese 4: Das Erkennen von Barrieren hängt davon ab, ob sie sich auf die Klassenführung oder die Lernunterstützung beziehen. Beides ist zu unterscheiden.

Frage 5: Die konstruierten Vignetten und Items sollen unterschiedliche Fähigkeitsausprägungen im Erkennen von Lernbarrieren erfassen. Ist die Beurteilungsaufgabe im Sinne der Kriteriumsvalidität sensibel genug, um Unterschiede zwischen Gruppen unterschiedlicher Expertise zu erfassen?

Hypothese 5: Es zeigen sich Unterschiede zwischen den Gruppen zugunsten der Sonderpädagogikstudierenden und Studierenden der höheren Fachsemester.

9.2 Konstruktion der Aufgaben

Das Erkennen von Lernbarrieren im oben definierten Sinn erfordert eine spezifische Beobachtung und Beurteilung von Lehr-Lern-Situationen. Es müssen Handlungen fokussiert werden, die relevant für den Unterricht sind. Damit stellt sich die Frage, ob dies durch die Art der Vignette und des Beurteilungsimpulses beeinflusst wird:

Zum einen soll geprüft werden, ob Studierende *spontan* Unterrichtsmerkmale fokussieren, um potenzielle Barrieren zu beobachten. Diese spontane Fokussierung soll durch eine offene

Frage erfasst werden, weil dadurch kein Fokus auf Unterrichtsmerkmale vorgegeben wird. Welche Rolle dafür die Vorstrukturierung der Vignette spielt, wird erprobt, indem zuerst die Kurzvignette in Kombination mit einer offenen Frage eingesetzt wird und dann die Vignette *Philipp* mit einer offenen Frage.

In der Kurzvignette werden keine konkreten Handlungen beschrieben, die direkt auf Probleme in der KF und LU schließen lassen. Der Impuls, dies zu hinterfragen, müsste bei den Studierenden also ohne äußeren Hinweis entstehen, i. S. von: *Wenn Lehr-Lernsituationen potenziell problematisch sein könnten, schaue ich mir gezielt die KF und LU an.* Dazu wurde zur oben vorgestellten Kurzvignette folgender Schreibimpuls gegeben: *Was würden Sie beobachten, um Philipps Situation genauer beschreiben zu können?* Damit können der Beobachtungsfokus und dessen Operationalisierung frei gewählt werden. Diese Aufgabe wurde von allen Studierenden bearbeitet.

Um den Einfluss des Impulses durch die Vignette zu prüfen, wurde die Vignette *Philipp* ebenfalls mit einer offenen Frage eingesetzt. Sie unterscheidet sich von der Kurzvignette darin, dass die in der Vignette beschriebenen Handlungen direkt auf Merkmale der KF und LU verweisen und somit einen Impuls für deren Beobachtung liefern. Die Studierenden erhielten die Instruktion *Stellen Sie sich nun die folgende Situation vor. Notieren Sie bitte alles, was Ihnen auffällt!*

Der Einfluss des Beurteilungsimpulses wurde geprüft, indem die Vignette *Philipp* außerdem mit einem quantitativen Ansatz erprobt wurde. Die Instruktion lautete: *Stellen Sie sich nun die folgende Situation vor. Beurteilen Sie bitte anschließend die Aussagen auf der nächsten Seite! Bitte kreuzen Sie an, lassen Sie bitte keine Antwort aus! (Sie dürfen zum Text zurückblättern)* Dazu waren Aussagen zu den potenziellen Lernbarrieren in der Vignette auf einer vierstufigen Likert-Skala zu beurteilen. Die Items wurden basierend auf den Vorüberlegungen zunächst sprachlich einfach, also im Aktiv ohne doppelte Verneinungen formuliert. Die Formulierungen bezogen sich auf die Merkmale der KF und LU, ohne diese explizit zu nennen. D. h., es wurde nicht formuliert *Die Lehrerin ist allgegenwärtig*, sondern *Die Lehrerin hat die ganze Klasse im Blick*. Einzig das Wort *Regeln* verweist auf das Merkmal *Regeln und Routinen*. Auf diese Weise wird der Fokus bereits auf relevante Aspekte gelenkt und es ist im Sinne der Kompetenz des Erkennens eine Interpretation der Handlungen hinsichtlich der Auswirkung auf die Lehr-Lernsituation notwendig, um die Items als mehr oder weniger zutreffend einordnen zu können. Die offene und geschlossene Befragung zur Vignette *Philipp* wurde jeweils von der Hälfte der Studierenden bearbeitet. Im Anschluss an die Vignettenbearbeitungen beantworteten außerdem alle Studierenden drei Fragen zur Akzeptanz (siehe Abbildung 4). Folgende Items wurden formuliert:

<i>Raum</i>	Die Arbeitssituation ist für Philipp angemessen.
<i>Inhaltliche Klarheit</i>	Die Aufgabenstellung ist für Philipp angemessen.
<i>Veranschaulichung</i>	Die Lehrerin stellt sicher, dass Philipp die Aufgabe versteht.
<i>Zielklarheit</i>	Die Lehrerin verdeutlicht Philipp das Vorgehen bei der Bearbeitung.
<i>Allgegenwärtigkeit</i>	Die Lehrerin hat die Gesamtsituation im Griff.
<i>Übergänge</i>	Die Lehrerin unterstützt Philipp bei der selbständigen Bearbeitung.
<i>Regeln und Routinen</i>	Die Gruppenarbeit in der Klasse läuft nach festen Regeln ab.
<i>Feedback</i>	Die Lehrerin gibt Philipp angemessene Rückmeldungen.
<i>Verständlichkeit</i>	Die Aufgabenstellung war verständlich.
<i>Subjektive Schwierigkeit</i>	Die Aufgabe war einfach.
<i>Interessanz</i>	Die Aufgabe war interessant.

Abbildung 4

Aufgaben und Aufbau des ersten Entwicklungsschritts

Aufgabe 1 (alle Studierenden)

Kurzvignette: Stellen Sie sich vor, Sie sind ein/e Lehrer/in in einer inklusiven Klasse, die aus 20 Schülerinnen und Schülern ohne Behinderungen und 4 Schülerinnen und Schülern mit Behinderungen besteht.... (Przibilla et al., 2016)

Offene Frage: Was würden Sie beobachten, um Philipps Situation genauer beschreiben zu können?

**Aufgabe 2 (zufällige Zuteilung)**

Vignette Philipp: Philipp sitzt während des Matheunterrichts in seiner Klasse an seinem Gruppentisch....

Offene Frage: Stellen Sie sich nun die folgende Situation vor. Notieren Sie bitte alles, was Ihnen auffällt!

oder

Vignette Philipp: Philipp sitzt während des Matheunterrichts in seiner Klasse an seinem Gruppentisch....

Geschlossenes Format: Stellen Sie sich nun die folgende Situation vor. Beurteilen Sie bitte anschließend die Aussagen auf der nächsten Seite! Bitte kreuzen Sie an, lassen Sie bitte keine Antwort aus! (Sie dürfen zum Text zurückblättern)

(Beurteilung von acht Items auf vierstufiger Likertskala)

**Akzeptanzbefragung (alle Studierenden)**

Drei Fragen zu Verständlichkeit, Schwierigkeit und Interessantheit

9.3 Stichprobe

Die Wahrnehmung von Aspekten der KF und LU wurde bislang in Gruppen untersucht, die sich hinsichtlich ihres Professionsgrades unterscheiden. Bewährt haben sich Stichproben von Studierenden unterschiedlicher Fächer und Fachsemester, von Referendarinnen und Referendaren, Lehrpersonen mit unterschiedlich viel Berufserfahrung sowie Schulinspektoren und -inspektorinnen (Baer et al., 2007; Barth, 2017; Gold & Holodynski, 2017; König & Kramer, 2016; Meschede et al., 2015; Seidel & Prenzel, 2007; Sherin & van Es, 2009; Star & Strickland, 2008; Stürmer & Seidel, 2017; Wolters, 2014).

Da sich die vorliegende Untersuchung auf Lehramtsstudierende konzentriert, besteht die Stichprobe aus Studierenden aus unterschiedlichen Fachsemestern und Studiengängen. Der Vorteil liegt dabei darin, dass der Ausbildungsstand der Studierenden anhand des Lehrangebots relativ gut eingeschätzt werden kann und innerhalb der Gruppen eher homogen ist.

Im ersten Entwicklungsschritt soll geprüft werden, ob sich unterschiedliche Fähigkeitsausprägungen zur Erkennung von Lernbarrieren abbilden lassen. Als Stichprobe werden deshalb Gruppen von Studierenden ausgewählt, für die unterschiedliche Expertise angenommen wird. Das Wahrnehmen und Interpretieren unterrichtsrelevanter Ereignisse erfordert als Disposition spezifisches Wissen sowie grundlegende Beobachtungskompetenzen. Da beides im Studienverlauf vermittelt wird, ist anzunehmen, dass auch die darauf basierende PW und damit das Erkennen von Lernbarrieren in höheren Semestern besser gelingt als zu Studienbeginn, so dass Studierende unterschiedlicher Fachsemester befragt werden.

In der Sonderpädagogik werden Studierende zudem für problematische Lernsituationen besonders sensibilisiert und sie lernen und üben das Beobachten als diagnostische Methode. Wegen dieser stärkeren Fokussierung auf Lernbarrieren wird angenommen, dass Studierende der Sonderpädagogik Barrieren in Vignetten besser erkennen als Studierende anderer Lehrämter, so dass auch Gruppen unterschiedlicher Lehramtsstudiengänge befragt werden (siehe Tabelle 4).

Die Studierenden wurden danach ausgewählt, welche Lehrenden kurzfristig in der Lage waren, den Vignettentest in ihrer Lehrveranstaltung durchzuführen, unabhängig vom jeweiligen Seminar- oder Vorlesungsthema. Innerhalb der Veranstaltungen nahmen alle Studierenden an der Befragung teil und Lehrende hatten keinen Einfluss auf die Durchführung, so dass nicht von starken Stichprobeneffekten auszugehen ist. Die Studiengangs- und Fachsemestergruppen werden als unabhängige Stichproben behandelt.

Tabelle 4

Häufigkeiten für die Stichprobe des ersten Entwicklungsschritts

Studiengang	Fachsemester			
	1. Bachelor	5. Bachelor	3. Master	Gesamt
Allgemeinpädagogik	38	22	21	81
Sonderpädagogik	54	22	58	134
Gesamt	92	44	79	215

9.4 Design und Durchführung des ersten Entwicklungsschritts

Die Aufgaben wurden von 215 Studierenden als Papier-und-Bleistift-Test innerhalb von Vorlesungen und Seminarveranstaltungen bearbeitet, um einen möglichst hohen Rücklauf zu generieren. Als Ankündigung erhielten die Studierenden die Information, dass es sich um die Entwicklung einer Aufgabe zur Einschätzung von Lernsituationen handelt. Das Deckblatt der Vignette gab zusätzlich mit folgendem Text Auskunft darüber:

In den erziehungswissenschaftlichen Studiengängen der Universität Hamburg sollen Studierende auf das Unterrichten in inklusiven Klassen vorbereitet werden. Die Lehre soll dahingehend optimiert werden. Um den Erfolg von Seminaren beurteilen zu können, ist es nötig, entsprechende Methoden zu entwickeln. Bei der vorliegenden Aufgabe handelt es sich um einen ersten Entwurf, der nun getestet werden soll. Mit der Bearbeitung der Aufgabe leisten Sie also einen wichtigen Beitrag!

Die Auswertung der Aufgabe erfolgt anonym und wird nur zu Forschungszwecken verwendet.

Die Ankündigung der Aufgabe erfolgte immer durch die gleiche Person, so dass von gleichen Bedingungen und einer vergleichbaren Durchführung ausgegangen werden kann. Auf dem Deckblatt gaben die Studierenden Studiengang und Fachsemester an. Rückschlüsse auf die Person waren nicht möglich. Die Bearbeitungszeit betrug jeweils 15 Minuten. Innerhalb dieser Zeit beantworteten alle Studierenden die erste Teilaufgabe zur Kurzvignette mit offener Frage. Die zweite Teilaufgabe zur Vignette *Philipp* beantworteten die Studierenden in zufälliger Zuteilung, entweder in der offenen oder der geschlossenen Variante. Die Studierenden hatten dabei die Gelegenheit, zum Text zurückzublättern. Die Bearbeitungszeit war dabei für alle Studierenden gleich, wobei diejenigen, die die offene Aufgabe bearbeiteten, Antworten formulieren mussten und einen höheren Bearbeitungsaufwand hatten. Dieser Umstand wurde bewusst in Kauf genommen, weil die Studierenden durch einen gewissen Zeitdruck dazu gebracht werden sollten,

keine zu langen Ausführungen zu schreiben, sondern nur das zu notieren, was ihnen zuerst und am meisten aufgefallen ist (vgl. Barth, 2017).

9.4.1 Auswertung der Akzeptanzbefragung

Für die Antworten der Akzeptanzbefragung werden jeweils Mittelwerte für die Gruppen berechnet. Es wird dabei in Studierende, die als zweite Aufgabe die offene bzw. geschlossene Variante bearbeitet hatten, unterschieden.

9.4.2 Auswertung der ersten Teilaufgabe

Für die Auswertung der Antworten der ersten Teilaufgabe wurden in Anlehnung an die qualitative Inhaltsanalyse nach Mayring (2008) inhaltstragende Textteile markiert und paraphrasiert. Es wurden dann Kategorien entworfen und die Aussagen zugeordnet. Dieses Verfahren wurde durch drei geschulte Raterinnen durchgeführt. Die Zuordnungen wurden anschließend überprüft und in einem Konsensverfahren zusammengeführt. Die Antworten der Studierenden wurden anhand des so entstandenen Kategoriensystems erneut ausgewertet, indem notiert wurde, ob die Kategorien jeweils genannt wurden (Wert 1) oder nicht (Wert 0). Für diese Auswertung wurde die Inter-Rater-Reliabilität mithilfe des Fleiss-Kappa bestimmt. Für die weitere Auswertung wurden nur die Fälle aufgenommen, bei denen alle drei Raterinnen in ihrem Urteil übereinstimmten (Fleiss-Kappa 1.00). Für die genannten Kategorien wurden Häufigkeiten bestimmt. Mit der Annahme, dass sich unterschiedliche Fokusse vor allem aus verschiedenen Studiengängen ergeben, wurde dabei zunächst nur in Studierende der Regelschullehrämter und Studierende der Sonderpädagogik unterschieden. Alle Berechnungen wurden mit dem Statistikprogramm SPSS (Version 28.0.1.1) durchgeführt.

9.4.3 Auswertung der offenen zweiten Teilaufgabe

Bei der zweiten Teilaufgabe wurde der Impuls durch die Vignette *Philipp* gegeben, die acht zu erkennende Lernbarrieren enthielt. Die Analysekatoren entsprachen diesen acht Barrieren und wurden durch Ankerbeispiele erläutert und in einem Codiermanual zusammengestellt (siehe Anhang B1). Die Antworten der Studierenden wurden durch die gleichen drei geschulten Raterinnen ausgewertet. In Anlehnung an Mayring (2008) wurden inhaltstragende Teile markiert, paraphrasiert und den vorgegebenen Kategorien zugeordnet. Dabei konnten unterschiedliche Werte vergeben werden:

<i>Kategorie wurde nicht genannt</i>	0
<i>Kategorie wurde beschreibend oder nacherzählend genannt</i>	1
<i>Kategorie wurde interpretierend oder beurteilend genannt</i>	2
<i>Es wurde ein Verbesserungsvorschlag gemacht</i>	3
<i>Es wurde auf eine Textstelle mit potenzieller Lernbarriere Bezug genommen, aber diese wurde nicht zutreffend interpretiert</i>	4

Diese Werte entsprechen zunächst nominalen Werten. Für die Prüfung der Inter-Rater-Reliabilität wurde erneut das Fleiss-Kappa bestimmt. Da in diesem Beurteilungsverfahren zwei Schritte übereinstimmen mussten, nämlich die Zuordnung zu einer Kategorie und die Zuordnung eines Werts für die Kategorie, wurde vorher festgelegt, dass im Falle einer zu geringen Anzahl völlig übereinstimmend beurteilter Fälle das Rating der Autorin als Masterrating der Auswertungen zugrunde gelegt wird. Dies geschah unter der Voraussetzung einer akzeptablen Forscher-Kodierer-Reliabilität (Lersch, 2006; Rössler, 2017) und mit dem Ziel, durch ein exploratives Vorgehen genügend Hinweise zur Vignette und zum Frageformat sammeln zu können.

9.4.4 Auswertung der geschlossenen zweiten Teilaufgabe

Bei den Items der geschlossenen Aufgabe handelt es sich jeweils um Sätze zu den in der Vignette enthaltenen Barrieren. Diese wurden von den Studierenden auf einer vierstufigen Likert-Skala (*trifft zu*, *trifft eher zu*, *trifft weniger zu*, *trifft nicht zu*) eingeschätzt. Für die Antworten wurden die Werte 1 bis 4 vergeben. Die Formulierung der Items enthält in allen Fällen eine positive Aussage wie z. B. *Die Lehrerin hat die Gesamtsituation im Griff*, so dass jeweils die Antworten *trifft weniger zu* (Wert 3) und *trifft nicht zu* (Wert 4) richtig sind. Für die Items werden Mittelwerte bestimmt.

Es wurde eine Verteilungsprüfung durchgeführt und die Daten werden anhand einer deskriptiven Statistik dargestellt. Anschließend wurden Trennschärfen betrachtet und eine Reliabilitätsanalyse durchgeführt. Die Konstruktvalidität wurde mittels explorativer Faktorenanalyse getestet. Zur Ermittlung der Kriteriumsvalidität wurden die Stichprobengruppen (verschiedene Fachsemester und Studiengänge) anhand eines Testscores verglichen. Der Testscore wurde gemäß den Ergebnissen der Faktorenanalyse anhand der gewichteten Mittelwerte (Itemrohwertermal Faktorladung) pro Faktor berechnet. Da es sich bei den unterschiedlichen Teilgruppen um kleine Stichproben handelt, für die nicht unbedingt Varianzhomogenität angenommen werden kann, wurden dafür non-parametrische Verfahren gewählt.

Während die Itemselektion dazu dient, nicht funktionierende Items zu eliminieren (Bühner, 2006), sollen in dieser Studie Items nicht ausgeschlossen werden. Für die Konstruktion der Vignette wurden die operationalisierten Merkmale der KF und LU genutzt. Die vorhandenen Items sollen sich auf diese Merkmale beziehen. Dies ist zu prüfen und ggf. zu verbessern und es sollen möglichst Items zu allen Merkmalen erhalten bleiben. Im nächsten Entwicklungsschritt sollen deshalb inhaltliche oder methodische *Veränderungen* der Items bzw. der dazugehörigen Textstellen in der Vignette Hinweise darauf liefern, wie sich die Beurteilungen der Studierenden dadurch beeinflussen lassen.

9.5 Ergebnisse und Interpretation des qualitativen Ansatzes

In den Ergebnissen des ersten Entwicklungsschritts wird die Fokussierung auf Unterrichtsmerkmale in Zusammenhang mit unterschiedlichen Vignetten und Frageformaten dargestellt und mit dem Blick auf Konsequenzen für das weitere Vorgehen verglichen. Die Ergebnisse der Akzeptanzbefragung werden vorangestellt, weil eine hohe Akzeptanz des Verfahrens Voraussetzung für die weiteren Entwicklungsschritte ist.

9.5.1 Ergebnisse der Akzeptanzbefragung

Frage 1: Führt die Umsetzung des Vignetten- und Aufgabenentwurfs dazu, dass Studierende diese gut akzeptieren?

Hypothese 1: Studierende akzeptieren die Arbeit mit Vignetten. Die unterschiedlichen Fragestellungen (offene Frage oder zu beurteilende Items) spielen dafür keine Rolle.

Für die Antworten wurden Werte von 1 = *trifft nicht zu* bis 4 = *trifft voll und ganz zu* vergeben. Es wurden jeweils Mittelwerte berechnet, wobei das arithmetische Mittel ($M = 2.5$) jeweils einer mittleren Verständlichkeit, einer mittleren Schwierigkeit und mittlerer Interessantheit entspricht.

Dabei ergibt sich für die Verständlichkeit ein Mittelwert von $M = 3.13$ ($SD = 0.81$) für die offene Version und von $M = 3.09$ ($SD = 0.91$) für die geschlossene. Der Mittelwert für den Schwierigkeitsgrad liegt bei der offenen Version bei $M = 2.61$ ($SD = 0.67$) und für die geschlossene bei $M = 2.69$ ($SD = 0.70$). Die Interessantheit der offenen Aufgabe wird im Mittel mit $M = 3.12$ ($SD = 0.88$) beurteilt, die geschlossene mit $M = 3.23$ ($SD = 0.82$) (siehe Tabelle 5). Die Ergebnisse für die unterschiedlichen Aufgabenformate unterscheiden sich geprüft mit dem Mann-Whitney-U-Test nicht signifikant (siehe Anhang B Tabelle B2).

Tabelle 5

Mittelwerte zur Einschätzung der Verständlichkeit, Schwierigkeit und Interessantheit

Aufgabenformat		Verständlichkeit	Schwierigkeitsgrad	Interessantheit
offene Frage	<i>M</i>	3.13	2.61	3.12
	<i>SD</i>	0.81	0.67	0.88
	<i>N</i>	104	106	106
geschlossene Fragen	<i>M</i>	3.09	2.69	3.23
	<i>SD</i>	0.90	0.70	0.82
	<i>N</i>	109	109	109

9.5.2 Interpretation der Akzeptanzbefragung

Die Ergebnisse stützen die erste Hypothese. Die nötige Akzeptanz des erprobten Verfahrens zeigt sich in den hohen Mittelwerten für die Verständlichkeit und Interessantheit der Aufgabe und einem subjektiven Schwierigkeitsgrad, der als mittelschwer eingeschätzt wird. Bedeutsame Unterschiede zwischen den beiden Frageformaten ergeben sich nicht, d. h., die Einschätzungen sind unabhängig vom Aufgabenformat.

9.5.3 Ergebnisse der ersten Teilaufgabe

Variation der Impulse

Frage 2: Ziehen Studierende den Fokus auf Unterrichtsmerkmale für ihre Beobachtung spontan in Betracht? Oder tun sie dies erst, wenn Handlungen zur Klassenführung und Lernunterstützung konkret in der Vignette beschrieben werden? Wie hängt das außerdem vom Frageformat ab?

Hypothese 2: Es hängt von der Art der Vignette und der Frage bzw. den Items ab, ob und wie stark Studierende Unterrichtsbedingungen fokussieren und Barrieren in Unterrichtsbedingungen erkennen. Je strukturierter und gezielter Barrieren dargestellt und abgefragt werden, umso deutlicher beziehen sich Urteile darauf.

Die Inhaltsanalyse der Antworten ergab 15 Oberkategorien mit insgesamt 30 Unterkategorien, die von den Studierenden genannt wurden (siehe Tabelle 6). Tabelle 7 zeigt für die übereinstimmend bewerteten Fälle die Häufigkeiten für diese Kategorien.

Tabelle 6*Ergebnis der Inhaltsanalyse der ersten Teilaufgabe: Von Studierenden genannte Kategorien*

Oberkategorie	Unterkategorie
Physische und kognitive Voraussetzungen	<p>Vorleseverständnis, Sprachverständnis</p> <p>Sehen (dicht an den Text gehen, Körperhaltung, Augen kneifen, Augenbewegungen, Mimik, Kopf bewegen, Text bewegen)</p> <p>Hören</p> <p>Sprechen und Sprache (Redeflüssigkeit, Wortschatz, SSES, Kommunikation, Erzählfähigkeit, Logopädie)</p> <p>Andere physische und kognitive Voraussetzungen (Arbeitsgedächtnis, Wahrnehmung, Konzentration, Aufmerksamkeit, Ablenkbarkeit, Händigkeit, Motorik, andere medizinische Aspekte)</p> <p>Vorläuferfähigkeiten</p>
Lesefähigkeiten	<p>Leseverständnis</p> <p>Phonem-Graphem-Korrespondenz („Alphabet beherrschen“)</p> <p>Silbenlesen (Synthese, Automatisierung)</p> <p>Lesestrategien (buchstaben- oder silbenweise, ganze Worte erkennen, laut oder halblaut lesen, Bekanntes zuerst, mit Finger mitlesen, selbstgewählte Herangehensweisen oder Hilfsmethoden)</p> <p>Allgemeine Problemanalyse Lesen (auf Lesefähigkeiten bezogen, nicht näher bezeichnet oder sehr allgemein, z. B. Lernstand, Fortschritte, Probleme mit bestimmten Buchstaben, Silben, Worten, Wortlängen, Sätzen, Textsorten, Leseleistung bei bestimmten Textarten, Pseudowörter, diagnostische Tests)</p> <p>Lesegeschwindigkeit und -flüssigkeit (auch: Einfluss der Zeit beim Lesen)</p>
Schriftsprachliche Voraussetzungen	Phase des Schriftspracherwerbs, Fähigkeiten oder Stand beim Schreiben (undifferenziert)
Interessen und Lesemotivation	Leseinhalte und Textsorten (Interessen und Hobbys, Interesse am Text oder am Lesen allgemein, Vorwissen, Sinnhaftigkeit des Lesens für Philipp, Einstellung zum Lesen)
Andere sprachliche Voraussetzungen	Mehrsprachigkeit
Häusliches Umfeld	Zugang zu Literatur, Lesebiographie, liest er überhaupt bzw. was liest er zu Hause, lebensweltliche Bedeutung des Lesens, Einbezug Eltern, allgemein Familiensituation
Lesebedingungen in der Schule	<p>Lesesituation sozial (wann, mit wem, in welcher Situation gelingt Lesen gut oder schlecht)</p> <p>Lesesituation Raum (Licht, Lärm, Sitzordnung, wo kann er gut lesen)</p> <p>Leseförderung, Nachteilsausgleiche, besondere Unterstützung und Aufgabendifferenzierung (Hilfs- oder Förderangebote und deren Akzeptanz und Nutzen)</p>
Soziale Situation, Interaktion und Integration in der Klasse	Integration, Interaktionen, soziale Beziehungen, Umgang mit Schwächen seitens der Mitschülerinnen und Mitschüler, Hilfsangebote durch Mitschülerinnen und Mitschüler, Zusammenarbeit mit anderen

Verhalten, Einstellungen oder beobachtbare Reaktionen	Verhalten allgemein (extro- oder introvertiert, allgemeiner Lernwille, Zuhören, störendes Verhalten, Lernverhalten bei anderen Aufgaben, sozial-emotionale Entwicklung) Verhalten bei Leseaufgaben (Hilfe suchen, Hilfe akzeptieren, Eigeninitiative und Beharrlichkeit, Frustrationstoleranz, Mitarbeit beim Lesen, Vermeidungsstrategien beim Lesen, Ängstlichkeit, schamhafte Reaktion beim Lesen)
Arbeitsorganisation	Arbeitsplatz, Vorhandensein Material, Arbeit allgemein von sich aus beginnen
Subjektive Sicht und affektiv-motivationale Aspekte	Selbsteinschätzung, schulisches Selbstkonzept, Affekt, Motivation
Aufgaben, Material	Aufgaben und Texte (Form, Darbietung, Textformat, Visualisierung, Schriftart, Hilfsmittel wie Zeilenhilfe)
Schulleistung	Vergleich zur Gruppe (im Fach Deutsch oder im Lesen) Leistungen in anderen Fächern (besondere Stärken/Kompetenzen/erfolgreiche Momente, Verständnis für Zahlen und Symbole, Verständnis allgemein)
Lehrkraft	Verhalten der Lehrkraft gegenüber Philipp (Feedback, Kommunikation) Selbstreflexion der Lehrkraft (Erwartungen, Vorurteile, Rollenverständnis, Beziehungen, Unterrichtsgestaltung, eigene Kompetenzen)
Grundbedürfnisse	Sicherung von Philipps Grundbedürfnissen (Toilettengang, Hunger, Durst)

Tabelle 7

Häufigkeiten für die Nennungen der Kategorien je Studienganggruppe

Kategorie	Gruppe 1			Gruppe 2		
	<i>n</i>	Nennungen absolut	Nennungen in %	<i>n</i>	Nennungen absolut	Nennungen in %
Vorleseverständnis, Sprachverständnis	75	6	8.00	131	5	3.82
Sehen	82	9	10.98	136	48	35.29
Hören	84	0	0.00	142	7	4.93
Sprechen und Sprache	81	5	6.17	137	28	20.44
Andere physische und kognitive Voraussetzungen	72	7	9.72	122	31	25.41
Vorläuferfähigkeiten	77	1	1.30	137	1	0.73
Leseverständnis	67	1	1.49	122	1	0.82
Phonem-Graphem-Korrespondenz	74	2	2.70	124	20	16.13

Silbenlesen, Synthese	79	0	0.00	127	11	8.66
Lesestrategien	73	7	9.59	114	15	13.16
Allgemeine Problem- analyse Lesen	61	32	52.46	106	66	62.26
Lesegeschwindigkeit und Leseflüssigkeit	78	2	2.56	137	1	0.73
Phase des Schrift- spracherwerbs	79	3	3.80	139	14	10.07
Interessen und Lese- motivation	75	9	12.00	125	25	20.00
Mehrsprachigkeit	84	4	4.76	142	4	2.82
Häusliches Umfeld	81	6	7.41	130	20	15.38
Lesesituation, sozial	72	11	15.28	118	19	16.10
Lesesituation, räum- lich	77	3	3.90	134	14	10.45
Förderung und Nach- teilsausgleiche	78	1	1.28	121	2	1.65
Soziale Situation und Integration	76	16	21.05	131	27	20.61
Verhalten allgemein	73	5	6.85	125	6	4.80
Verhalten bei Leseauf- gaben	66	16	24.24	120	12	10.00
Arbeitsorganisation	82	1	1.22	143	0	0.00
Affektiv-motivationale Aspekte	71	3	4.23	127	5	3.94
Aufgaben und Mate- rial	76	3	3.95	128	23	17.97
Schulleistung Deutsch im Vergleich zur Klasse	80	5	6.25	137	0	0.00
Schulleistung in ande- ren Fächern	78	8	10.26	120	10	8.33
Verhalten der Lehr- kraft	82	0	0.00	139	2	1.44
Selbstreflexion der Lehrkraft	83	1	1.20	141	1	0.71
Sicherung von Grund- bedürfnissen	84	1	1.19	143	0	0.00

Anmerkung. n = Anzahl der von drei Raterinnen übereinstimmend beurteilten Fälle.

Gruppe 1 = Studierende der Allgemeinpädagogik, Gruppe 2 = Studierende der Sonderpädagogik.

Unter den Kategorien finden sich verschiedene Faktoren, die Lehr-Lernsituationen beeinflussen können: Unterrichtsbedingungen (räumliche, prozessuale und inhaltliche Gestaltung von Unterricht), individuelle Bedingungen des Kindes (physische, kognitive, motivationale und affektive), das Handeln der Lehrkraft sowie die Klassenatmosphäre (soziale Situation und Integration). Die genannten Kategorien wurden diesen Einflussfaktoren zugeordnet (siehe Anhang B Tabelle B3) und zusammengefasst. Die Verteilung zeigt, dass in beiden Gruppen etwa 79 % der Studierenden die individuellen Voraussetzungen des Kindes fokussieren, aber nur in geringem Maße Unterrichtsbedingungen und die Lehrkraft so gut wie gar nicht (siehe Tabelle 8).

Tabelle 8

Häufigkeiten für die Nennungen der zusammengefassten Kategorien nach Einflussfaktoren je Gruppe

Zusammengefasste Kategorien	Gruppe 1		Gruppe 2	
	<i>N</i>	%	<i>N</i>	%
Individuelle Voraussetzungen	133	79.17	330	78.95
Unterrichtsbedingungen	18	10.71	58	13.88
Lehrkraft	1	0.60	3	0.72
Klassenatmosphäre	16	9.52	27	6.46
Gesamt	168	100.00	418	100.00

Anmerkung. *n* = Anzahl der von drei Raterinnen übereinstimmend beurteilten Fälle. Gruppe 1 = Studierende der Allgemeinpädagogik, Gruppe 2 = Studierende der Sonderpädagogik.

9.5.4 Interpretation der Ergebnisse der ersten Teilaufgabe

Die erste explorative Analyse zeigt, dass die eingesetzte Kurzvignette und eine offene Abfrage dazu führen, dass Studierende KF und LU oder Barrieren der KF und LU kaum in den Blick nehmen. Von den Kategorien, die sich aus der Analyse der Studierenden-Antworten ergeben, lassen sich nur wenige der Gestaltung der Lehr-Lernsituation zuordnen. Viele dagegen beziehen sich auf andere Voraussetzungen, vor allem die Lesefähigkeit des Kindes wird fokussiert. Studierende der Sonderpädagogik fokussieren zudem stärker individuelle physische und kognitive Voraussetzungen, was zum Beispiel in den Kategorien Sehen, Sprechen und Sprache oder andere physische und kognitive Voraussetzungen auffällt. Aber auch die Lesebedingungen wie die räumliche Situation oder Aufgaben und Material beim Lesen werden von den Studierenden der Sonderpädagogik etwas stärker berücksichtigt. Studierende der Regelschullehrämter nehmen dagegen im Vergleich etwas mehr das Verhalten des Kindes in den Blick.

Die Zuordnungen der Kategorien zu Gelingensbedingungen zeigen, dass Studierende bei der Beurteilung von Lehr-Lernprozessen die Umsetzung von Unterrichtsmerkmalen spontan quasi gar nicht als relevant betrachten. Vielmehr werden hauptsächlich individuelle Voraussetzungen fokussiert. Dies deckt sich mit bisherigen Befunden (vgl. Kopmann, 2022; Thiel, 2021) und stützt Hypothese 2.

9.5.5 Ergebnisse der zweiten offenen Teilaufgabe

Für die Auswertung der zweiten Teilaufgabe (Vignette *Philipp* mit offener Frage) wurde zunächst erneut das Fleiss-Kappa für die Inter-Rater-Reliabilität für die einzelnen Kategorien bestimmt. Das Fleiss-Kappa für die einzelnen Kategorien lag dabei zwischen .266 und .569 (siehe Anhang B Tabelle B4). Die Werte liegen zwar nach Landis und Koch (1977) im akzeptablen Bereich (außer für die Kategorien *Veranschaulichung* und *Übergänge*), wurden aber dennoch stichprobenartig auf Konsens geprüft. Dabei wurde deutlich, dass eine Raterin ein selbst erstelltes Kurzmanual zur besseren Übersicht benutzt hatte, das zu Unterschieden in der Zuordnung führte. Nach Ausschluss dieser Raterin verbesserten sich die Werte und wiesen eine ausreichende Übereinstimmung zwischen Forscherin und Codiererin aus (Landis & Koch, 1977; Rössler, 2017) (siehe Anhang B Tabelle B5).

Legt man hier den gleich strengen Maßstab wie in der Auswertung der ersten offenen Teilaufgabe an und arbeitet nur mit den Fällen weiter, die eine völlige Übereinstimmung zwischen den Raterinnen zeigen, ergibt sich jedoch eine zu geringe Fallzahl. Für weitere Analysen werden deshalb die Codierungen der Autorin verwendet, da diese dem Verständnis entsprechen, auf dessen Basis die Hypothesen geprüft werden (vgl. Barth, 2017).

Bei der Analyse der Antworten wurde der Wert 0 vergeben, wenn eine Kategorie nicht genannt wurde. Die Häufigkeitsverteilung (siehe Anhang B Tabelle B6) zeigt, dass die Studierenden insgesamt Kommentierungen zu allen Kategorien vornahmen und keine der Kategorien unkommentiert blieb. Der Wert 1 wurde vergeben, wenn Studierende zwar Bezug auf eine Textstelle genommen hatten, die eine Barriere beschreibt, jedoch keine entsprechende Interpretation vornahmen. Um die Lernbarrieren in den Textvignetten zu *erkennen*, mussten diese wahrgenommen und hinsichtlich der potenziell negativen Auswirkung auf die Lehr-Lernsituation zutreffend interpretiert werden. Von einem Erkennen der Lernbarrieren konnte also nur ausgegangen werden, wenn für die Antworten der Wert 2 (*beurteilt*) oder 3 (*Verbesserungsvorschlag gemacht*, was eine Beurteilung voraussetzt) vergeben wurde. Der Wert 4 (*relevante Stelle wahrgenommen, aber falsch interpretiert*) wurde ebenfalls als *nicht erkannt* gewertet.

Die Werte wurden dichotomisiert, indem die Werte 2 (*beurteilt*) und 3 (*Verbesserungsvorschlag*) als *erkannt* (Wert 1) rekodiert wurden und die Werte 0 (*nicht erkannt*), 1 (*nacherzählend*

genannt) und 4 (*genannt und falsch interpretiert*) als *nicht erkannt* (Wert 0). Tabelle 9 zeigt, dass die *inhaltliche Klarheit* jeweils am häufigsten erkannt wurde (Sonderpädagogikstudierende 65.7%, Allgemeinpädagogikstudierende 40.4 %), das *Feedback* erkannten Studierende der Allgemeinpädagogik am seltensten (2.4 %), Studierende der Sonderpädagogik erkannten die Barriere in den *Regeln und Routinen* am seltensten (5.7 %).

Tabelle 9

Häufigkeiten für die dichotomisierten Variablen, erkannte Barrieren nach Studiengängen

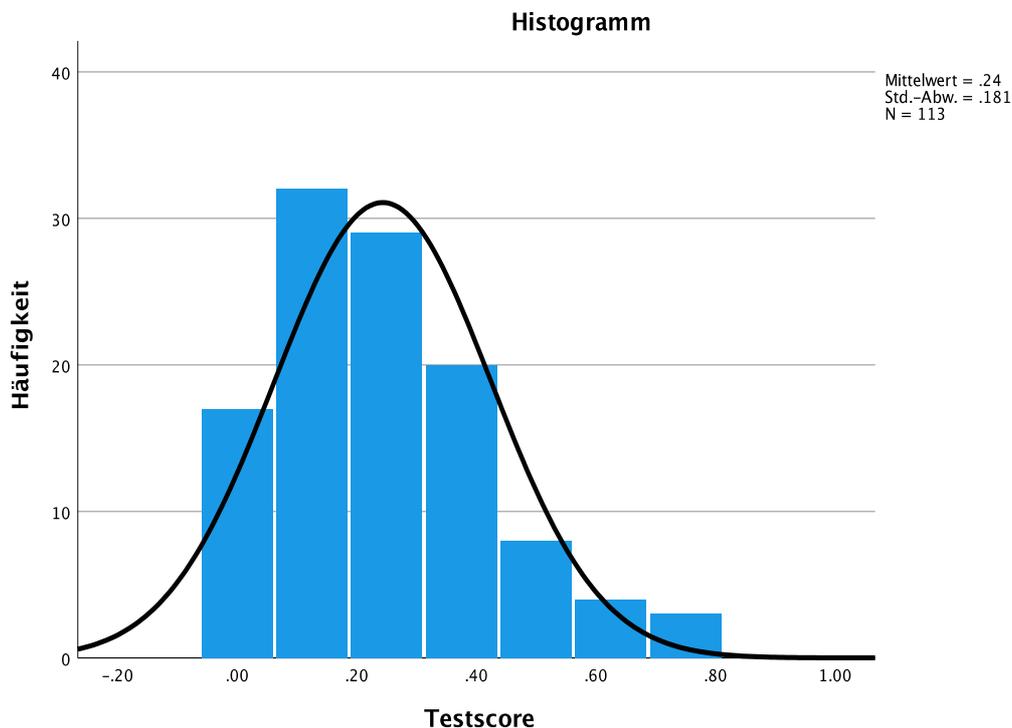
Item	Gruppe 1	Gruppe 2
	%	%
Raum	38.1	50.0
Inhaltliche Klarheit	40.4	65.7
Zielklarheit	19.0	20.0
Allgegenwärtigkeit	4.8	14.3
Übergänge	7.1	18.6
Veranschaulichung	21.4	27.2
Regeln und Routinen	7.2	5.7
Feedback	2.4	27.1

Anmerkung. Gruppe 1 = Studierende der Allgemeinpädagogik ($n = 42$), Gruppe 2 = Studierende der Sonderpädagogik ($n = 70$). Angaben in gültigen Prozent, höchste und niedrigste Werte sind fett.

Errechnet man den Mittelwert aller dichotomisierten Items (= Testscore), erreichen 15 % aller Studierenden einen Wert von 0, d. h. dass sie keine der Barrieren erkennen. Rund 86.7% aller Studierenden erkennen weniger als die Hälfte der Barrieren ($M < 0.50$), den Höchstwert ($M = 1.00$) erreicht niemand. Das Histogramm verdeutlicht diese Linksverteilung der Testscores (siehe Abbildung 5).

Abbildung 5

Histogramm zur Häufigkeitsverteilung der Testscores im Erkennen von Lernbarrieren

**9.5.6 Interpretation der Ergebnisse der zweiten offenen Teilaufgabe**

In der Vignette *Philipp* werden Interaktionen beschrieben, die Merkmale der KF und LU beschreiben. So sollte der Blick der Studierenden auf Unterrichtsbedingungen gelenkt werden, um Beurteilungsprozesse hinsichtlich dieser Merkmale in Gang zu setzen. Es sollte geprüft werden, ob der Impuls durch die spezifisch konstruierte Vignette in Kombination mit einer offenen Frage ausreicht. Die Analyse der Antworten zeigt, dass die Studierenden tatsächlich Bezug auf alle Stellen in der Vignette nehmen, die relevante Handlungen beschreiben, und diese zumindest von einigen Studierenden bemerkt werden. Die Vignettenabschnitte zu den Kategorien *Raum, Regeln und Routinen* oder *inhaltliche Klarheit* beispielsweise werden in beiden Studiengangsgruppen von mindestens der Hälfte der Studierenden kommentiert (siehe Anhang B Tabelle B6). Andere Lernbarrieren wie das *Feedback, Übergänge* oder die *Zielklarheit* werden seltener kommentiert. Insgesamt stützt dieses Ergebnis die Hypothese 2.

Allerdings zeigt die Analyse der Qualität der Antworten durch die Dichotomisierung, dass das Bemerkens und Kommentieren relevanter Interaktionen häufig noch nicht einem Erkennen entspricht. Äußerungen wie z. B. *Philip ist frustriert, weil die anderen schneller sind und reinrufen* zeigen zwar, dass hier eine problematische Situation bemerkt wird. Eine Interpretation im Sinne des Erkennens einer Lernbarriere, die in den Unterrichtsmerkmalen liegt und nicht bei Philip,

wird hier aber noch nicht deutlich. Gerade für das Merkmal *Regeln und Routinen* ist der Anteil zutreffender Beurteilungen gering (Sonderpädagogikstudierende: 5.7 %, Allgemeinpädagogikstudierende: 7.2 %). Insgesamt erkennen nur 6.2 % aller befragten Studierenden mehr als die Hälfte der Barrieren und niemand erkennt alle Lernbarrieren. Dagegen erkennen 15 % aller Studierenden keine der Barrieren, so dass sich hier ein Bodeneffekt zeigt und eine Kompetenzerfassung nicht angemessen möglich ist (Lienert & Raatz, 2011).

Die Erprobung der Vignette in Kombination mit einer offenen Frage gibt also Hinweise auf die Tauglichkeit der Vignette, zeigt aber auch, dass eine offene Frage den Blick noch zu wenig auf Unterrichtsbedingungen lenkt. Zudem wird deutlich, dass die Art und Weise der Formulierungen Einfluss auf die Genauigkeit der Codierung hat. Die einzelnen Merkmale der KF und LU, die bei der Konstruktion der Vignette und der Ankerbeispiele des Codierleitfadens genutzt werden, lassen sich in den Antworten nicht so detailliert wiederfinden, dass diese sich sicher einzelnen Lernbarrieren zuordnen lassen. Wenn Antworten weniger elaboriert und eher global formuliert sind, führen komplexere Situationen eher zur Beschreibung eines Gesamteindrucks ohne ausreichende Begründung. Aussagen wie *Die Lehrerin unterstützt Philipp nicht genug, so dass er nicht ins Arbeiten kommt* sprechen dafür, dass eine oder mehrere Barrieren erkannt wurden. Es wird jedoch nicht deutlich, ob das z. B. auf einen schlechten Übergang, mangelnde Zielklarheit, zu wenig inhaltliche Klarheit oder durch fehlende Verständnissicherung bedingt ist. Anders als im theoretischen Modell, in dem einzelne Merkmale beschrieben werden, sind diese in der Situation ineinander übergehend und miteinander verbunden. Das führt dazu, dass die Bewertung der Antworten nicht immer eindeutig möglich ist und es zu einer nur mäßigen und teils nicht zufriedenstellenden Interrater-Reliabilität kommt (Landis & Koch, 1977).

Auch wenn das Wahrnehmen relevanter Situationen als Voraussetzung für das Erkennen durch eine strukturierte Vignette ausgelöst werden kann, scheint es nötig zu sein, die Beurteilung dieser Situationen durch ebenso gezielte Impulse anzustoßen. Die Items der geschlossenen zweiten Teilaufgabe waren deshalb so konstruiert, dass sie eine Beurteilung der jeweiligen Barrieren direkt abfragten.

9.6 Ergebnisse und Interpretation des quantitativen Ansatzes

In der geschlossenen Version der zweiten Teilaufgabe wird die Vignette *Philipp* in Kombination mit Items präsentiert und die Frage untersucht, ob eine solche fokussierte Befragung die gewünschten Beurteilungsprozesse auslöst und beurteilbar macht.

9.6.1 Ergebnisse zu mittleren Schwierigkeitsgraden, Trennschärfen und Reliabilität

Frage 3: Weisen die konstruierten Items akzeptable Werte hinsichtlich der Schwierigkeit, Trennschärfe und Reliabilität auf?

Hypothese 3: Die konstruierten Items sind angemessen schwierig, trennscharf und reliabel.

Die erzeugte Schwierigkeit wird anhand einer deskriptiven Statistik dargestellt (siehe Tabelle 10). Es zeigt sich, dass die vierstufige Beurteilungsskala voll ausgenutzt wird, außer beim Item zur *Allgegenwärtigkeit*, für das nicht der Wert 1 vergeben wurde. Allerdings wurde insgesamt sehr selten der Extremwert 1 vergeben und häufig die Werte 3 und 4, so dass die Mittelwerte für die Items überwiegend deutlich über der Skalenmitte (2.50) liegen. Den niedrigsten Mittelwert ($M = 2.81$, $SD = 0.70$) erreicht das Item zu *Regeln und Routinen*, den höchsten Mittelwert ($M = 3.57$, $SD = 0.61$) erreicht das Item zur *Raumbarriere*. Die Mittelwerte aller Items liegen zwischen 1.88 und 4.00 ($M = 3.18$, $SD = 0.45$). Die Daten weisen dadurch zwar eine gewisse Streuung auf, verteilen sich jedoch rechtssteil.

Tabelle 10

Deskriptive Statistik zu geschlossenen Items zu Barrieren

Item	<i>n</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Var</i>
Raum	115	1	4	3.57	0.61	.37
inhaltliche Klarheit	114	1	4	3.36	0.80	.64
Veranschaulichung	115	1	4	3.16	0.89	.80
Zielklarheit	114	1	4	3.09	0.89	.79
Allgegenwärtigkeit	113	2	4	3.23	0.66	.43
Übergänge	114	1	4	2.82	0.84	.71
Regeln und Routinen	108	1	4	2.81	0.70	.49
Feedback	116	1	4	3.41	0.72	.52

Anmerkung. Höchste und niedrigste Mittelwerte sind fett markiert.

Zur Prüfung der Item-Kennwerte werden außerdem Cronbachs Alpha und Trennschärfen dargestellt. Die Trennschärfen der Items liegen zwischen $r_{j(t-j)} = .20$ und $r_{j(t-j)} = .61$. Bei einer ersten Itemrevision sollten die Trennschärfen der Items bei $r_{j(t-j)} = .25$ oder höher liegen (Lienert & Raatz, 2011). Dies trifft auf alle Items zu, außer auf das Item zur Barriere in den *Regeln und Routinen* ($r_{j(t-j)} = .20$). Der Ausschluss dieses Items würde die Reliabilität allerdings nicht deutlich erhöhen (siehe Tabelle 11). Das Cronbachs Alpha liegt für acht Items bei .73.

Tabelle 11*Trennschärfen der Items zu Barrieren*

Item	Skalenmittelwert, wenn Item wegge- lassen	Skalenvarianz, wenn Item weg- gelassen	Korrigierte Item- Skala-Korrelation	Cronbachs Alpha, wenn Item weg- gelassen
Raum	21.83	11.75	.25	.73
Inhaltliche Klarheit	22.06	11.23	.25	.73
Veranschaulichung	22.29	9.25	.57	.66
Zielklarheit	22.31	9.71	.48	.68
Allgegenwärtigkeit	22.18	11.00	.40	.70
Übergänge	22.63	9.47	.58	.66
Regeln und Routinen	22.60	11.69	.20	.74
Feedback	22.01	9.85	.61	.66

Anmerkung. Trennschärfen $r_{j(t-j)} \geq .25$ sind fett markiert.

9.6.2 Interpretation der mittleren Schwierigkeitsgrade, Trennschärfen und Reliabilität

Die Items der geschlossenen zweiten Teilaufgabe waren so konstruiert, dass sie eine Beurteilung der jeweiligen Barrieren direkt abfragten. Dies führte dazu, dass ein sehr großer Teil der Studierenden die Barrieren erkannte und die Mittelwerte für alle Items über der Skalenmitte (2.50) liegen. Diese Ergebnisse stützen die Hypothese 2, da durch die Kombination der Vignette *Philipp* (gezielte Impulse) mit einem geschlossenen Frageformat eine Beurteilung der Barrieren erreicht werden konnte.

Damit die Fähigkeit zum Erkennen von Lernbarrieren auf diese Weise erfasst werden kann, muss der Schwierigkeitsgrad der Items aber geeignet sein, verschieden ausgeprägte Fähigkeiten zu erfassen. Es darf also kein Deckeneffekt auftreten. Die rechtssteile Verteilung der Daten zeigt jedoch, dass die Beurteilung auf diese Weise zu leicht ist oder zumindest sehr eindeutig erfolgt, insbesondere in der Gruppe der Sonderpädagogikstudierenden. Es kann auf diese Weise nicht für alle Items genügend Varianz erzeugt werden, so dass unterschiedliche Fähigkeiten nicht sichtbar werden. Ziel des nächsten Entwicklungsschritts muss also sein, Aspekte zu überdenken, die den Schwierigkeitsgrad der Beurteilung beeinflussen könnten.

Mit einem Cronbachs Alpha von .73 kann die Reliabilität als akzeptabel bezeichnet werden. Angesichts der Tatsache, dass ein relativ breites Konstrukt mit verschiedenen Merkmalen erfasst wird, kann Cronbachs Alpha sogar als gut bewertet werden (Streiner, 2003), so dass die Daten Hypothese 3 stützen.

Das Item zu *Regeln und Routinen* wird insgesamt am seltensten richtig beurteilt und erreicht den geringsten Mittelwert. Zudem weist es als einziges Item eine zu geringe Trennschärfe ($r_{j(t-j)} = .20$) auf, so dass die Darstellung der Barriere in der Vignette geprüft werden muss.

9.6.3 Ergebnisse zur Validität

Zur Prüfung der Validität wird zum einen die Konstruktvalidität des Verfahrens erfasst. Aufgrund des pilotierenden Vorgehens wird die Struktur der Daten durch eine explorative Datenanalyse bestimmt. Außerdem wird die Kriteriumsvalidität überprüft und es wird dargestellt, ob durch die unterschiedlichen Ansätze gleiche oder verschiedene Konstrukte erfasst werden.

9.6.3.1 Ergebnisse zur Konstruktvalidität

Frage 4: Lässt sich das Erkennen von Barrieren in der Klassenführung von dem Erkennen von Barrieren in der Lernunterstützung unterscheiden?

Hypothese 4: Das Erkennen von Barrieren hängt davon ab, ob sie sich auf die Klassenführung oder die Lernunterstützung beziehen. Beides ist zu unterscheiden.

Um die Struktur der Daten zu untersuchen, wurde eine explorative Faktorenanalyse durchgeführt. Es wird angenommen, dass bei Vorliegen von mehreren Faktoren diese verschiedene Dimensionen der übergeordneten Fähigkeit *Erkennen von Lernbarrieren* darstellen und deshalb miteinander korrelieren. Als Rotationsmethode wird deshalb eine Promax-Rotation gewählt. Zwar ist die Datenmenge mit einer Vignette und acht Items klein. Durch die Faktorenanalyse sollte aber sichtbar werden, ob sich die Items latenten Konstrukten zuordnen lassen. Das Kaiser-Meyer-Olkin-Kriterium (.71) und die Signifikanz nach Bartlett ($p < .001$) weisen darauf hin, dass sich die Daten für eine Faktorenanalyse eignen. Die Anzahl der Faktoren wurde dabei so weit reduziert, dass nur Faktoren mit mindestens drei Faktorladungen über .30 entstanden (Lienert & Raatz, 2011), die auch inhaltlich interpretierbar waren. Es ergeben sich zwei Faktoren. Der erste Faktor erklärt dabei 35.86 % der Gesamtvarianz und der zweite 18.53 % (siehe Tabelle 12 und Anhang B Tabelle B7). Die Faktoren korrelieren mit $r = .20$ (siehe Anhang B Tabelle B8).

Tabelle 12*Ergebnis der Faktorenanalyse*

Item	Faktorladung	
	1	2
Faktor 1		
Veranschaulichung	.91	
Zielklarheit	.79	
Feedback	.72	
Übergänge	.70	
Inhaltliche Klarheit	.35	
Faktor 2		
Regeln und Routinen		.86
Raum		.68
Allgegenwärtigkeit		.60

Anmerkung. $N = 103$. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung. Faktorladungen unter .30 werden nicht dargestellt.

9.6.3.2 Interpretation der Konstruktvalidität. Aufgrund bisheriger Studien und des zugrundegelegten theoretischen Modells wurde ein zweifaktorielles Modell vermutet, das die Fähigkeit *Erkennen von Lernbarrieren* abbildet und dabei Items der KF und LU unterscheidet (Stefensky et al., 2015). Die Mustermatrix zeigt zwei Faktoren: Mit den Items zu *Raum*, *Regeln und Routinen* sowie *Allgegenwärtigkeit* wird der Faktor 2 von Items zur KF gebildet. Faktor 1 wird von Items zur LU gebildet (*Veranschaulichung*, *Zielklarheit*, *inhaltliche Klarheit*). Das *Feedback* kann laut theoretischer Modellierung je nach Gestaltung zur KF oder zur LU gezählt werden. Damit ordnet sich lediglich das Item zu *Übergängen* nicht dem Faktor zu, wie es der theoretischen Modellierung entspricht. Allerdings ist davon auszugehen, dass Merkmale nicht isoliert wahrgenommen werden, sondern in Handlungszusammenhängen. Zwar gehören reibungslose Übergänge als Merkmal zur prozessualen Strukturierung und damit zur KF, in der konkreten Umsetzung gelingt der Übergang in der beschriebenen Situation für Philipp aber auch wegen mangelnder LU nicht. Gerade Übergänge sind auch von anderen Merkmalen wie z. B. Regeln und Routinen oder Klarheit abhängig. Zusätzlich kommt in der Formulierung des Items *Die Lehrerin unterstützt Philipp bei der selbständigen Bearbeitung* das Verb *unterstützt* vor. Gemeint ist hier

der Übergang in die selbständige Bearbeitung. Die Formulierung könnte aber auch auf mangelnde LU verweisen, so dass eine Zuordnung der Übergänge zur LU hier inhaltlich gut nachvollziehbar ist.

Zu den Items *Raum* und *Regeln und Routinen* ist anzumerken, dass diese das leichteste bzw. schwerste Item sind und die geringsten Trennschärfen aufweisen. Die Umsetzung in der Vignette und die Erfassung durch die Items sind hier besonders zu prüfen.

Die Daten stützen Hypothese 4, sollten aber mit Vorsicht interpretiert werden. Strengere Angaben zur Interpretation von Faktoren fordern 4 Variablen $>.6$ oder 10 Variablen $>.4$, damit ein Faktor zuverlässig interpretiert werden kann und nicht mit zufälligen Ladungsmustern gerechnet werden muss. Ansonsten sollte die Stichprobe mind. 300 Personen umfassen oder die Datenstruktur in anderen Stichproben replizierbar sein (Guadagnoli & Velicer, 1988, zitiert nach Bortz, 1999).

9.6.3.3 Ergebnisse zur Kriteriumsvalidität

Frage 5: Ist die Beurteilungsaufgabe im Sinne von Kriteriumsvalidität sensibel genug, um Unterschiede zwischen Gruppen unterschiedlicher Expertise zu erfassen?

Hypothese 5: Es zeigen sich Unterschiede zwischen den Gruppen zugunsten der Sonderpädagogikstudierenden und Studierenden der höheren Fachsemester.

Zur Darstellung unterschiedlicher Fähigkeitsausprägungen wurden unter Berücksichtigung der Ergebnisse der Faktorenanalyse gewichtete Mittelwerte pro Item und pro Faktor berechnet. Vergleiche werden zwischen Beurteilungen der Fachsemester- und Studienganggruppen mit non-parametrischen Verfahren durchgeführt.

Zunächst wurde mit dem Kruskal-Wallis-Test geprüft, ob es in den Studienganggruppen signifikante Unterschiede in Abhängigkeit von der Studiendauer gibt. Das Signifikanzniveau wird auf 10 % festgelegt, um auch schwächere Effekte zu erkennen, die Hinweise für die weitere Entwicklung geben könnten. Es wird für die Post-hoc-Tests jeweils die angepasste zweiseitige Signifikanz berichtet.

Der Kruskal-Wallis-Test zeigt, dass nur in der Gruppe der Allgemeinpädagogikstudierenden das Ergebnis für Faktor 1 (LU) von der Studiendauer beeinflusst wird ($\chi^2(2) = 5.341$, $p = .069$) (siehe Anhang B Tabelle B9). Die Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass der Unterschied zwischen dem 1. Semester im Bachelor und 3. Semester im Master knapp signifikant ist ($z = -2.125$, $p = .101$) (siehe Anhang B Tabelle B10). Es handelt sich mit $r = .38$ um einen mittleren Effekt nach Cohen (1992).

Unterschiede zwischen den Urteilen der Studiengangsgruppen wurden mit dem Mann-Whitney-U-Test geprüft. Signifikante Unterschiede zeigen sich nur im 1. Bachelorsemester sowohl für Faktor 1 (LU), $z = 2.943$, $p = .003$, als auch für Faktor 2 (KF), $z = 2.025$, $p = .043$ (siehe Anhang B Tabelle B11 und B12). Es handelt sich für Faktor 1, $r = .44$ und für Faktor 2, $r = .30$ um mittlere Effekte nach Cohen (1992).

9.6.3.4 Interpretation der Kriteriumsvalidität. Für die verschiedenen Teilstichproben wurde angenommen, dass Studierende der Sonderpädagogik und in höheren Fachsemestern besser abschneiden. In der Gruppe der Sonderpädagogikstudierenden zeigt sich aber, dass sie bereits zu Beginn des Studiums hohe Werte erzielen und sich diese von höheren Fachsemestern wenig unterscheiden. Aufgrund der zu geringen Schwierigkeit und des Deckeneffekts in dieser Gruppe lässt sich über den Einfluss der Studiendauern auf die Fähigkeit zum Erkennen von Lernbarrieren also keine Aussage treffen.

In der Gruppe der Allgemeinpädagogikstudierenden zeigt sich dieser Effekt jedoch besonders zwischen dem 1. Bachelor- und 3. Mastersemester, so dass zumindest in dieser Gruppe die Studiendauer eine Rolle zu spielen scheint. Der signifikante Unterschied zwischen den Studiengangsgruppen zu Beginn des Studiums lässt vermuten, dass eine unterschiedliche Fokussierung eine Rolle spielen könnte. Möglicherweise sind Studierende der Sonderpädagogik schon von Beginn an mehr darauf konzentriert, problematische Lehr-Lernprozesse wahrzunehmen, und beurteilen diese eindeutiger als schlecht. Hypothese 5 wird durch die Ergebnisse also gestützt, es bleibt aber zu prüfen, ob sich die gezeigten Unterschiede zwischen den Fachsemestern auch für den Studiengang Sonderpädagogik zeigen lassen und ob bei insgesamt höherem Schwierigkeitsgrad auch über die Studiendauer hinweg Unterschiede zwischen den Studiengängen erkennbar werden.

9.7 Zusammenfassung und Konsequenzen für den nächsten Entwicklungsschritt

Der Vergleich der Ansätze mit unterschiedlichen Vignetten und Frageformaten sollte Aufschluss darüber geben, wie Vignetten konstruiert werden müssen und welche Frageformate geeignet sind, Urteilsprozesse hinsichtlich potenzieller Lernbarrieren in der KF und LU so auszulösen, dass das Erkennen dieser erfassbar und beurteilbar wird.

Die Ergebnisse zeigen, dass bei einem allgemeinen Impuls und einer offenen Abfrage zunächst verstärkt andere Bedingungen als Unterrichtsmerkmale in den Blick genommen werden. Als Beobachtungsfokus werden von der großen Mehrheit der Studierenden vor allem die individuellen Bedingungen des Kindes gewählt. Besonders wird die Lesefähigkeit des Kindes fokussiert, vermutlich ausgehend von der Kontextinformation, dass das Kind Leseschwierigkeiten hat. Auch wenn eine Vignette als Impuls gewählt wird, die insofern spezifischer ist, als dass sie in

konkreten Interaktionen Barrieren in Unterrichtsmerkmalen darstellt, konzentrieren sich viele Kommentare zur Situation auf das Kind. Zudem bleiben diese häufig auf einer nacherzählenden Ebene und beinhalten weder einen Theoriebezug noch eine Beurteilung hinsichtlich der möglichen Auswirkung auf den Lernprozess, so dass ein Erkennen potenzieller Barrieren nicht zum Ausdruck kommt. Beispielsweise wird zwar beschrieben, dass Philipp negativ auf den Zwischenruf eines Kindes während der Gruppenarbeit reagiert (*Wir sind schon fertig!*). Es wird jedoch keine Erklärung für dieses Verhalten gesucht, das in mangelnden Regeln und Routinen für die Gruppenarbeit, also in der KF liegen könnte. Der Beobachtungsfokus bleibt demnach verstärkt auf dem Kind, ohne Erklärungen in der Unterrichtsgestaltung zu vermuten. Wie im modifizierten Kompetenzmodell gezeigt, ist das aber die Voraussetzung für das Erkennen von Lernbarrieren. Eine weitere Voraussetzung für das Erkennen ist, dass über die Wahrnehmung der kritischen Situation hinaus eine Hypothesenbildung über die Ursachen stattfindet. Beides findet mit offenen Befragungen als Impuls nicht in ausreichendem Maß statt.

Der quantitative Ansatz fragt Barrieren gezielt ab und es werden auf diese Weise Daten generiert, die annehmen lassen, dass die Konstruktion der quantitativen Vignettenaufgabe grundsätzlich geeignet ist, reliabel und valide die Fähigkeit zum Erkennen von Lernbarrieren zu erfassen. Gleichzeitig zeigte sich eine gute Akzeptanz des Aufgabenformats durch die Studierenden. Probleme zeigen sich aber noch bzgl. des zu geringen Schwierigkeitsgrades, vor allem in der Gruppe der Sonderpädagogikstudierenden.

Im nächsten Schritt sollen Vignette und Items weiterentwickelt werden. Dies betrifft die inhaltliche und sprachliche Gestaltung der Vignette sowie methodisch das Format der Items. Beides mit dem Ziel, Schwierigkeiten zu erhöhen und Varianzen zu erzeugen, die eine Abbildung unterschiedlicher Fähigkeiten ermöglichen.

10 Zweiter Entwicklungsschritt

Die ersten Ergebnisse sprachen für eine Weiterarbeit an der quantitativen Version. Im nächsten Entwicklungsschritt waren dafür sowohl die Vignette als auch die Items weiter zu qualifizieren.

Zur Weiterentwicklung des Verfahrens stellt sich die Frage, wie es durch gezielte Änderungen verbessert werden kann. Das betrifft eine inhaltliche und sprachliche Überarbeitung der Darstellung der Barrieren in der Vignette sowie methodische Änderungen bzgl. der Itemkonstruktion und Instruktion. Ziele sind eine angemessene Deutlichkeit der Barrieren in der Vignette sowie Instruktions- und Itemformate, die eine genaue Beurteilung provozieren, um insgesamt einen guten Schwierigkeitsgrad zu erzeugen.

10.1 Überarbeitung der Vignette und Items

Da sich in der ersten Datenerhebung zum Teil bereits akzeptable Ergebnisse bzgl. Reliabilität, Faktorenstruktur und Hinweise für eine Kriteriumsvalidität gezeigt hatten, sollten inhaltliche Änderungen möglichst filigran und unter Beibehaltung der Grundstruktur erfolgen.

Als Zielsetzung wurde festgelegt, dass der Schwierigkeitsgrad insgesamt steigen und sich die Varianz erhöhen sollten, damit der Deckeneffekt in der Gruppe der Sonderpädagogikstudierenden abnimmt und unterschiedliche Fähigkeitsausprägungen abbildbar sind. Dazu sollten Barrieren nicht zu einfach zu erkennen sein, aber dennoch eindeutig durch die beschriebenen Handlungen repräsentiert werden, damit die Beurteilungen nicht zufällig vorgenommen werden. Dafür sollten folgende Punkte geprüft werden: (1) die Wirkung der Kontextinformation (*Philipp hat eine spezifische Lernstörung...*), (2) die sprachliche Gestaltung der Vignette, (3) die inhaltliche Gestaltung der Vignette hinsichtlich der Frage, ob die Barrieren in den Vignetten in angemessener Deutlichkeit dargestellt sind. D. h., ob sie gut erkennbar, aber nicht zu offensichtlich sind. Ferner sollte geprüft werden, ob (4) die Items repräsentativ für die Barrieren in den Merkmalen der KF und LU sind und (5) ob das Itemformat und die Instruktion methodisch so verändert werden können, dass eine genauere Beurteilung der einzelnen Barrieren angeregt werden kann, die über ein Pauschalurteil hinausgeht.

In Interviews können in der Entwicklungsphase mehr Hinweise für die Gestaltung der Vignetten und Items erarbeitet werden als in Ratings, da Nachfragen, Diskussionen und Erläuterungen von Entscheidungen möglich sind. Zu erwarten ist, dass Personen, die sich ausführlich mit

Unterrichtsprozessen befassen, geeignete Expertinnen und Experten sind, um die in der Vignette beschriebenen Unterrichtsausschnitte, enthaltene Barrieren und Itemformulierungen sowie Instruktionen zu diskutieren.

Es konnte eine Expertengruppe aus Lehrenden der Universität Hamburg, die innerhalb des Projekts *ProfaLe* mit der Grundidee der Studie vertraut waren, gewonnen werden. Es wurden acht Expertinnen und Experten über das theoretische Modell und die Operationalisierung der Lernbarrieren informiert und gebeten, die Vignette analog zu den Studierenden in einer Papier- und Bleistift-Version innerhalb von 15 Minuten zu bearbeiten. In einer anschließenden Gruppendiskussion wurden Beurteilungen verglichen sowie die Vignette, Instruktion und Items diskutiert. Dabei wurde Konsens zu den folgenden Vorschlägen erzielt, die dann in die Überarbeitung aufgenommen wurden (siehe Tabelle 13).

Tabelle 13

Ergebnisse der Expertendiskussion und Überarbeitungen von Vignette und Items

Ergebnisse der Expertendiskussion	Überarbeitung von Vignette und Items
Kontextinformation	
<i>Feedback, Veranschaulichung:</i> Vermutlich starker Einfluss der Kontextinformation (<i>Philipp ist ein Kind mit spezifischer Lernstörung</i>) insbesondere hinsichtlich der Angemessenheit des Feedbacks und der fehlenden Veranschaulichung der Aufgabe.	Kontextinformationen fallen weg
Sprachlich	
Die Bezeichnung „Textaufgabe“ ist nicht korrekt.	Vorher: <i>Wir beide gucken jetzt mal zusammen, wie du diese Textaufgabe lösen kannst.</i> Nachher: <i>Wir beide gucken jetzt mal zusammen, wie du diese <u>Sach</u>aufgabe lösen kannst.</i>
Vorgehen bei der selbstständigen Bearbeitung soll stärker betont werden und weniger der Rechenweg.	Vorher: <i>Ich würde ganz gut finden, wenn du laut denkst, damit ich schauen kann, welche Rechenstrategien du dir überlegst und wie wir die nutzen können.</i> Nachher: <i>Ich würde ganz gut finden, wenn du laut denkst, damit ich schauen kann, welche <u>Strategien</u> du dir überlegst und wie wir die nutzen können.</i>
Die drei Punkte am Ende der Vignette sind „irritierend“.	Vorher: <i>Wenn du dich schon wieder von vornherein weigerst, dann kommen wir nie weiter...“</i>

Nachher: *Na, komm. Wenn du dich schon wieder von vornherein weigerst, dann kommen wir nie weiter!*“

Inhaltlich

Raumsituation: Dass Philipp am Gruppentisch sitzt, aber einzeln arbeiten soll, ist zu deutlich wahrnehmbar.

Vorher: *Philipp soll die Aufgabe mit Hilfe seiner Lehrerin lösen.*

Nachher: *Luise arbeitet noch nicht. Die Lehrerin geht zu ihr.*

Inhaltliche Klarheit: Die Klarheit wird an der Aufgabe selbst festgemacht. Es wird nicht deutlich, dass das Kind auch aufgrund der Darbietung der Aufgabe nicht mit der Arbeit beginnt.

Vorher: *Vor ihm liegt ein Arbeitsblatt mit einer Textaufgabe.*

Nachher: *Vor ihr liegt ein Arbeitsblatt mit einem Text, den sie nicht anschaut.*

Regeln und Routinen: Hinweise zu *Regeln und Routinen* werden kaum gefunden.

Vorher: *Wir sind schon fertig!*

Nachher: *Wir sind schon fertig! Was sollen wir jetzt machen?*

Methodisch

Die Formulierungsrichtung der Items ergibt ein einheitliches Antwortmuster (bei richtigem Erkennen alle Kreuze links).

Änderung der Formulierungsrichtung einiger Items.

Alle Items beziehen sich auf Barrieren, was ein „Triggern“ nach Barrieren zu suchen zur Folge haben könnte.

Es werden fünf Distraktoren hinzugefügt, die im Zusammenhang mit der geschilderten Situation sinnvoll erscheinen, aber inhaltlich nicht im Zusammenhang mit der Beurteilung der in der Vignette enthaltenen Lernbarrieren stehen.

Die Lernunterstützung wird nicht genügend durch die Items repräsentiert.

Es wird das Item *Die Lehrerin bietet Luise passende Lernunterstützung an* hinzugefügt.

Die vierstufige-Likert-Skala führt möglicherweise einerseits zu weniger sorgfältigen Urteilen, andererseits weisen Kreuze zwischen den Antwortmöglichkeiten darauf hin, dass die Studierenden nicht genug Ausdrucksmöglichkeiten für ihre Einschätzungen finden.

Die Likert-Skala wird zu einer sechsstufigen Skala geändert: *Trifft ganz und gar nicht zu, Trifft kaum zu, Trifft weniger zu, Trifft eher zu, Trifft weitgehend zu, Trifft voll und ganz zu*

Das Item zur Allgegenwärtigkeit soll sowohl das Im-Blick-Haben der Klasse als auch die passende Reaktion auf Störungen beinhalten.

Vorher: *Die Lehrerin hat die Gesamtsituation im Griff.*

Nachher: *Die Lehrerin hat die Gesamtsituation im Griff und reagiert passend auf Störungen.*

Das Item zu Übergängen soll deutlicher von der Lernunterstützung abgegrenzt werden und klarer

Vorher: *Die Lehrerin unterstützt Philipp bei der selbständigen Bearbeitung.*

machen, dass es um den Übergang in die selbständige Arbeit geht.

Nachher: *Die Lehrerin bietet Luise keine Anleitung zur selbstständigen Bearbeitung der Aufgabe an.*

Die Formulierung des Items zu Regeln und Routinen soll einen stärkeren Hinweis auf die Barriere enthalten.

Vorher: *Die Gruppenarbeit in der Klasse läuft nach festen Regeln ab.*

Nachher: *Der Ablauf der Gruppenarbeit erfolgt nach festen Regeln und Routinen.*

Formal

Kreuze zwischen den Antwortmöglichkeiten sollen vermieden werden.

In der Instruktion bekommen die Studierenden zusätzlich den Hinweis, Kreuze nur „IN“ die Kästchen zu setzen.

Personenangaben sind durch offene Abfrage zu oft nicht eindeutig zuordenbar.

Personenangaben werden durch Multiple-Choice-Fragen erfasst.

Anmerkung. Unterstreichungen markieren geänderte Textabschnitte. Der Name des Kindes wurde von Philipp zu Luise geändert.

Im Ergebnis wurden inhaltlich Änderungen in der Vignette vorgenommen und Kontextinformationen gestrichen, um die Deutlichkeit in der Darstellung der Barrieren zu beeinflussen. Die Präzisierungen sollen dazu führen, dass alle Merkmale und dazugehörigen Items beurteilbar sind und nicht aufgrund eines insgesamt negativen Gesamteindrucks im Einklang mit den anderen Items negativ beurteilt werden. Die Veränderungen der Items und der Urteilsskala sollten die Anforderungen der Bearbeitung dahingehend erhöhen, dass im Urteilsprozess ein genaueres Hinsehen und Abwägen des Urteils erzeugt werden sollte. Für den Fall, dass Studierende bereits die Vignette *Philipp* bearbeitet hatten, wurde das Kind in der überarbeiteten Vignette *Luise* genannt, um Wiedererkennungseffekte von z. B. Kontextinformationen zu Philipp, wie *Philipp hat eine Leseschwäche*, zu verringern.

Es wurden somit von der ersten zur zweiten Datenerhebung mehrere Veränderungen in einem Schritt unternommen. Dies ist dem Umstand geschuldet, dass ein begrenzter Zugang zu unterschiedlichen Kohorten von Studierenden bestand und Datenerhebungen dadurch limitiert waren.

Vignette und Items lauteten nach der Überarbeitung wie folgt (Masterlösung analog zu Vignette *Philipp*, Bepunktung der Items siehe Anhang C1):

Stellen Sie sich die folgende Situation vor. Beurteilen Sie bitte anschließend die Aussagen auf den nächsten Seiten!

Die Schülerin Luise sitzt während des Matheunterrichts in ihrer Klasse an einem Gruppentisch. Vor ihr liegt ein Arbeitsblatt mit einem Text, den sie nicht anschaut. Die anderen Kinder an ihrem Tisch haben bereits in Gruppenarbeit mit der Bearbeitung der Aufgabe begonnen und diskutieren lebhaft. Luise arbeitet noch nicht. Die Lehrerin geht zu ihr.

Lehrerin: „Wir beide gucken jetzt mal zusammen, wie du diese Sachaufgabe lösen kannst. Hier steht:

„Für die Musik-Aufführung in der Schulturnhalle stellt der Hausmeister 280 Stühle auf. Außerdem benötigt der Elternbeirat noch 20 Stühle für den Verkauf. Im Orchester spielen 24 Kinder, die während der Aufführung stehen. Die Stühle sind in 6er Stapeln im Material-Lager gestapelt.“

Okay?“ (wartet kurz, Luise nickt) „Dann wollen wir jetzt mal überlegen, wie du das angehen kannst. Ich würde ganz gut finden, wenn du laut denkst, damit ich schauen kann, welche Strategien du dir überlegst und wie wir die nutzen können. Und dann kann ich dir immer dann helfen, wenn du allein nicht weiterkommst. Was haben wir denn da jetzt alles für Informationen? Vielleicht fängst du mal mit der Frage an?“

Die Lehrerin bleibt neben Luise stehen und wartet, Luise schaut auf den Text und bleibt still.

Dann rempelt Luise ihren Sitznachbarn an und schimpft: „Mann, jetzt rück doch mal ein Stück weg!“

Lehrerin: „Na, na, Luise, du hast so viel Platz wie immer. Nun fang mal an!“

Luise: „Und wie?“

Lehrerin: „Wie gesagt, vielleicht mit der Frage. Versuch mal.“

Luise: „Wie geht die denn?“

Lehrerin: „Das wollen wir ja jetzt rausfinden. Wie könnten wir vorgehen?“

Luise schaut ihre Lehrerin erwartungsvoll an, diese wartet geduldig.

Ein Kind vom Nachbartisch ruft derweil: „Wir sind schon fertig! Was sollen wir jetzt machen?“

Luise knallt sauer den Stift auf den Tisch: „Oh Mann, ich hab' keine Lust! Ich mach das nicht!“

Lehrerin: „Na, komm. Wenn du dich schon wieder von vornherein weigerst, dann kommen wir nie weiter!“

<i>Raum</i>	Die Arbeitssituation ist für Luise angemessen.
<i>Inhaltliche Klarheit</i>	Die Aufgabenstellung ist für Luise angemessen.
<i>Distraktor</i>	Die Lehrerin ist freundlich zu Luise.
<i>Veranschaulichung</i>	Die Lehrerin stellt nicht sicher, dass Luise die Aufgabe versteht.
<i>Lernunterstützung</i>	Die Lehrerin bietet Luise passende Lernunterstützung an.
<i>Distraktor</i>	Luise bemüht sich, die Aufgabe zu bearbeiten.
<i>Zielklarheit</i>	Die Lehrerin versäumt es, Luise das Vorgehen bei der Bearbeitung zu verdeutlichen.
<i>Distraktor</i>	Luise lässt sich beim Bearbeiten der Aufgabe ablenken.
<i>Allgegenwärtigkeit</i>	Die Lehrerin hat die Gesamtsituation im Griff und reagiert passend auf Störungen.
<i>Übergänge</i>	Die Lehrerin bietet Luise keine Anleitung zur selbstständigen Bearbeitung der Aufgabe an.
<i>Distraktor</i>	Die besondere Schwierigkeit der Aufgabe liegt darin, dass die Kinder die Frage selbst formulieren sollen.
<i>Regeln und Routinen</i>	Der Ablauf der Gruppenarbeit erfolgt nach festen Regeln und Routinen.
<i>Feedback</i>	Die Lehrerin gibt Luise unangemessene Rückmeldungen.
<i>Distraktor</i>	Es handelt sich bei der Sachaufgabe um ein gängiges Aufgabenformat.

10.2 Fragestellungen und Hypothesen

Ziel der zweiten Datenerhebung war, durch gezielte inhaltliche und methodische Änderungen bei gleichzeitiger Beibehaltung der Grundstruktur Probleme der ersten Datenerhebung zu beseitigen. Daraus ergaben sich folgende Fragestellungen und Hypothesen:

Mittlere Schwierigkeitsgrade, Trennschärfen und Reliabilität

Frage 1: Die Überarbeitungen der Vignette dienten der Präzisierung und insofern auch der besseren Erkennbarkeit der Barrieren. Die Beurteilung sollte dadurch aber nicht einfacher werden, sondern in Kombination mit den methodischen Veränderungen genauer. Es sollte so einem pauschal negativen Beurteilen entgegengewirkt werden. Lassen sich durch diese gezielten Änderungen die mittleren Schwierigkeitsgrade, die Trennschärfen und die Reliabilität im Vergleich zur ersten Version verbessern?

Hypothese 1.1: Die mittleren Schwierigkeitsgrade erhöhen sich im Vergleich zur ersten Version, so dass sich zutreffende von weniger zutreffenden Beurteilungen unterscheiden lassen.

Hypothese 1.2: Die Trennschärfen der Items und die Reliabilität bleiben mindestens erhalten oder erhöhen sich.

Validität

Frage 2: Die Ergebnisse der ersten Datenerhebung sprachen dafür, dass sich die Faktoren Klassenführung und Lernunterstützung trennen lassen. Nur das Item zu *Übergängen* gehörte, anders als im theoretischen Modell postuliert, zur Lernunterstützung. Führen die Veränderungen und Präzisierungen dazu, dass sich weiterhin zwei Faktoren zeigen, die der theoretischen Modellierung entsprechen?

Hypothese 2: Klassenführung und Lernunterstützung zeigen sich in einem zweifaktoriellen Modell.

Frage 3: Ziel war, durch Präzisierungen und methodische Veränderungen genauere Beurteilungen zu erhalten. Führen diese dazu, dass sich die Kriteriumsvalidität erhöht, indem sich Einschätzungen verschiedener Fachsemester in der Gruppe der Sonderpädagogikstudierenden unterscheiden?

Hypothese 3: Die Veränderungen führen dazu, dass sich Unterschiede zwischen den Fachsemestern auch in der Gruppe der Sonderpädagogikstudierenden zugunsten höherer Fachsemester zeigen.

10.3 Stichprobe

Die überarbeitete Fassung wurde im Folgesemester nach der ersten Datenerhebung mit einer Kohorte von 142 Studierenden im 2. und 6. Bachelorsemester sowie im 4. Mastersemester als Paper-and-Pencil-Test an der Universität Hamburg durchgeführt. Die Erhebung beschränkte sich auf Studierende der Sonderpädagogik, weil in dieser Gruppe der Deckeneffekt auftrat. Die

Auswahl und Befragung der Studierenden erfolgte analog zur ersten Datenerhebung danach, welche Lehrenden kurzfristig in der Lage waren, den Vignettentest in ihrer Lehrveranstaltung durchzuführen, unabhängig vom jeweiligen Seminar- oder Vorlesungsthema. Innerhalb der Veranstaltungen nahmen alle Studierenden an der Befragung teil und Lehrende hatten keinen Einfluss auf die Durchführung, so dass erneut nicht von starken Stichprobeneffekten ausgegangen werden muss und die Fachsemesterguppen als unabhängige Stichproben behandelt werden (siehe Tabelle 14).

Tabelle 14

Häufigkeiten der Fachsemester für die Stichprobe der zweiten Datenerhebung

Studiengang	Fachsemester			
	2. Bachelor	6. Bachelor	4. Master	Gesamt
Sonderpädagogik	61	43	38	142

10.4 Design und Durchführung

Analog zur ersten Durchführung erfolgte die Bearbeitung innerhalb von Seminarzeiten als Papier-und-Bleistift-Test. Die Ankündigung erfolgte ebenfalls in gleicher Weise und durch die gleiche Person, so dass eine hohe Durchführungsähnlichkeit gegeben war. Wie auch in der ersten Durchführung durften die Studierenden bei der Beantwortung der Items zum Text zurückblättern. Die Zeit war auf 15 Minuten beschränkt. Diese Zeitspanne wurde im ersten Durchlauf von einigen Studierenden, die die offene Aufgabe bearbeitet hatten, als zu kurz beurteilt. Für die geschlossene Variante schien die Bearbeitungszeit angemessen. Innerhalb dieser Zeit bearbeiteten alle Studierenden nur die Vignette *Luise* und die dazugehörigen Items.

Die Auswertung erfolgte analog der ersten Datenerhebung. Alle Antworten wurden in SPSS (Version 28.0.1.1) übertragen und in Werte übersetzt. Je nach Formulierungsrichtung wurden diese rekodiert, so dass wieder ein hoher Wert dem Erkennen der Barriere entspricht (je nach Formulierungsrichtung des Items *trifft ganz und gar nicht zu* bzw. *trifft voll und ganz zu* = 6).

Die Daten wurden deskriptiv und hinsichtlich der Itemeigenschaften Schwierigkeiten und Trennschärfen ausgewertet. Die Struktur der Daten wurde durch eine Faktorenanalyse geprüft. Zur Ermittlung der Testscores wurde entsprechend der Faktorenanalyse der gewichtete Mittelwert aller Items pro Faktor berechnet. Unterschiede zwischen den Fachsemestern wurden durch non-parametrische Verfahren geprüft.

10.5 Ergebnisse und Interpretation

Nachfolgend sollen Ergebnisse fokussiert werden, die die gezielten Änderungen an Vignette und Items zu Verbesserungen hinsichtlich der genannten Fragestellungen im Vergleich zur ersten Studie untersuchen. Die Ergebnisse werden deshalb analog dazu dargestellt und verglichen.

10.5.1 Ergebnisse zu mittleren Schwierigkeitsgraden, Trennschärfen und Reliabilität

Frage 1: Lassen sich durch die gezielten Änderungen die mittleren Schwierigkeitsgrade, Trennschärfen und die Reliabilität im Vergleich zur ersten Version verbessern?

Hypothese 1.1: Die mittleren Schwierigkeitsgrade erhöhen sich im Vergleich zur ersten Version, so dass sich zutreffende von weniger zutreffenden Beurteilungen unterscheiden lassen.

Hypothese 1.2: Die Trennschärfen der Items und die Reliabilität bleiben mindestens erhalten oder erhöhen sich.

Die deskriptive Statistik (siehe Tabelle 15) zeigt, dass die sechsstufige Skala bei 7 von 9 Items ausgenutzt wurde. Für das Item *Allgegenwärtigkeit* wurde der Wert 1 nicht genutzt, für das Item *Raum* wurden die Werte 1 und 2 nicht genutzt. Die Extremwerte 1 und 6 wurden insgesamt seltener genutzt. Am häufigsten kommen die Werte 4 oder 5 vor (siehe Anhang C Tabelle C2). Bei einem Skalenmittelwert von 3.5 weist das Item zu *Feedback* den geringsten Mittelwert auf ($M = 3.91$, $SD = 1.27$) und das Item zu *Raum* den höchsten ($M = 5.15$, $SD = 0.79$). Der Itemmittelwert für die ganze Stichprobe liegt deutlich über der Skalenmitte ($M = 4.48$, $SD = 0.53$). Die Daten zeigen eine geringe Streuung für die leichtesten Items zu *Raum* und *inhaltlicher Klarheit* und etwas bessere Streuungen für die übrigen Items.

Tabelle 15*Deskriptive Statistik zu Items zu Barrieren*

Item	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>MW</i>	<i>SD</i>	<i>Var</i>
Raum	141	3	6	5.15	0.80	0.64
Inhaltliche Klarheit	140	1	6	4.88	0.98	0.96
Veranschaulichung	142	1	6	4.13	1.53	2.33
Lernunterstützung	142	1	6	4.68	1.00	1.00
Zielklarheit	139	1	6	4.42	1.23	1.51
Allgegenwärtigkeit	141	2	6	4.63	1.01	1.02
Übergänge	141	1	6	4.38	1.29	1.65
Regeln und Routinen	132	1	6	4.20	1.26	1.58
Feedback	138	1	6	3.91	1.30	1.68

Anmerkung. Niedrigste und höchste Mittelwerte sind fett markiert.

Um die Varianzen aus der ersten und zweiten Datenerhebung vergleichen zu können, wurde jeweils der Variationskoeffizient für die Items berechnet (siehe Tabelle 16). Der Variationskoeffizient gibt das prozentuale Verhältnis der Standardabweichung zum arithmetischen Mittel an, so dass sich die Werte der unterschiedlichen Skalen vergleichen lassen. Für den Großteil der Items zeigen sich nur geringfügige Unterschiede zwischen der ersten und zweiten Datenerhebung von 0.– 5 %. Für das Item *Veranschaulichung* steigt der Variationskoeffizient um rund 15 % und für das Item zum *Feedback* um rund 16 %. Dies bedeutet, dass die Beurteilungen dieser Items in der zweiten Datenerhebung nicht ganz so homogen erfolgten. Für den Mittelwert aller Items von 3.29 ($SD = 0.35$) liegt der Variationskoeffizient in der ersten Datenerhebung bei rund 11 % und in der 2. Datenerhebung ($M = 4.49$, $SD = 0.54$) bei rund 12 %. Bezogen auf den Gesamtmittelwert zeigt sich also nur ein sehr geringer Unterschied. Damit zeigen sich in beiden Datenerhebungen nur moderate Streuungen der Gesamtmittelwerte in der Gruppe der Sonderpädagogikstudierenden. Größere Streuungen sind für die einzelnen Items erkennbar.

Tabelle 16

Mittelwerte, Standardabweichung und Variationskoeffizienten der Items zu Barrieren für die Gruppe der Sonderpädagogikstudierenden, 1. und 2. Datenerhebung im Vergleich

Item	Datenerhebung 1				Datenerhebung 2			
	<i>N</i>	<i>M</i>	<i>SD</i>	VK	<i>n</i>	<i>M</i>	<i>SD</i>	VK
Raum	72	3.62	.49	13.54 %	141	5.15	.80	15.53 %
Inhaltliche Klarheit	70	3.49	.70	20.06 %	140	4.88	.98	20.08 %
Veranschaulichung	72	3.29	.72	21.88 %	142	4.13	1.53	37.05 %
Lernunterstützung	0	-	-	-	142	4.68	1.00	21.37 %
Zielklarheit	70	3.20	.83	25.94 %	139	4.42	1.23	27.83 %
Allgegenwärtigkeit	69	3.32	.63	19.69 %	141	4.63	1.01	21.81 %
Übergänge	70	2.97	.72	24.24 %	141	4.38	1.29	29.45 %
Regeln und Routinen	65	2.88	.72	25.00 %	132	4.20	1.26	30.00 %
Feedback	72	3.54	.60	16.95 %	138	3.91	1.30	33.25 %
Gesamt	72	3.29	0.35	10.64 %	142	4.49	0.54	12.03 %

Anmerkung. VK = Variationskoeffizient. Unterschiede des Variationskoeffizienten > 5 % sind fett markiert.

Zur Prüfung der Reliabilität des Verfahrens wurden Item-Kennwerte sowie Cronbachs Alpha berechnet. Wie oben erwähnt, sollten bei einer ersten Itemrevision die Trennschärfen der Items bei $r_{j(t-j)} = .25$ oder höher liegen (Lienert & Raatz, 2011). Dies trifft in der zweiten Datenerhebung auf 5 von 9 Items zu. Keins der Items liegt bei $r_{j(t-j)} > .5$ und alle außer dem Item *Raum* weisen eine geringere Trennschärfe als in der ersten Datenerhebung auf (siehe Tabelle 17). Cronbachs Alpha liegt in der zweiten Datenerhebung bei .49.

Tabelle 17*Trennschärfen der Items zu Barrieren*

	Skalenmittelwert, wenn Item wegge- lassen	Skalenvarianz, wenn Item wegge- lassen	Korrigierte Item- wenn Item wegge- Skala-Korrelation	Cronbachs Alpha, wenn Item wegge- lassen
Raum	35.13	19.16	.33	.44
Inhaltliche Klarheit	35.44	20.18	.10	.50
Veranschaulichung	36.16	16.84	.23	.46
Lernunterstützung	35.65	18.21	.32	.43
Zielklarheit	35.91	17.25	.31	.43
Allgegenwärtigkeit	35.70	18.44	.29	.44
Übergänge	35.98	16.71	.34	.41
Regeln und Routinen	36.08	20.51	.00	.54
Feedback	36.43	19.38	.09	.51

Anmerkung. Trennschärfen $r_{j(t-j)} \geq .25$ sind fett markiert.

10.5.2 Interpretation der mittleren Schwierigkeitsgrade, Trennschärfen und Reliabilität

Die deskriptive Statistik zeigt, dass nach den Änderungen der Vignetten und Items die Itemmittelwerte über der Skalenmitte ($M = 3.50$) liegen und die Beurteilung damit nach wie vor zu leicht ist. Allerdings wird der Extremwert 4 in der ersten Datenerhebung für 5 von 8 Items am häufigsten vergeben. In der zweiten Datenerhebung wird der Extremwert 6 dagegen für keines der Items am häufigsten vergeben. Es zeigen sich geringe bis moderate Veränderungen bzgl. der Streuung der Daten, dennoch ergibt sich eine rechtssteile Verteilung. Die Konzentration der Antworten ohne eine hohe Anzahl von Ausreißern spricht für eine eindeutige Beurteilbarkeit der Items. Es zeigt sich aber, dass die Breite der Skala die Varianzen nicht nennenswert beeinflusst und diese methodische Variation sowie die Verwendung von Distraktoren somit keinen entscheidenden Einfluss zu haben scheinen.

Bezüglich der Schwierigkeiten der einzelnen Items wird deutlich, dass die Änderungen nicht dazu geführt haben, dass die Items zu *Raum* und *inhaltlicher Klarheit* schwerer wurden. Änderungen in der Schwierigkeit und bzgl. des Variationskoeffizienten ergeben sich aber für die Items zum *Feedback* und zur *Veranschaulichung*, so dass sich die Rangfolge der Itemschwierigkeiten ändert (siehe Tabelle 18). Nach dem Wegfall der Kontextinformation wird das *Feedback* nicht mehr so deutlich wie zuvor als unangemessen beurteilt, sondern rutscht in der Rangfolge auf den letzten Platz – auch wenn es nach wie vor eher ein leichtes Item ist. Möglicherweise hat hier

aber nicht nur die fehlende Kontextinformation einen Einfluss, sondern auch die Umformulierung des Items, das nun negativ formuliert ist (*Die Lehrerin gibt Luise unangemessene Rückmeldungen*). Eine Bestätigung, dass das Feedback unangemessen ist, beinhaltet möglicherweise ein negativeres oder extremeres Urteil als die Verneinung der Aussage *Die Lehrerin gibt Luise angemessenes Feedback*. Aufgrund der unterschiedlichen Formulierungsrichtungen könnten Zustimmung- oder Ablehnungstendenzen eine Rolle spielen, die vom Iteminhalt unabhängig sind (Rostampour, 1998). Auch das Item zur *Veranschaulichung*, das vom 5. auf den vorletzten Rang absteigt, ist negativ formuliert und lautet in der überarbeiteten 2. Fassung *Die Lehrerin stellt nicht sicher, dass Luise die Aufgabe versteht*. Das gilt allerdings nicht für die ebenfalls negativ formulierten Items zu *Zielklarheit* und *Übergängen*, die nicht an Schwierigkeit gewinnen, so dass auch die Formulierungsrichtung der Items keinen eindeutigen Effekt zeigt.

Die Items *Feedback* und *Veranschaulichung* platzieren sich nun nach dem Item zu *Regeln und Routinen*, das somit in der Rangfolge vom letzten auf den 6. Rang vorrückt. Nach den inhaltlichen Änderungen und dem zusätzlichen Hinweis zu den fehlenden Regeln und Routinen in der Vignette (*Was sollen wir jetzt machen?*) ist das Item im Verhältnis zu den anderen Items also etwas leichter und weist einen etwas höheren Variationskoeffizienten auf. Es lässt sich aber nicht sicher beantworten, ob die zusätzlichen Hinweise in der Vignette zu diesem Effekt führen. Ein Einfluss des Geschlechts des Kindes sowie des Namens können ebenfalls nicht ausgeschlossen werden (Denn, 2021; Denn et al., 2015).

Insgesamt führen die Änderungen nicht zu zufriedenstellenden mittleren Schwierigkeitsgraden, so dass die Hypothese 1.1 verworfen werden muss.

Tabelle 18

Rangfolge der Items zu Barrieren gemessen am Schwierigkeitsgrad, 1. und 2. Datenerhebung im Vergleich

Rang	Items	
	Datenerhebung 1	Datenerhebung 2
1	Raum	Raum
2	Feedback	Inhaltliche Klarheit
3	Inhaltliche Klarheit	Allgegenwärtigkeit
4	Allgegenwärtigkeit	Zielklarheit
5	Veranschaulichung	Übergänge
6	Zielklarheit	Regeln und Routinen
7	Übergänge	Veranschaulichung
8	Regeln und Routinen	Feedback

Anmerkung. Je höher Items platziert sind, desto geringer ist die Itemschwierigkeit. Da es in Datenerhebung 1 kein Item Lernunterstützung gab, wird dieses im Vergleich der Rangfolge nicht dargestellt.

Reliabilität und Trennschärfen verschlechtern sich in der zweiten Datenerhebung deutlich, so dass auch Hypothese 1.2 verworfen werden muss. Das Item zum *Feedback*, das vorher die höchste Trennschärfe aufwies ($r_{j(t-j)} = .61$) und nach den Veränderungen zwar einen höheren mittleren Schwierigkeitsgrad und mehr Varianz zeigte, ist nun nicht mehr als trennscharf zu bewerten ($r_{j(t-j)} = .094$). Auch das Item zur *Veranschaulichung* (vorher $r_{j(t-j)} = .57$) hat erheblich an Trennschärfe verloren ($r_{j(t-j)} = .23$). Inhaltliche Veränderungen der Textstellen, die Hinweise auf diese Barrieren enthalten, wurden nicht vorgenommen, so dass hier ein starker Einfluss der Gesamtwahrnehmung, die durch Kontextinformationen beeinflusst wird, zu vermuten ist.

Relativierend muss hinzugefügt werden, dass die Reliabilität in der Gruppe der Sonderpädagogikstudierenden auch im ersten Durchgang nur bei einem Cronbachs Alpha von .57 liegt. Hier ist insofern ein Einfluss der Expertise auf die Reliabilität zu vermuten, als es in dieser Gruppe in der ersten wie zweiten Datenerhebung Deckeneffekte gibt, so dass eine Differenzierung kaum möglich ist (Streiner, 2003).

Insgesamt kann in dieser überarbeiteten Vignettenversion nicht mehr von einem in sich konsistenten Test ausgegangen werden, sondern es scheinen mit den Merkmalen unterschiedliche Konstrukte erfasst zu werden.

10.5.3 Ergebnisse zur Konstruktvalidität

Frage 2: Führen die Veränderungen und Präzisierungen dazu, dass sich weiterhin zwei Faktoren zeigen, die der theoretischen Modellierung entsprechen?

Hypothese 2: Klassenführung und Lernunterstützung zeigen sich in einem zweifaktoriellen Modell.

Wie in der ersten Datenerhebung wurde zur Prüfung der Datenstruktur eine explorative Hauptkomponentenanalyse mit Promax-Rotation durchgeführt. Das Kaiser-Meyer-Olkin-Kriterium (.62) und die Signifikanz nach Bartlett ($p < .001$) weisen darauf hin, dass sich die Daten grenzwertig für eine Faktorenanalyse eignen. Die Anzahl der Faktoren wurde so vorgegeben, dass nur Faktoren mit mindestens drei Faktorenladungen über .30 entstanden (Lienert & Raatz, 2011). Es entstanden so zwei Faktoren, von denen der erste rund 22 % der Gesamtvarianz erklärt und der zweite rund 15 % (siehe Tabelle 19 und Anhang C Tabelle C3). Die beiden Faktoren korrelieren mit $r = .16$ (siehe Anhang C Tabelle C4).

Tabelle 19

Ergebnis der Faktorenanalyse

Item	Faktorladung	
	1	2
Faktor 1		
Raum	.70	
Inhaltliche Klarheit	.62	-.42
Übergänge	.56	
Veranschaulichung	.54	
Zielklarheit	.47	
Faktor 2		
Feedback		.63
Lernunterstützung		.59
Allgegenwärtigkeit		.51
Regeln und Routinen		.39

Anmerkung. $N = 126$. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung. Faktorladungen unter .30 werden nicht dargestellt.

10.5.4 Interpretation der Konstruktvalidität

Die Fähigkeit zum Erkennen von Lernbarrieren beinhaltet, dass negative Auswirkungen von Interaktionen bzgl. der KF und LU erkannt werden. Diese Fähigkeit sollte sich als zweifaktorielles Modell in den Daten zeigen. Die entstandenen zwei Faktoren lassen sich aber nicht gemäß dem theoretischen Modell interpretieren. Während zuvor mit *Regeln und Routinen*, *Raum* und *Allgegenwärtigkeit* Items zur KF einen Faktor bildeten, mischen sich nun in beiden Faktoren Items zur KF und LU. Das theoretische Modell liefert hier keine Erklärung für die Datenstruktur. Abgesehen von der fehlenden inhaltlichen Interpretierbarkeit weisen die Faktorenladungen darauf hin, dass die Ladungsmuster auch zufällig sein könnten, da pro Faktor nur ein bis zwei Faktorladungen $>.6$ sind und die Stichprobe relativ klein ist (Guadagnoli & Velicer, 1988, zitiert nach Bortz, 1999). Hypothese 2 muss demnach verworfen werden bzw. kann nicht sicher beantwortet werden.

10.5.5 Ergebnisse zur Kriteriumsvalidität

Frage 3: Führen die Änderungen dazu, dass sich die Kriteriumsvalidität erhöht, indem sich Einschätzungen verschiedener Fachsemester in der Gruppe der Sonderpädagogikstudierenden unterscheiden?

Hypothese 3: Die Veränderungen führen dazu, dass sich Unterschiede zwischen den Fachsemestern auch in der Gruppe der Sonderpädagogikstudierenden zugunsten höherer Fachsemester zeigen.

Die Ergebnisse der Reliabilitäts- und Faktorenanalyse weisen nicht auf eine konsistente Skala hin. Analog zu Schritt 1 sollen trotzdem gewichtete Mittelwerte pro Item und pro Faktor berechnet und Gruppenunterschiede geprüft werden, um mögliche Tendenzen zu erkennen. Das Signifikanzniveau wird auf 10 % festgelegt. Es wird für die Post-hoc-Tests jeweils die angepasste zweiseitige Signifikanz berichtet.

Der Kruskal-Wallis-Test zeigt, dass die Ergebnisse bzgl. Faktor 1 von der Studiendauer beeinflusst werden, $\chi^2(2) = 8.090$, $p = .018$ (siehe Anhang C Tabelle C5). Die Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass der Unterschied zwischen dem 2. und 6. Bachelorsemester signifikant ist, $z = -2.799$, $p = .015$ (siehe Anhang C Tabelle C6). Es handelt sich mit $r = .27$ um einen schwachen Effekt nach Cohen (1992).

10.5.6 Interpretation zur Kriteriumsvalidität

Ziel war es, durch gezielte Veränderungen unterschiedliche Fähigkeitsausprägungen auch in der Gruppe der Sonderpädagogikstudierenden abbilden zu können. Auch wenn festgestellt wurde, dass mittlere Schwierigkeitsgrade und Varianzen noch nicht zufriedenstellend waren,

wurden Unterschiede zwischen den Fachsemestern für Faktor 1 gefunden: Studierende am Anfang des Bachelorstudiums unterscheiden sich auf diesem Faktor von Studierenden am Ende des Bachelorstudiums, was die 3. Hypothese stützen könnte. In dieser Studienphase war auch im ersten Entwicklungsschritt der größte Anstieg bei den Studierenden der Allgemeinpädagogik zu verzeichnen. Das Ergebnis zeigt aber nur eine schwache Effektstärke und ist angesichts der geringen Reliabilität und Trennschärfen sowie der wenig aussagekräftigen Faktoren nur bedingt interpretierbar. Dennoch wurden alle Items genutzt, auch die mit nicht idealen Werten, um zu prüfen, ob das Verfahren insgesamt mit der Abfrage aller Merkmale funktioniert. Ob sich Expertiseunterschiede tatsächlich abbilden lassen und durch die Studiendauer beeinflusst werden, ist zu prüfen.

Auffällig ist, dass die Mittelwerte für die Masterstudierenden etwas niedriger ausfallen als für die Bachelorstudierenden des 6. Semesters. Möglicherweise nutzen die Studierenden in den höheren Fachsemestern mehr Wissen, das zum Gelingen oder Misslingen einer Situation beitragen kann, und urteilen deshalb zurückhaltender. Zudem haben sie Lehrerfahrung im Kernpraktikum gesammelt und urteilen deshalb möglicherweise zurückhaltender über Lehrkräfte. Es fällt auf, dass die beiden Items mit dem höchsten Mittelwert keine Formulierungen zur Lehrkraft enthalten (*Die Arbeitssituation ist für Luise angemessen* und *Die Aufgabenstellung ist für Luise angemessen*). Die niedrigste Zustimmung geben die Masterstudierenden (das Item zu *Regeln und Routinen* ausgenommen) zu negativ formulierten Items zur Lehrkraft: *Die Lehrerin stellt nicht sicher, dass Luise die Aufgabe versteht*, *Die Lehrerin gibt Luise unangemessene Rückmeldungen* sowie *Die Lehrerin bietet Luise keine Anleitung zur selbstständigen Bearbeitung der Aufgabe an*. Möglicherweise müssen diesbezüglich Itemformulierungen bedacht werden und dahingehende Effekte überprüft werden.

10.6 Diskussion und Konsequenzen für das weitere Vorgehen

Ziel der zweiten Datenerhebung war es, zu prüfen, welche Effekte durch gezielte inhaltliche und methodische Änderungen bewirkt werden können. Dabei konnten die mittleren Schwierigkeitsgrade nicht im gewünschten Maß verbessert werden. Trennschärfen und Reliabilität verringerten sich und die Faktorenstruktur zeigte möglicherweise nur zufällige Ladungsmuster. Im Vergleich der Fachsemestergruppen zeigte sich zwar ein signifikanter Unterschied zwischen Studierenden des 2. und 6. Fachsemesters (Bachelor), der aber vor dem Hintergrund der schlechten Itemkennwerte mit Vorbehalt zu interpretieren ist.

Die Übersicht (siehe Tabelle 20) zeigt die jeweiligen Veränderungen, die im zweiten Entwicklungsschritt unternommen wurden, wobei das Item zur *LU* nicht aufgeführt wird, weil es in der ersten Datenerhebung noch nicht vorkam. Es fällt auf, dass alle drei Items, bei denen ein starker Kontexteffekt vermutet wurde, nach dem Wegfall der Kontextinformation deutlich an Trennschärfe verlieren.

Wie Avramidis und Norwich (2002) zeigen konnten, variieren Lehrkrachteinstellungen bzgl. Schülerinnen und Schülern mit funktionalen und sensorischen Behinderungen, kognitiven Beeinträchtigungen und sozial-emotionalen Verhaltensstörungen beträchtlich. Die Erfassung der Einstellungen von Lehrpersonen im Inklusionskontext ist daher besonders instruktiv, wenn sie spezifisch auf verschiedene Beeinträchtigungsformen bezogen wird. Dies könnte ein Hinweis darauf sein, dass auch bei der Vignettenbeurteilung entsprechende Kontextinformationen zum Kind eine Rolle spielen.

Die Items, die durch Änderungen innerhalb der Vignette beeinflusst werden sollten, ändern sich ebenfalls. Das Item zu *Regeln und Routinen* war bereits in der ersten Datenerhebung nicht trennscharf, weist aber in der zweiten Datenerhebung eine noch deutlich geringere Item-Skala-Korrelation auf. Die Items zu *Raum* und *Feedback* laden im Vergleich zur ersten Vignette *Philipp* jeweils auf dem anderen Faktor. Die Items zu *Zielklarheit*, *Allgegenwärtigkeit* und *Übergängen*, die nur von methodischen Veränderungen wie Item-Polung und -Formulierung betroffen waren, veränderten sich in den Kennwerten dagegen weniger stark.

Tabelle 20

Veränderungen der Trennschärfe und Faktorenzugehörigkeit in Bezug zu Überarbeitungen von Kontext, Vignette, Item-Polung und Item-Formulierung

Item	Vermuteter Einfluss der Kontextreduktion	Änderung der Barrierenbeschreibung in der Vignette	Änderung der Item-Polung	Änderung der Item-Formulierung	Verlust von Trennschärfe	Änderung der Faktorenzugehörigkeit
Raum		X				X
Inhaltliche Klarheit	X	X			X	
Veranschaulichung	X		X		X	
Zielklarheit			X			
Allgegenwärtigkeit				X		
Übergänge			X	X		
Regeln und Routinen		X		X	X	
Feedback	X		X		X	X

Anmerkung. Kreuze geben an, welche Items jeweils betroffen sind. Veränderungen, die zum Verlust von Trennschärfe oder einer Änderung der Faktorenzugehörigkeit führen, sind fett markiert.

Es liegt deshalb die Annahme nahe, dass methodische Änderungen weniger Einfluss auf die Beurteilung haben. Der Kontext und/oder geringfügige inhaltliche Informationen innerhalb der Vignette scheinen aber zu einem Gesamteindruck beizutragen, der die Beurteilungen beeinflusst. Insofern sind die Informationen, die bei der Vignettenkonstruktion gegeben werden, genau abzuwägen.

Zusammenfassend gibt es Hinweise darauf, dass das Erkennen von Lernbarrieren eher holistisch zu funktionieren scheint und dass methodische Variationen der Items eher weniger Einfluss nehmen, während Kontextinformationen, die auf den Gesamteindruck wirken, etwas sensibler zu sein scheinen, so dass sich Veränderungen hier stärker auswirken könnten.

Bei der Konstruktion neuer Vignetten muss deshalb Sorgfalt auf die Auswahl von Beispielen gelegt werden, die eindeutig sind und gut zugeordnet werden können. Zudem sollte deren inhaltliche Validierung im Auge behalten werden. Das methodische Vorgehen kann hingegen grundsätzlich erhalten bleiben. Gleichzeitig ist das Ziel, die Fähigkeit zum Erkennen einzelner Lernbarrieren losgelöst von einem Gesamteindruck zu erfassen. Insofern müssen Beurteilungsprozesse ausgelöst werden, die geeignet sind, eine genauere Wahrnehmung und Interpretation einzelner *Handlungen* hervorzurufen. Im Unterricht müssen Urteilsprozesse innerhalb kurzer Zeit in komplexen Situationen vollzogen werden. Würde man eine genauere Wahrnehmung der Lernbarrieren in Vignetten über eine Variation der Instruktion anregen, wie z. B. *Schauen Sie sich die einzelnen Vignettenabschnitte genau an und entscheiden Sie für jede einzelne Aussage, ob diese zutrifft oder nicht* wäre das ökologisch weniger valide. Vielmehr ist es ja gerade Teil der Kompetenz des Erkennens von Lernbarrieren, relevante Situationen wahrzunehmen und differenziert zu beurteilen und eben nicht Pauschalurteile zu fällen, die einzelne Lernbarrieren oder gelingende Situationen gewissermaßen schlucken. Insofern liegt in der Information, ob pauschale oder differenzierte Urteile gefällt werden, auch eine Aussage über die zu erfassende Kompetenz. Damit der Impuls geeignet ist, solche differenzierten Urteile zu ermöglichen, muss also die Gestaltung der Vignetten und Items entsprechend überdacht werden. Die Umsetzung dieser Ziele wird im folgenden Kapitel beschrieben.

11 Dritter Entwicklungsschritt

Der vergleichsweise starke Effekt der Kontextreduktion im zweiten Entwicklungsschritt führt zu der Frage, welche Informationen Studierende nutzen, um Vignetten zu beurteilen und Lernbarrieren zu erkennen. Möglicherweise beurteilen sie in Unterrichtssituationen nicht einzelne Handlungsabschnitte aufgrund differenzierter Informationen, sondern fällen eher Pauschalurteile aufgrund eines Gesamteindrucks. In Anbetracht der diagnostischen Aufgabe der sehr *spezifischen* Beobachtung potenzieller Lernbarrieren kommt dieser Frage eine hohe Bedeutung mit großer Praxisrelevanz zu. Daraus ergeben sich die Kernfragen für den dritten Entwicklungsschritt.

Erkenntnisse aus den vorherigen beiden Datenerhebungen sollen der Konstruktion neuer Vignetten und Items dienen. Neben Fragen nach der Testgüte stellt sich vor allem die Frage nach der Art der Wahrnehmung in Bezug auf das Erkennen von Lernbarrieren. Entspricht diese einem spezifischen Beobachten oder eher der Gesamtwahrnehmung einer Situation?

Die vorherigen Ergebnisse haben gezeigt, dass Barrieren in Unterrichtsmerkmalen durch ausreichende Hinweise in der Darstellung zwar gut erkennbar und eindeutig sein müssen, aber gleichzeitig nicht zu offensichtlich und damit zu leicht zu beurteilen sein sollten. Durch ein mehrschrittiges Konstruktions- und Validierungsverfahren soll dies erreicht werden.

11.1 Konstruktion der Vignetten und Items

Mit der Frage nach der Art der Wahrnehmung und möglichen Einflüssen auf das Erkennen von Lernbarrieren ist die Verwendung einer einzelnen Vignette nicht mehr ausreichend. Es sollen deshalb mehrere Vignetten konstruiert werden, die anhand gleicher Kriterien konstruiert, in ausgewählten Aspekten aber variiert werden, um mögliche Unterschiede bzw. Einflüsse hinsichtlich der Beurteilung erfassen zu können.

Alle Vignetten sollen Unterrichtssituationen beschreiben, in denen die sechs ausgewählten Merkmale der KF und LU vorkommen. Unterschieden wird jedoch zum einen, wie die jeweiligen Merkmale durch die beschriebenen Handlungen repräsentiert werden. Unterschiedliche Fähigkeitsausprägungen zum Erkennen von Lernbarrieren sollen davon unabhängig sichtbar werden.

Zum anderen wird in den Vignetten unterschieden, wie die Situierung hinsichtlich Klassenstufe, Unterrichtsfach, Geschlechtern der handelnden Personen und Arbeitsphase dargestellt wird. Studienergebnisse deuten darauf hin, dass die PW unter anderem auch vom fachlichen Kontext abhängig ist (Gold & Holodynski, 2017; Meschede, 2014; Sunder et al., 2016). Lernbarrieren wurden in dieser Arbeit jedoch so operationalisiert, dass sie auch ohne Fachwissen oder

fachdidaktisches Wissen beurteilbar sein sollen. Bei einer differenzierten Beurteilung soll sich also eine gewisse Konsistenz der Urteile über die verschiedenen Situationen hinweg zeigen. Folgende Kriterien werden inhaltlich und methodisch zur Konstruktion der Vignetten und Items überlegt:

Inhaltlich: Es sollen eindeutige und den Unterrichtsmerkmalen gut zuordenbare Unterrichtsbeispiele für Lernbarrieren gefunden werden. Die Deutlichkeit und Zuordenbarkeit sollen in einem mehrschrittigen Verfahren und durch wiederholte Einschätzungen von Expertinnen und Experten entwickelt und inhaltlich validiert werden (Moosbrugger & Kelava, 2007; Seidel et al., 2010). Die Beispiele für die Darstellung der Merkmale der KF und LU sollen bei guter Zuordenbarkeit dennoch nicht zu offensichtlich sein. So soll zum einen ein mittlerer Schwierigkeitsgrad erreicht werden. Zum anderen soll vermieden werden, dass Antworten auf Items aufgrund fehlender Hinweise in der Vignette geraten werden oder durch Orientierung an den anderen Items beantwortet werden (einheitlich negative Beurteilung).

Vignetten und Items sollen zudem geeignet sein, eine *differenzierte* Beurteilung anzuregen. Inhaltlich soll dies erreicht werden, indem nicht nur misslingende Unterrichtssituationen dargestellt wurden, die sich als Barriere auswirken könnten, sondern auch Situationen, in denen Maßnahmen der KF und LU gut gelingen. Bei einer differenzierten Beobachtung müssten diese Situationen dann anders als Barrieren beurteilt werden.

Methodisch: Es wird beibehalten, dass Items unterschiedlich gepolt und durch Distraktoren ergänzt werden, um einen genauen Beurteilungsprozess anzuregen. Für die Vignetten sollen jeweils unterschiedliche Itemsets verwendet werden, um Wiederholungseffekte durch identische Abfragen zu vermeiden. Es werden jeweils die gleichen Merkmale der KF und LU abgefragt, jedoch mit unterschiedlichen Formulierungen der Items. Da die sechsstufige Skala keine zufriedenstellende Verbesserung zeigte, wurde wieder eine vierstufige Skala verwendet, um die Komplexität zu reduzieren.

Die Bearbeitung von insgesamt vier Vignetten erschien dabei für die Studierenden zumutbar, um zu gewährleisten, dass die Bearbeitungszeit nicht viel mehr als 20 Minuten beträgt.

11.1.1 Konstruktion der Vignetten

Um authentische Beispiele für gut zuordenbare Situationen zu den Merkmalen der KF und LU zu finden, wurden Unterrichtsszenen auf die folgende Weise gesammelt:

Masterstudierende der Sonderpädagogik (nicht Teil der Stichprobe), die in der Kernpraktikumsphase bereits Praxiserfahrung gesammelt hatten, bekamen innerhalb eines Seminars eine

Schulung zu den Merkmalen der KF und LU und deren Operationalisierung wie sie für ErLe vorgenommen wurde. Anschließend sollten die Studierenden in Gruppenarbeit Unterrichtsbeispiele zu den einzelnen Merkmalen sammeln und notieren, die sie in der Praxisphase erlebt hatten. Diese Beispiele wurden vorgetragen und diskutiert und von den Seminarteilnehmerinnen und -teilnehmern auf ihre Zuordnung zu den Merkmalen überprüft und von den Gruppen ggf. überarbeitet. Die notierten Beispiele wurden anschließend eingesammelt und von der Autorin in Form von Kurzvignetten aufbereitet. Diese Kurzvignetten beinhalteten einzelne Handlungssequenzen, in denen jeweils Situationen zu einem Merkmal der KF oder LU beschrieben waren.

Im weiteren Seminarverlauf und mit Abstand von einigen Wochen bearbeiteten die gleichen Masterstudierenden einen Test, in dem diese Kurzvignetten den Merkmalen der KF und LU zugeordnet werden sollten. Dabei sollten sie die Zuordnung zunächst allein vornehmen und danach in kleinen Gruppen vergleichen. Die Kurzvignetten sollten diskutiert und auf ihre Zuordnung überprüft werden. Zusätzlich zu den zugeordneten Merkmalen der KF und LU gaben die Studierenden an, wie einig sich die Gruppe beim Zuordnen war. Bei der Auswertung wurde berücksichtigt, wie viele Studierende die jeweilige Kurzvignette einem Merkmal zugeordnet hatten und wie sicher diese Zuordnung erfolgte. Dabei stellte sich heraus, dass *Inhaltliche Klarheit* und *Zielklarheit* schlecht zu trennen waren. Diese wurden deshalb zu einem Merkmal (*Klarheit*) zusammengefasst. Außerdem waren Studierende unsicher, wie sich *Lernunterstützung* (allgemein) zeigt. Für dieses allgemeinere Item fehlte eine Operationalisierung. Die LU wirkt auf kognitive Prozesse, die nicht sichtbar sind. Maßnahmen mit dem Potenzial zur LU sind oft in einen komplexeren und auch fachlichen Kontext eingebettet und deshalb in einer kurzen Textvignette zum einen schwer zu beschreiben (Lipowsky & Bleck, 2019; Praetorius et al., 2014). Zum anderen sollten die Vignetten möglichst auch ohne spezifisches Fachwissen und fachdidaktisches Wissen beurteilbar sein. Auf das Item *Lernunterstützung* wurde deshalb bei der Konstruktion der neuen Vignetten und Items verzichtet. Die LU wird durch die Facette der *inhaltlichen Strukturierung* repräsentiert, die durch Beispiele für sprachliche Klarheit, die Klarheit bezüglich der Ziele von Lernaktivitäten und die Reduktion komplexer Lernsituationen dargestellt wird. Die LU war somit nur noch mit einem Item zur Klarheit vertreten. Es verblieben insgesamt sechs Merkmale: *Raum, Übergänge, Klarheit, Feedback, Regeln und Routinen* sowie *Allgegenwärtigkeit*.

Die Beispiele, die am sichersten und eindeutigsten für die Studierenden zuzuordnen waren, wurden dann für die Gestaltung von neuen Vignetten genutzt. Es entstanden vier Texte, die nach dem gleichen Prinzip aufgebaut sind, hinsichtlich der Geschlechter sowie beschriebener Unterrichtsfächer und -situationen aber variieren. In jeder Vignette werden Handlungen beschrieben, die die sechs Merkmale der KF und LU repräsentieren sollen. Diese sind jedoch nicht mehr in jedem Fall als misslungen und somit als potenzielle Barrieren beschrieben, sondern zum Teil

auch als gelingende Lehr-Lernsituation. Insgesamt wird jedes Merkmal dreimal als Barriere und einmal als gelingende Situation dargestellt. In den Vignetten 1 und 4 ist jeweils nur eine gelingende Situation beschrieben, in den Vignetten 2 und 3 jeweils zwei (siehe Tabelle 21).

11.1.2 Konstruktion und Zuordnung der Items

Die Studierenden sollten in jeder Vignette zum Nachdenken und differenzierten Beurteilen animiert werden, ohne sich an vorangegangenen Antworten zu orientieren. Zu den Vignetten wurde deshalb ein Pool unterschiedlich formulierter Items und Distraktoren entworfen. Alle Items wurden in einer Gruppe von drei Expertinnen diskutiert und geprüft und anschließend auf die passende Anzahl Items reduziert, die dann zufällig den Vignetten zugeordnet und in eine zufällige Reihenfolge gebracht wurden. Zu jeder Vignette wurden sechs Items zu den Merkmalen der KF und LU sowie drei Distraktoren zugeordnet. Insgesamt wurde jedes Merkmal zweimal durch ein positiv und zweimal durch ein negativ formuliertes Item abgefragt (siehe Tabelle 21).

Tabelle 21

Übersicht über die Verteilung der Barrieren und Itempolungen auf die vier Vignetten

Item	Vignette 1	Vignette 2	Vignette 3	Vignette 4
Raum	Barriere	Barriere (negativ formuliert)	gelingend (negativ formuliert)	Barriere
Übergänge	Barriere	gelingend	Barriere (negativ formuliert)	Barriere (negativ formuliert)
Klarheit	gelingend (negativ formuliert)	Barriere	Barriere	Barriere (negativ formuliert)
Feedback	Barriere	Barriere (negativ formuliert)	gelingend (negativ formuliert)	Barriere
Regeln und Routinen	Barriere (negativ formuliert)	gelingend	Barriere	Barriere (negativ formuliert)
Allgegenwärtigkeit	Barriere (negativ formuliert)	Barriere (negativ formuliert)	Barriere	gelingend

Es entstanden folgende Vignetten und Items:

Vignette 1 (Lösung siehe Anhang D1)

Unterrichtsbeginn Deutsch 4. Klasse. Die Kinder sitzen an Partnertischen.

Der Lehrerin fällt auf, dass Murat noch Kaugummi kaut. Lehrerin: „Murat, spuck bitte dein Kaugummi aus!“ Murat steht auf, spuckt sein Kaugummi in den Müll, setzt sich wieder hin. Lehrerin: „Hanna, spuck dein Kaugummi aus!“ Hanna steht auf, spuckt ihr Kaugummi aus. In der Zwischenzeit fängt die Klasse an zu kichern – Murat hat sich ein neues Kaugummi in den Mund gesteckt. Lehrerin: „Murat, spuck dein Kaugummi aus, muss ich das jedem einzeln sagen? Wir wollen uns heute noch mal mit den Wortarten beschäftigen. Ich schreib euch drei Sätze an die Tafel und unterstreiche einige Wörter. Ihr müsst die nicht abschreiben. Anschließend sucht ihr euch einen Satz aus und überlegt erst mal allein für euch, mit welcher Frage ihr nach den unterstrichenen Wörtern fragen könnt (erklärt ein Beispiel). Jetzt schreib ich erst mal an.“

Die Lehrerin schreibt an die Tafel. Es wird immer unruhiger. Tom beginnt, Papierkügelchen zu werfen, Luise zeigt Franz, die hinter ihr sitzt, ihre neuen Sticker. Jonas ruft schließlich laut: „Oh Mann, Leute!“ Lehrerin: „Ruhe jetzt! Jonas! Sucht einen Satz aus und notiert die Frage zu den unterstrichenen Wortarten. Ihr habt 5 Minuten Zeit.“

Die Stillarbeit beginnt. Emma stöhnt: „Oh Mann, Paul redet dauernd vor sich hin, das macht der immer! Das nervt!“ Lehrerin: „Versucht jetzt bitte mal, 5 Minuten still zu arbeiten.“

Nach 5 Minuten beendet die Lehrerin die Stillarbeit und beginnt, Ergebnisse zu sammeln, ein paar Kinder schreiben noch. Lehrerin: „Legt mal die Stifte hin und macht mit, sonst werden wir nicht fertig, ihr sollt das nachher auch noch geordnet aufschreiben.“ Paul: „Das ist doof, ich kann das irgendwie nicht.“ Lehrerin: „Du passt ja auch nie auf und quasselst immer!“

Items Vignette 1:

<i>Ablenker</i>	Die technische Ausrüstung in der Klasse ist ausreichend.
<i>Raum</i>	Die Lehrerin hat die Sitzordnung im Blick.
<i>Allgegenwärtigkeit</i>	Die Lehrerin reagiert zum falschen Zeitpunkt und unpassend auf Störungen.
<i>Ablenker</i>	Die Klassengemeinschaft ist stabil.
<i>Regeln und Routinen</i>	Für einen reibungslosen Ablauf fehlen Regeln und Routinen.
<i>Struktur und Übergänge</i>	Der Unterricht ist sinnvoll strukturiert.
<i>Feedback</i>	Die Lehrerin gibt Rückmeldungen zum Lernprozess.
<i>Ablenker</i>	In der Unterrichtsgestaltung fehlt Methodenvielfalt.
<i>Klarheit</i>	Den Schülerinnen und Schülern sind einzelne Teilschritte unklar.

Vignette 2 (Lösung siehe Anhang D2)

Sachunterricht Grundschule, Thema Magnetismus. Die Kinder sitzen an Gruppentischen. Kati kritzelt an ihrem Platz auf einem Zettel rum, während der Lehrer an der Tafel den Ablauf für die Gruppenarbeit erklärt.

Lehrer: „Kati, jetzt dreh dich doch bitte mal zur Tafel und schau hin.“ (erklärt weiter) „Also, ihr stellt erst mal Hypothesen auf und dann überprüft ihr die im Experiment und dokumentiert die Ergebnisse und macht euch vielleicht auch schon Notizen dazu, wenn euch was Bestimmtes auffällt, zum Beispiel, ähm, also einfach was euch auffällt. Die Aufteilung für die Gruppen kennt ihr, wer schon fertig ist, bleibt am Gruppentisch und kann wie immer Zeichnungen zu den Notizen machen. Los geht's!“

Die SuS beginnen eifrig, sich in den Gruppen zu beschäftigen. Der Lehrer wendet sich einer Gruppe zu, in der es heftige Diskussionen gibt, und bemüht sich intensiv um Klärung. Mit der Zeit wird es immer unruhiger in der Klasse. Lehrer: „Jetzt seid doch bitte mal noch einen Moment leise, damit die letzte Gruppe auch die Chance hat, das fertig zu machen! Seid ihr überhaupt schon alle fertig?“

Die Gruppenarbeit wird nach einiger Zeit durch einen Gong beendet und die SuS kommen im Kreis zusammen, um ausführlich die Ergebnisse zu besprechen. Eine Gruppe berichtet, dass Geldmünzen magnetisch sind (1 Cent), eine andere Gruppe hat das Gegenteil festgestellt (50 Cent). Die SuS sollen ihre Vermutungen zu den Gründen äußern.

Önder: „Vielleicht hat die andere Gruppe den Magneten falsch gehalten?“ Lehrer: „Was meinst du, Linn?“ Linn: „Das ist doch beides Metall. Vielleicht haben die den Magneten auch zu weit weg gehalten!“ Lehrer: „Der nächste, Louis.“ Louis: „Aber 50 Cent sind viel schwerer!“ Lehrer: „Kay, was meinst du?“ (...)

Items Vignette 2

<i>Ablenker</i>	In der Klasse gibt es Aufmerksamkeitsprobleme.
<i>Feedback</i>	Die Rückmeldungen des Lehrers beziehen sich zu wenig auf den Lernfortschritt.
<i>Struktur und Übergänge</i>	Einzelne Unterrichtsphasen sind gut aufeinander abgestimmt.
<i>Raum</i>	Der Lehrer sorgt unzureichend dafür, dass alle gut an ihrem Platz arbeiten können.
<i>Ablenker</i>	Die Aufgabenstellung ist sinnvoll erklärt.
<i>Ablenker</i>	Für die anschauliche Unterrichtsgestaltung fehlen technische Hilfsmittel.
<i>Allgegenwärtigkeit</i>	Der Lehrer versäumt, Unterrichtsprozesse aufmerksam zu steuern.
<i>Klarheit</i>	Der Lehrer strukturiert Inhalte.
<i>Regeln und Routinen</i>	Der Unterrichtsfluss gelingt durch eingeübte Rituale.

Vignette 3 (Lösung siehe Anhang D3)

Kunstunterricht 5. Klasse Stadtteilschule, die Klasse ist zu Unterrichtsbeginn noch sehr unruhig. Lehrerin: „Wir machen jetzt zum Locker werden mal 5 Minuten Bodypercussion!“ Sie fängt an, die SuS zögern noch. Lehrerin: „Los, alle mitmachen! Einfach machen, was ich mache! (...) Und jetzt im Loop!“ Ein paar SuS steigen wieder aus. Lehrerin: „Na gut, hab ich vielleicht nicht so gut vorgemacht. Aber so ist es ja jetzt toll ruhig hier, so dass wir anfangen können. Wir haben letzte Woche mit dem Thema „Zentangle“ begonnen und Konturen gezeichnet, heute geht es um die Muster. Ich zeig euch ein paar Beispiele am Smartboard, da seht ihr das am besten. Antoni, setz dich heut ruhig hier vorne hin. (...) Fangt schon mal an, ich geh noch mal ins Lehrerzimmer, da hab ich noch Schritt-für-Schritt-Anleitungen für ein paar Muster.“

Die Lehrerin verlässt den Klassenraum, die meisten Kinder arbeiten, aber es ist wieder unruhig. Nach einer Weile kommt die Lehrerin zurück. Antoni hat noch nicht mit Zeichnen begonnen. Die Lehrerin beginnt, ihm noch mal den nächsten Arbeitsschritt zu erklären. Sie unterbricht: „Moment, Antoni.“ (Zur Klasse:) „Ihr habt euch eben so toll beruhigt, bitte versucht mal, euch weiter auf eure Arbeit zu konzentrieren. Zentangle ist ja meditatives Zeichnen, da solltet ihr euch also drauf konzentrieren und das geht am besten, wenn ihr dabei nicht redet.“ (erklärt Antoni weiter)

Die Kinder zeichnen eine Weile. Alina und Carolin sind schnell fertig und fragen: „Dürfen wir jetzt Stempelbilder machen?“ Lehrerin: „Oh, ihr wart ja schnell! Ja, dann könnt ihr stempeln oder was ihr wollt.“ Alina und Carolin laufen los und wollen Stempel holen. Lehrerin: „Alina, Carolin, bitte räumt die Zeichenstifte weg.“ Alina und Carolin gehen zurück und räumen die Stifte weg und wollen dann stempeln. Lehrerin: „Bitte auch die Bilder in die Mappen tun!“ Alina und Carolin laufen wieder zurück und räumen die Bilder weg und laufen dann wieder zum Regal mit den Stempeln. Antoni: „Oh Mann, habt ihr's bald?“

Items Vignette 3

<i>Klarheit</i>	Die Lehrerin macht deutlich, was die Schülerinnen und Schüler tun sollen.
<i>Ablenker</i>	Die Binnendifferenzierung wird durch einen Mangel an Materialien erschwert.
<i>Feedback</i>	Die Lehrerin gibt unpassende Rückmeldungen.
<i>Regeln und Routinen</i>	Die Schülerinnen und Schüler sind geübt im Wechsel von Unterrichtsaktivitäten.
<i>Ablenker</i>	Die Lehrerin bereitet das Thema zu uninteressant auf.
<i>Struktur und Übergänge</i>	Der Unterrichtsverlauf ist stockend.
<i>Ablenker</i>	Der Umgangston der Klasse ist freundschaftlich.
<i>Raum</i>	Die Lehrerin lässt räumliche Gegebenheiten unbeachtet.
<i>Allgegenwärtigkeit</i>	Die Lehrerin hat das gesamte Unterrichtsgeschehen im Griff.

Vignette 4 (Lösung siehe Anhang D4)

5. Stunde Mathe. Der Lehrer kommt rein und begrüßt die Kinder. Lehrer: „Heute ist es so schön ruhig hier! Ach, der Paul ist nicht da!“ Gelächter, der Lehrer sorgt schnell für Ruhe und kündigt eine Aufgabe an. Die Kinder sitzen an Gruppentischen und sollen sich gegenseitig erklären, welche „schlau Rechenwege“ sie nutzen, um Aufgaben mit Zehnerübergang zu rechnen.

Lehrer: „Ihr habt jetzt genau 15 Minuten Zeit. Anna, du machst das mit mir, kannst da sitzen bleiben.“ (Geht zu Anna) „Erklär' mal wie du vorgehst.“ Anna: „Ich weiß ja eben nicht, wie ich das machen soll.“ Lehrer: „Dann versuch mal so: 30 plus 10 ist 40, dann denkst du an die Zahlengeschwister: 8 plus 2 ist 10, macht zusammen 50, und dann bleiben von der 7 noch 5 übrig, macht zusammen 55.“

(schreibt auf ein Blatt):

$$38+17=$$

$$30+10=40$$

$$8+2=10$$

$$40+10=50$$

$$50+5=55$$

Anna: „Das ist total unfair! Dann muss ich ja fünf Aufgaben rechnen!“ Am Tisch entsteht Unruhe, die Kinder kichern. Der Lehrer macht den lachenden Kindern ein Zeichen, still weiterzuarbeiten und erklärt gleichzeitig weiter. Anschließend geht er zu den anderen Gruppen.

Nach 15 Minuten stellt sich der Lehrer vor die Klasse und wartet. Die Kinder reden noch. Lehrer: „Die 15 Minuten sind um, ich warte! (...) Ich möchte jetzt, dass jede Gruppe mal einen Rechenweg vorstellt. Wer fängt an?“ Die Kinder am vordersten Tisch schweigen zunächst und schubsen sich dann gegenseitig an. Lehrer: „So, wer trägt vor?“ Jan: „Okay, dann stell ich meinen Weg vor.“ Lucas murmelt leise: „Der will immer bestimmen.“ (...)

Items Vignette 4

<i>Raum</i>	Die räumlichen Arbeitsbedingungen sind für alle Schülerinnen und Schüler gut.
<i>Struktur und Übergänge</i>	Es kommt zu unnötigen Verzögerungen im Unterrichtsverlauf.
<i>Klarheit</i>	Den Schülerinnen und Schülern ist das Ziel der Stunde unklar.
<i>Ablenker</i>	Den Schülerinnen und Schülern mangelt es an Lernmotivation.
<i>Feedback</i>	Der Lehrer kann Unterrichtsgeschehen durch passendes Feedback steuern.
<i>Ablenker</i>	Der Lehrer handelt ergebnisorientiert.
<i>Regeln und Routinen</i>	Den Schülerinnen und Schülern fehlen Routinen, was sie wann und wie zu tun haben.
<i>Allgegenwärtigkeit</i>	Der Lehrer kann verschiedene Unterrichtsprozesse gleichzeitig steuern.
<i>Ablenker</i>	Die Schülerinnen und Schüler haben genügend Materialien zur Verfügung.

11.1.3 Überarbeitung der Instruktion und der Skala

Im zweiten Entwicklungsschritt und der Expertinnen- und Expertendiskussion sowie in den Gruppendiskussionen der Masterstudierenden wurde deutlich, dass Unsicherheit bezüglich der Eindeutigkeit von Beurteilungen herrschte. So wurde bemängelt, dass sich von einer einzelnen Unterrichtssituation nicht darauf schließen lasse, ob Verhalten und Fähigkeiten der Lehrkraft danach beurteilt werden könnten. Auch wenn nach den Fähigkeiten der Lehrkraft nicht gefragt wurde und nur die momentanen Auswirkungen der Interaktionen beurteilt werden sollten, entstand bei den Masterstudierenden der Eindruck, die Lehrkraft persönlich zu beurteilen. Da die beschriebenen Situationen außerhalb eines Zusammenhangs stünden, sei eine Beurteilung nicht oder nur unter Vorbehalt möglich. Zudem ließe ein kurzer Ausschnitt keine eindeutige Beurteilung zu, ob sich die beschriebene Situation tatsächlich als Barriere auswirke. Die in den Vignetten als Barrieren gedachten Stellen seien eher als *potenzielle* Störungen zu verstehen, die im weiteren Verlauf zu Lernhindernissen führen *könnten*.

Die Instruktion wurde deshalb folgendermaßen umformuliert: *Die folgenden Aspekte könnten maßgeblich zum Gelingen oder Misslingen der Lernsituation beitragen. Bitte beurteilen Sie,*

für wie wahrscheinlich Sie die folgenden Aussagen aufgrund der beschriebenen Lernsituation halten.

Die Skalen wurden nicht mehr mit *trifft voll und ganz zu* etc. betitelt, sondern mit *trifft sehr wahrscheinlich zu*, *trifft eher wahrscheinlich zu*, *trifft eher unwahrscheinlich zu* und *trifft sehr unwahrscheinlich zu*.

11.2 Fragestellungen und Hypothesen

Für die so konstruierten Vignetten ergeben sich die folgenden Fragestellungen und Hypothesen:

Mittlere Schwierigkeitsgrade

Frage 1: Ziel des Konstruktionsprozesses mit der Validierung der einzelnen Beispiele ist es, dass Merkmale gut zuordenbar, aber nicht zu leicht zu beurteilen sind. Es stellt sich damit die Frage, ob auf diese Weise bessere Schwierigkeitsgrade als in den ersten Vignetten erzeugt werden können.

Hypothese 1: Es zeigen sich für die Items und die Vignetten mittlere Schwierigkeitsgrade und angemessene Verteilungen und damit bessere Werte als für die vorher analysierten Vignetten und Items.

Trennschärfe und Reliabilität der Items

Frage 2: Die neuen vier Vignetten wurden nach dem gleichen Konstruktionsprinzip verfasst und enthalten jeweils die gleichen Merkmale mit dem Ziel, durch diese Vervielfachung die Reliabilität und die Trennschärfen der Items zu erhöhen. Außerdem sollte durch die Validierung der Vignetten vermieden werden, dass Items durch ein zufälliges Ankreuzen beurteilt werden. Die Items sollen die enthaltenen Merkmale gezielt abfragen. Konnten auf diese Weise trennscharfe und reliable Items erzeugt werden?

Hypothese 2: Die Items sind trennscharf und reliabel.

Konstruktvalidität

Frage 3: Die Konstruktionsprinzipien der Vignetten sollten dazu führen, dass situationsübergreifend gelingende und misslingende Merkmale der KF und LU erkannt und differenziert beurteilt werden. Zeigt sich dieses Konstruktionsprinzip in der Struktur der Daten?

Hypothese 3: Lernbarrieren und gelingende Situationen werden differenziert voneinander beurteilt. Eine hohe Fähigkeit im Erkennen zeigt sich sowohl im Erkennen von Barrieren als auch im Erkennen gelingender Situationen.

Kriteriumsvalidität

Frage 4: Mit der Erhöhung der mittleren Schwierigkeitsgrade sollen sich unterschiedliche Fähigkeitsausprägungen darstellen lassen. Lassen sich Expertiseunterschiede zwischen Studierenden verschiedener Studiengänge und Fachsemester sowie zwischen Studierenden und Lehrkräften zeigen?

Hypothese 4: Es zeigen sich Expertiseunterschiede zugunsten der Studierenden der Sonderpädagogik, der Lehrkräfte und der höheren Fachsemester.

Kreuzvalidität

Frage 5: Bisherige Erkenntnisse beruhen auf der Erprobung mit Studierenden der Universität Hamburg. Daten aus anderen Stichproben sollen zeigen, ob sich diese generalisieren lassen und so die Aussagekraft der Ergebnisse verbessern. Lässt sich zeigen, dass die Datenstruktur und Befunde zu Gruppenunterschieden nicht zufällig oder stichprobengebunden sind, sondern sich replizieren lassen?

Hypothese 5.1: Die Datenstrukturen gleichen sich in den verschiedenen Stichproben.

Hypothese 5.2: Gruppenunterschiede lassen sich in verschiedenen Stichproben replizieren.

11.3 Stichprobe

Der neue Satz von vier Textvignetten wurde an acht Universitäten erprobt. Es wurden mit den gleichen Annahmen zu Expertiseunterschieden wie in ersten Entwicklungsschritten Studierende der Sonderpädagogik und anderer Lehrämter befragt sowie Studierende verschiedener Fachsemester. Die Auswahl der Seminare richtete sich erneut nach den Durchführungsmöglichkeiten seitens der Lehrenden, so dass nicht von einem bedeutenden Stichprobeneffekt ausgegangen werden muss. Zudem erfolgte die Befragung zu Semesterbeginn, so dass die Studierenden noch keinen Input bzgl. der jeweiligen Seminarthemen hatten und von Seminareffekten ebenfalls nicht auszugehen ist. Um angemessene Gruppengrößen zu erzielen, wurden die Fachsemester zu Gruppen zusammengefasst. Zusätzlich wurde eine Gruppe von Lehrkräften im Rahmen einer Fortbildung befragt (siehe Tabelle 22).

Tabelle 22*Häufigkeiten für die Stichprobe der dritten Datenerhebung*

Universität	Studiengang	Fachsemester				n
		1. und 2. Bachelor	4. bis 6. Bachelor	Master	Fortbildung	
Hamburg	Sonderpädagogik	27	65	43	-	135
	Regelschullehrämter	72	4	11	-	87
Wuppertal	Sonderpädagogik	-	38	2	-	40
	Regelschullehrämter	-	-	-	-	0
Rostock	Sonderpädagogik	-	-	-	-	0
	Regelschullehrämter	1	9	22	-	32
Potsdam	Sonderpädagogik	-	-	-	18	18
	Regelschullehrämter	-	-	-	-	0
Flensburg	Sonderpädagogik	3	34	-	-	37
	Regelschullehrämter	-	-	-	-	0
Bremen	Sonderpädagogik	-	-	-	-	0
	Regelschullehrämter	4	22	10	-	36
Gesamt		107	172	88	18	385

11.4 Design und Durchführung

Die Beurteilungsaufgaben wurden in Hamburg, Wuppertal und in der Fortbildung in Potsdam als Papier-und-Bleistift-Test ausgefüllt und an den anderen Universitäten als Online-Befragung. Um möglichst ähnliche Bedingungen zu schaffen, bekamen alle Teilnehmenden vorab die folgende gleiche Information:

Liebe Studierende,

im Rahmen des Projekts „Professionelles Lehrerhandeln zur Förderung fachlichen Lernens unter sich verändernden gesellschaftlichen Bedingungen (ProfaLe)“ werden derzeit an der Universität Hamburg verschiedene Verfahren entwickelt, die die Beurteilungsfähigkeit von Studierenden erfassen sollen. Dies soll zur Evaluation und somit zur Verbesserung der Lehre beitragen. Um die Verfahren erproben und verbessern zu können, sind wir auf Ihre Mithilfe angewiesen! Aus diesem Grund möchten wir Sie bitten, an dieser Befragung teilzunehmen, die 4 Aufgaben enthält.

Ausfüll-Hinweis:

Im Folgenden sind Ausschnitte aus verschiedenen Lehr-Lernsituationen beschrieben. Lesen Sie die Beispiele aufmerksam durch und beurteilen Sie bitte im Anschluss an jeden Text die dazu formulierten Aussagen. Beziehen Sie sich bei der Beurteilung der Aussagen allein auf das Beispiel.

Im Papier-und-Bleistift-Test gab es außerdem noch den zusätzlichen Hinweis, Kreuzchen nur *in* die Kästchen zu setzen und nicht dazwischen. Anschließend bearbeiteten alle Probandinnen und Probanden alle vier Textvignetten in zufälliger Reihenfolge. Da die Bearbeitungszeit in der Online-Version nicht limitiert war, wurde die Bearbeitungszeit des Papier-und-Bleistift-Test dieses Mal auch nicht beschränkt. Dies sollte zudem unterstützen, dass die Situation nicht pauschal beurteilt wird.

Zur Auswertung wurden wie zuvor alle Antworten in SPSS übertragen und in Werte von 1 bis 4 übersetzt. Diese Werte wurden so rekodiert, dass sich ein hoher Wert (3 und 4) immer auf das Erkennen der Barriere bzw. der gelingenden Situation bezog.

Verfahren zur Ermittlung der Testgüte werden zunächst nur mit der Kohorte der Universität Hamburg berechnet. Kohorten anderer Universitäten werden zur Überprüfung der Kreuzvalidität genutzt.

Zur Itemanalyse werden mittlere Schwierigkeitsgrade und die Verteilungen betrachtet sowie eine Reliabilitätsanalyse durchgeführt und Trennschärfen dargestellt. Die Struktur der Daten wird durch eine explorative Faktorenanalyse dargestellt. Zur Ermittlung von Testscores werden die Ergebnisse der Faktorenanalyse genutzt und gewichtete Mittelwerte für die Items berechnet, indem die Itemrohwerte mit der Höhe der jeweiligen Faktorladung des Items multipliziert werden. Der Mittelwert aller gewichteten Items (je nach Ergebnis für die Gesamtskala oder pro Faktor) entspricht dem Testscore. Die Kriteriumsvalidität wird ermittelt, indem die Gruppen anhand dieses Testscores mittels non-parametrischer Verfahren verglichen werden. Zur Überprüfung der Kreuzvalidität wird überprüft, ob sich die Datenstruktur sowie die Kriteriumsvalidität in verschiedenen Stichproben mit den gleichen Methoden replizieren lassen.

11.5 Ergebnisse und Interpretation

Mit der Annahme, dass die vier Vignetten als Subtests die gleiche Fähigkeit erfassen, werden die Analysen sowohl vignettenweise dargestellt, um zu prüfen, ob die einzelnen Vignetten die Testgütekriterien erfüllen, als auch insgesamt, um zu prüfen, ob ein Gesamtwert genutzt werden kann. Alle Berechnungen werden zunächst mit der Teilstichprobe der Studierenden der Universität Hamburg vorgenommen.

11.5.1 Ergebnisse zu mittleren Schwierigkeitsgraden

Frage 1: Ziel des Konstruktionsprozesses mit der Validierung der einzelnen Beispiele war es, dass die Merkmale gut zuzuordnen, aber nicht zu leicht zu beurteilen sind. Es stellt sich damit die Frage, ob auf diese Weise bessere Schwierigkeitsgrade als in den ersten Vignetten erzeugt werden können.

Hypothese 1: Es zeigen sich für die Items und die Vignetten mittlere Schwierigkeitsgrade und angemessene Verteilungen und damit bessere Werte als für die vorher analysierten Vignetten und Items.

In Tabelle 23 ist die deskriptive Statistik für Items zu Lernbarrieren und gelingenden Situationen dargestellt. Bei der Beurteilung wird in allen Fällen die Skala von 1–4 ausgenutzt, wobei die Extremwerte seltener gewählt werden (siehe Anhang D Tabelle D5). Die Itemmittelwerte liegen zwischen 1.94 ($SD = 0.75$) für das Item zu Klarheit in Vignette 1 (gelingend) und 3.18 ($SD = 0.77$) für das Item zu Feedback in Vignette 1 (Barriere). Items zu Barrieren werden insgesamt etwas besser erkannt ($M = 2.70$, $SD = 0.32$) als Items zu gelingenden Situationen ($M = 2.47$, $SD = 0.45$). Vignette 3 ist am schwierigsten zu beurteilen ($M = 2.51$, $SD = 0.67$) und Vignette 1 am einfachsten ($M = 2.85$, $SD = 0.35$) (siehe Tabelle 23).

Tabelle 23

Deskriptive Statistik zu Items zu Barrieren und Items zu gelingenden Situationen

Items und Vignetten	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Var</i>
V1 Raum Barriere	219	1	4	2.86	.76	.58
V1 Allgegenwärtigkeit Barriere	220	1	4	3.05	.90	.82
V1 Regeln und Routinen Barriere	221	1	4	3.16	.76	.57
V1 Übergänge Barriere	218	1	4	2.92	.68	.46
V1 Feedback Barriere	220	1	4	3.18	.77	.60
V1 Klarheit gelingend	222	1	4	1.94	.75	.60
V2 Feedback Barriere	218	1	4	2.80	.79	.62
V2 Übergänge gelingend	222	1	4	2.71	.72	.52
V2 Raum Barriere	221	1	4	2.43	.82	.67
V2 Allgegenwärtigkeit Barriere	222	1	4	2.79	.80	.64
V2 Klarheit Barriere	219	1	4	2.60	.76	.57
V2 Regeln und Routinen gelingend	222	1	4	2.60	.94	.88
V3 Klarheit Barriere	221	1	4	2.24	.94	.88
V3 Feedback gelingend	221	1	4	2.59	.86	.74

V3 Regeln und Routinen Barriere	221	1	4	2.51	.87	.75
V3 Übergänge Barriere	217	1	4	2.66	.88	.78
V3 Raum gelingend	218	1	4	2.43	.87	.75
V3 Allgegenwärtigkeit Barriere	219	1	4	2.65	.90	.81
V4 Raum Barriere	217	1	4	2.49	.70	.48
V4 Struktur Barriere	220	1	4	2.43	.75	.56
V4 Klarheit Barriere	220	1	4	2.47	.82	.68
V4 Feedback Barriere	218	1	4	2.78	.79	.62
V4 Regeln und Routinen Barriere	219	1	4	2.57	.69	.48
V4 Allgegenwärtigkeit	218	1	4	2.53	.85	.72
V1 gesamt	222	1.80	4.00	2.85	.35	.12
V2 gesamt	223	1.33	3.50	2.65	.34	.12
V3 gesamt	221	1.83	3.50	2.51	.36	.13
V4 gesamt	221	1.50	3.50	2.54	.37	.14
Items zu Barrieren gesamt	223	1.89	3.78	2.70	.32	.10
Items zu gelingenden Situationen gesamt	223	1.33	3.67	2.47	.45	.20

Anmerkung. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4. Niedrigste und höchste Mittelwerte sind fett markiert.

11.5.2 Interpretation zu mittleren Schwierigkeitsgraden

Die Mittelwerte für die Items liegen in der dritten Erhebung insgesamt näher an der Skalenmitte von 2.5 und es zeigt sich keine rechtssteile Verteilung wie in den vorherigen beiden Datenerhebungen. Schwierigkeiten verteilen sich über die Vignetten und Items besser und es zeigt sich eine zufriedenstellende Streuung der Items. Die Mittelwerte für die Vignetten zeigen, dass diese enger um den Mittelwert streuen, da Mittelwerte für Items zu gelingenden Situationen niedriger und für Items zu Barrieren höher sind und sich Gesamtwerte somit mehr der Skalenmitte annähern. Wie bei der Konstruktion beabsichtigt, sind die Barrieren und gelingenden Situationen also gut, aber nicht zu offensichtlich zu erkennen, so dass die Ergebnisse die Hypothese 1 stützen.

Für das Erkennen der verschiedenen Barrieren zeigt sich, dass sich die Schwierigkeiten in den neuen Vignetten anders verteilen. Die Barriere im *Feedback* wird wie in der ersten Studie – obwohl keine Kontextinformationen zu Kindern gegeben werden – in allen Vignetten sicher erkannt. Anders als in den vorherigen Entwicklungsschritten wurden auch die Hinweise auf mangelnde *Regeln und Routinen* gut erkannt, ebenso wie die *Allgegenwärtigkeit*. Die *Übergänge* wurden in Abhängigkeit von der Vignette erkannt: Während sie in Vignette 3 am deutlichsten

gesehen wurden, waren sie in Vignette 4 am schwierigsten zu erkennen. Im Gegensatz zu den Vignetten aus der ersten und zweiten Datenerhebung gehörten die Barrieren bzgl. *Klarheit* und *Raumbedingungen* in allen neuen Vignetten zu den Situationen, die weniger sicher erkannt wurden. Hier lässt sich vermuten, dass das Erkennen der Barrieren weniger von der Zugehörigkeit zu einem Merkmal und stärker von der konkreten Darstellung in den jeweiligen Vignetten abhängig ist.

11.5.3 Ergebnisse zu Trennschärfen und Reliabilität

Frage 2: Konnten durch die Konstruktionskriterien trennscharfe und reliable Items erzeugt werden?

Hypothese 2: Die Items sind trennscharf und reliabel.

Die Trennschärfen und Reliabilitäten werden zunächst vignettenweise und dann insgesamt dargestellt. Dabei zeigt sich in allen Vignetten, dass jeweils die Items zu gelingenden Situationen eine negative Item-Skala-Korrelation aufweisen (siehe Anhang D Tabellen D6 bis D9). Diese Items werden deshalb zunächst ausgeschlossen und Trennschärfen sowie die Reliabilität für die verbleibenden Items zu Barrieren berechnet. In Vignette 1 zeigen zwei Items eine Trennschärfe von $r_{j(t-j)} > .25$ und kein Item eine Trennschärfe von $r_{j(t-j)} = .50$ oder höher (vgl. Lienert & Raatz, 2011) (siehe Tabelle 24). Cronbachs Alpha liegt für Vignette 1 bei .43.

Tabelle 24

Trennschärfen der Items zu Barrieren in Vignette 1

Item	Skalenmittelwert, wenn Item weglassen	Skalenvarianz, wenn Item weglassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weglassen
Raum	12.30	3.50	.17	.42
Allgegenwärtigkeit	12.13	2.97	.24	.37
Regeln und Routinen	12.03	3.22	.28	.34
Übergänge	12.26	3.21	.35	.29
Feedback	11.97	3.67	.10	.46

In Vignette 2 liegen die Trennschärfen für die Items zwischen Trennschärfe von $r_{j(t-j)} = .33$ und $r_{j(t-j)} = .60$ (siehe Tabelle 25). Die Skala weist ein Cronbachs Alpha von .68 auf.

Tabelle 25*Trennschärfen der Items zu Barrieren in Vignette 2*

Item	Skalenmittelwert, wenn Item wegge- lassen	Skalenvarianz, wenn Item wegge- lassen	Korrigierte Item- Skala-Korrelation	Cronbachs Alpha, wenn Item wegge- lassen
Feedback	7.83	3.10	.50	.59
Raum	8.21	3.43	.33	.70
Allgegenwärtigkeit	7.82	2.90	.56	.53
Klarheit	8.02	3.33	.45	.63

Die Trennschärfen in Vignette 3 liegen zwischen $r_{j(t-j)} = .44$ und $r_{j(t-j)} = .67$ (siehe Tabelle 26) und das Cronbachs Alpha bei .75.

Tabelle 26*Trennschärfen der Items zu Barrieren in Vignette 3*

Item	Skalenmittelwert, wenn Item wegge- lassen	Skalenvarianz, wenn Item weg- gelassen	Korrigierte I- tem-Skala-Kor- relation	Cronbachs Al- pha, wenn Item weggelassen
Klarheit	7.82	4.28	.56	.67
Regeln und Routinen	7.55	4.92	.44	.74
Übergänge	7.40	4.71	.49	.71
Allgegenwärtigkeit	7.40	4.12	.67	.61

In der Vignette 4 weisen die Items eine Trennschärfe von $r_{j(t-j)} = .33$ bis $r_{j(t-j)} = .52$ auf (siehe Tabelle 27) und Cronbachs Alpha liegt bei .70.

Tabelle 27*Trennschärfen der Items zu Barrieren in Vignette 4*

Item	Skalenmittelwert, wenn Item wegge- lassen	Skalenvarianz, wenn Item weggelassen	Korrigierte I- tem-Skala-Kor- relation	Cronbachs Alpha, wenn Item wegge- lassen
Raum	10.26	4.96	.33	.70
Struktur	10.32	4.36	.50	.63
Klarheit	10.27	4.27	.44	.66
Feedback	9.97	4.20	.50	.63
Regeln und Routinen	10.18	4.45	.52	.62

Betrachtet man die Vignetten als zusammengehörig (unter Ausschluss der Items zu gelingenden Situationen sowie Vignette 1 wegen zu geringer Trennschärfen und Reliabilität), so zeigen sich für die Gesamtskala Trennschärfen zwischen $r_{j(t-j)} = .18$ und $r_{j(t-j)} = .49$, von 13 Items zeigen 8 Trennschärfen $r_{j(t-j)} > .25$ (siehe Tabelle 28). Die Reliabilität liegt für die Skala mit 13 Items bei $\alpha = .68$.

Tabelle 28

Trennschärfen der Items zu Barrieren für die Gesamtskala der Vignetten 2, 3 und 4

Item	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
V2 Feedback	30.76	21.52	.18	.69
V2 Raum	31.12	21.27	.21	.68
V2 Allgegenwärtigkeit	30.74	19.88	.42	.65
V2 Klarheit	30.95	21.39	.23	.68
V3 Klarheit	31.28	19.51	.38	.66
V3 Regeln und Routinen	31.03	20.93	.23	.68
V3 Übergänge	30.87	20.36	.30	.67
V3 Allgegenwärtigkeit	30.87	18.97	.48	.64
V4 Raum	31.05	21.24	.29	.67
V4 Struktur	31.11	21.05	.28	.67
V4 Klarheit	31.06	21.02	.24	.68
V4 Feedback	30.76	19.43	.49	.64
V4 Regeln und Routinen	30.98	21.04	.31	.67

Anmerkung. V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4

11.5.4 Interpretation der Trennschärfen und Reliabilität

Bei der Betrachtung der Trennschärfen und Reliabilitäten fällt zunächst auf, dass die Items zu gelingenden Situationen negativ mit den übrigen Items korrelieren. Items zu Barrieren und gelingenden Situationen bilden demnach nicht eine Skala. Unterschiede in den Trennschärfen und Reliabilitäten der Items zu Barrieren zeigen sich in Abhängigkeit davon, ob diese Kennwerte für die einzelnen Vignetten oder über alle Vignetten hinweg betrachtet werden. Nach Ausschluss der Items zu gelingenden Situationen zeigt Vignette 1 die geringste Reliabilität und die schlechtesten Item-Kennwerte. Hingegen sind Reliabilitäten und Trennschärfen für die anderen Vignetten akzeptabel bis gut (Lienert & Raatz, 2011; Streiner, 2003). Während die Skalen innerhalb der Vignetten 2, 3 und 4 konsistent sind, sind die Korrelationen zwischen den Vignetten

jeweils geringer. Dass die Fähigkeit zum Erkennen von Lernbarrieren anhand verschiedener Merkmale und unterschiedlicher Vignetten erfasst wird, bringt es in diesem Fall mit sich, dass hohe Trennschärfen für die Gesamtskala nicht erreicht werden (Streiner, 2003). Die Hypothese 2 wird insofern nur von den Ergebnissen für die Vignetten 2, 3 und 4 und nur für Items zu Barrieren gestützt.

11.5.5 Ergebnisse zur Konstruktvalidität

Frage 3: Die Konstruktionsprinzipien der Vignetten sollten dazu führen, dass situationsübergreifend gelingende und misslingende Merkmale der KF und LU erkannt und differenziert beurteilt werden. Zeigt sich dieses Konstruktionsprinzip in der Struktur der Daten?

Hypothese 3: Lernbarrieren und gelingende Situationen werden differenziert voneinander beurteilt. Eine hohe Fähigkeit im Erkennen zeigt sich sowohl im Erkennen von Barrieren als auch im Erkennen gelingender Situationen.

Bereits in der Reliabilitätsanalyse zeigt sich, dass Items zu gelingenden Situationen und zu Barrieren nicht konsistent beantwortet werden. Um dies sichtbar zu machen, wird eine Faktorenanalyse unter Berücksichtigung aller Items der Vignetten 2, 3 und 4 zunächst für die Hamburger Stichprobe dargestellt. Vignette 1 wird aufgrund der schlechteren Itemkennwerte ausgeschlossen. Es wird eine Interkorrelation der Faktoren angenommen, so dass eine Promax-Rotation genutzt wird. Die Anzahl der Faktoren wird auf Faktoren mit mehr als drei Ladungen über .3 reduziert (Lienert & Raatz, 2011). Das Maß der Stichprobeneignung nach dem Kaiser-Meyer-Olkin-Kriterium (.78) und die Signifikanz nach Bartlett ($p < .001$) zeigen, dass sich die Daten für eine Faktorenanalyse eignen.

Es entstehen drei Faktoren, wobei die Vignetten Ladungen auf jeweils einem Faktor zeigen. Die Items zu gelingenden Situationen korrelieren in allen Vignetten negativ mit den übrigen Items (siehe Tabelle 29). Die Faktoren erklären rund 22 %, 14 % und 12 % der Gesamtvarianz (siehe Anhang D Tabelle D10) und korrelieren mit $r = .16$ bis $r = .24$ (siehe Anhang D Tabelle D11).

Tabelle 29*Ergebnis der Faktorenanalyse*

Items	Faktorladungen		
	1	2	3
Faktor 1: Vignette 3			
V3 Klarheit Barriere	.72		
V3 Feedback gelingend	-.68		
V3 Regeln und Routinen Barriere	.65		
V3 Übergänge Barriere	.73		
V3 Raum gelingend	-.60		
V3 Allgegenwärtigkeit Barriere	.82		
Faktor 2: Vignette 4			
V4 Raum Barriere		.54	
V4 Übergänge Barriere		.77	
V4 Klarheit Barriere		.62	
V4 Feedback Barriere		.67	
V4 Regeln und Routinen Barriere		.75	
V4 Allgegenwärtigkeit gelingend		-.66	
Faktor 3: Vignette 2			
V2 Feedback Barriere			.69
V2 Übergänge gelingend			-.70
V2 Raum Barriere			.46
V2 Allgegenwärtigkeit Barriere			.76
V2 Klarheit Barriere			.76
V2 Regeln und Routinen gelingend			-.62

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung. V3 = Vignette 3, V4 = Vignette 4, V2 = Vignette 2.

11.5.6 Interpretation der Konstruktvalidität

Die Reliabilitätsanalysen geben bereits Hinweise darauf, dass die Barrieren in Abhängigkeit von den Vignetten beurteilt werden. Die Datenstruktur bestätigt Vignetteneinflüsse: Es zeigt sich, dass die Beurteilung der Items jeweils von der spezifischen Situation abhängig ist und sich sogenannte Testlets bilden, also Itemgruppierungen innerhalb des Instruments (Gold & Holodynski, 2017; Meschede, 2014). Es zeigt sich zudem, dass die Items zu gleichen Merkmalen jeweils ähnliche Ladungsmuster innerhalb der Faktoren aufweisen: Während beispielsweise die Items

zu *Übergängen* immer hohe Ladungen aufweisen, tragen die Items zu *Raum* jeweils die geringsten Ladungen. Die Faktoren korrelieren jedoch geringfügig, was bedeutet, dass die Vignetten relativ unabhängig voneinander beurteilt werden. Merkmale der KF und LU lassen sich nicht unterscheiden.

Einerseits spricht das dafür, dass das Verfahren kein in sich konsistentes Konstrukt abbildet und mit den vier Vignetten unterschiedliche Fähigkeiten erfasst werden – nämlich das Erkennen der Barrieren in den jeweiligen Vignetten. Andererseits entspricht die Situationsspezifität der Realität von Unterricht, so dass es sinnvoll erscheint, die Fähigkeit zum Erkennen von Lernbarrieren anhand unterschiedlicher Subskalen zu erfassen, die solche unterschiedlichen Situationen repräsentieren. In der pädagogisch-psychologischen Diagnostik sind solche Testlet-basierten Verfahren verbreitet, um verschiedene Dimensionen der gleichen Fähigkeit zu erfassen (Eckes, 2015). Die geringen Korrelationen zwischen den Faktoren weisen aber auf einen sehr starken Einfluss der Vignetten und weniger auf einen Einfluss einer übergeordneten Fähigkeit hin.

Bezüglich der Beurteilungen scheint außerdem ein positiver oder negativer Gesamteindruck der geschilderten Situationen einen starken Einfluss zu haben. Die negativen Korrelationen innerhalb der Vignetten zeigen dies. Items werden also insgesamt eher als einheitlich positiv oder negativ beurteilt, unabhängig davon, ob sie sich auf Barrieren oder gelingende Situationen beziehen.

Um dies zu verdeutlichen, werden Mittelwertvergleiche für Teilstichproben dargestellt. Die Teilstichproben unterscheiden sich darin, welche Mittelwerte diese für die Items zu Barrieren erreichen. Maßgeblich für die Cut-Off-Werte ($M = 2.30$, $M = 2.50$, $M = 2.80$ und $M = 3.00$) sind angemessene Stichprobengrößen. Verglichen werden die Mittelwerte zu Items zu Barrieren und Items zu gelingenden Situationen. Es zeigt sich, dass diejenigen, die Barrieren im Mittel besonders zutreffend beurteilen ($M > 3.00$), gelingende Situationen im Mittel nicht zutreffend beurteilen ($M = 1.92$). Diese Personengruppe beurteilt also alles tendenziell eher als nicht gelingend. Studierende, die Barrieren weniger zutreffend beurteilen ($M < 2.30$), erreichen dagegen deutlich höhere Mittelwerte für die Items zu gelingenden Situationen ($M = 3.02$) (siehe Tabelle 30). Es scheinen also Urteilertendenzen zu insgesamt positiven oder negativen Urteilen eine Rolle zu spielen.

Tabelle 30

Vergleich der Mittelwerte von Items zu Barrieren und zu gelingenden Situationen für Teilstichproben mit unterschiedlich zutreffenden Beurteilungen von Barrieren

Anzahl Teilstichprobe	Items zu Barrieren	Items zu gelingenden Situationen
<i>n</i>	<i>M</i>	<i>M</i>
18	> 3.00	1.92
54	> 2.80	2.19
88	< 2.50	2.90
48	< 2.30	3.02

Mit der Darstellung sowohl von Barrieren als auch gelingenden Situationen in den Vignetten sollte eine genaue Beobachtung angeregt werden bzw. sichtbar werden, ob differenziert beobachtet und beurteilt wird. Die Ergebnisse deuten aber darauf hin, dass eher Gesamteindrücke die Beurteilungen beeinflussen und diese jeweils durch die Vignetten verursacht werden. Die Hypothese 3 muss deshalb verworfen werden. Für die Weiterentwicklung des Verfahrens bedeutet das, dass Überarbeitungskriterien überlegt werden müssen, die eine differenzierte Beurteilung stärker anregen. Außerdem stellt sich die Frage, welche Faktoren zu solchen negativen oder positiven Gesamteindrücken führen.

11.5.7 Ergebnisse zur Kriteriumsvalidität

Frage 4: Lassen sich Expertiseunterschiede zwischen Studierenden verschiedener Studiengänge und Fachsemester sowie zwischen Studierenden und Lehrkräften zeigen?

Hypothese 4: Es zeigen sich Expertiseunterschiede zugunsten der Studierenden der Sonderpädagogik, der Lehrkräfte und der höheren Fachsemester.

Die Stichproben der Universität Hamburg und der Lehrkräfte ermöglichen Vergleiche zwischen Studierenden der Sonderpädagogik und der Allgemeinpädagogik im 1. und 2. Bachelorsemester sowie Vergleiche zwischen den Fachsemestergruppen der Sonderpädagogikstudierenden. Außerdem können Studierende mit Lehrkräften verglichen werden.

Die bisherigen Ergebnisse legen nahe, Werte für die einzelnen Vignetten zu bilden. Auch wenn Items zu Barrieren und gelingenden Situationen einer Vignette jeweils auf dem gleichen Faktor laden, sollten diese aufgrund der negativen Korrelationen auch getrennt betrachtet werden. Mit der Annahme, dass eine *differenzierte* Beurteilung gelingender und nicht gelingender

Situationen für eine höhere Fähigkeitsausprägung spricht als ein einseitiges Pauschalurteil, erscheinen Gesamtmittelwerte für die Vignetten vor dem Hintergrund der bisherigen Ergebnisse nicht ausreichend aussagekräftig.

Tabelle 30 legt nahe, dass es unterschiedliche Urteilertypen gibt, die jeweils die einzelnen Vignetten insgesamt als eher negativ oder positiv beurteilen. Diese sollen sichtbar gemacht werden, indem nicht nur Faktorenwerte berechnet werden, sondern gewichtete Mittelwerte auch differenziert für Items zu gelingenden Situationen und zu Barrieren betrachtet werden. Im dritten Entwicklungsschritt werden erstmalig in der Untersuchung Items zu gelingenden Situationen erprobt. Diese sind im Vergleich zu Items zu Barrieren aber unterrepräsentiert und kommen pro Vignette jeweils nur ein- oder zweimal vor, so dass Mittelwerte zu gelingenden Situationen für die Vignetten auf nur ein oder zwei Items beruhen. Dennoch sollen diese Informationen zur weiteren Entwicklung genutzt werden. Als Signifikanzniveau wird die Grenze von 10 % festgesetzt.

Für die Studierenden im 1. und 2. Bachelorsemester ergeben sich im Mann-Whitney-U-Test keine signifikanten Unterschiede zwischen den Studiengängen (siehe Anhang D Tabelle D12). Auch für die Fachsemestergruppen im Studiengang Sonderpädagogik ergeben sich, geprüft mit dem Kruskal-Wallis-Test, keine signifikanten Unterschiede (siehe Anhang D Tabelle D13).

Der Vergleich zwischen Studierenden und Lehrkräften zeigt, geprüft mit dem Mann-Whitney-U-Test, keine signifikanten Unterschiede bzgl. der Faktorenwerte. Differenziert man aber nach Items zu Barrieren und gelingenden Situationen, so zeigt sich, dass Lehrkräfte signifikant höhere Werte bei der Beurteilung von Barrieren in Vignette 2 und 4 erzielen und signifikant niedrigere Werte bei der Beurteilung von gelingenden Situationen in allen Vignetten (siehe Anhang D Tabelle D14). Die signifikanten Ergebnisse im Einzelnen:

Vignette 2 Barrieren: Lehrkräfte unterscheiden sich signifikant von Studierenden ($z = 2.811$, $p = .005$). Es handelt sich mit $r = .19$ um einen schwachen Effekt nach Cohen (1992) (siehe Anhang D Tabelle D15).

Vignette 2 gelingende Situationen: Lehrkräfte unterscheiden sich signifikant von Studierenden ($z = -2.922$, $p = .003$). Es handelt sich mit $r = .20$ um einen schwachen Effekt nach Cohen (1992) (siehe Anhang D Tabelle D16).

Vignette 3 gelingende Situationen: Lehrkräfte unterscheiden sich signifikant von Studierenden ($z = -3.085$, $p = .002$). Es handelt sich mit $r = .21$ um einen schwachen Effekt nach Cohen (1992) (siehe Anhang D Tabelle D17).

Vignette 4 Barrieren: Lehrkräfte unterscheiden sich signifikant von Studierenden ($z = 1.735$, $p = .083$). Es handelt sich mit $r = .12$ um einen schwachen Effekt nach Cohen (1992) (siehe Anhang D Tabelle D18).

Vignette 4 gelingende Situationen: Lehrkräfte unterscheiden sich signifikant von Studierenden ($z = -1.699$, $p = .089$). Es handelt sich mit $r = .12$ um einen schwachen Effekt nach Cohen (1992) (siehe Anhang D Tabelle D19).

11.5.8 Interpretation der Kriteriumsvalidität

Die Stichproben wurden mit der Annahme gewählt, dass sich Unterschiede durch Studienfach und Studiendauer ergeben bzw. Expertise durch Berufserfahrung. Befunde der ersten beiden Datenerhebungen lassen vermuten, dass eine unterschiedliche Fokussierung bereits zu Studienbeginn einen Unterschied zwischen Studierenden der Sonder- und der Allgemeinpädagogik bedingen und dass Urteile mit fortschreitender Studiendauer sicherer werden. Die Daten der dritten Datenerhebung stützen diese Annahme jedoch nicht.

Die signifikant höheren Werte für Items zu Barrieren sowie signifikant niedrigeren Werte für Items zu gelingenden Situationen in der Gruppe der Lehrkräfte im Vergleich zu Studierenden zeigen jedoch Unterschiede in den Beurteilungen. Bei diesen Unterschieden kann jedoch eher nicht von einer Expertise in genauem Beobachten ausgegangen werden. Vielmehr legt der Unterschied der Werte nahe, dass in dieser Datenerhebung eine Fokussierung auf und stärkere Gewichtung von Störungen mit vorhandener Praxiserfahrung einhergeht. Einerseits ist das Ziel dieser Untersuchung, die Fokussierung auf misslingende Unterrichtsprozesse auszulösen und zu erfassen. Andererseits ist mit Blick auf die Praxisrelevanz aber keine *einseitige* Fokussierung auf misslingende Prozesse erstrebenswert. Sollen durch spezifische Beobachtung diagnostisch verwertbare Daten generiert werden, müssen diese Beobachtungen so spezifisch und präzise sein, dass sie konkrete Hypothesenbildungen bzgl. möglicher Interventionen zulassen. Der Eindruck, dass die Lehr-Lernsituation insgesamt schlecht verläuft, bietet dafür keine Ansatzpunkte. Die Fähigkeit zum Erkennen von Lernbarrieren muss demnach präzisiert werden: Es muss eine Differenzierung unterschiedlicher Interaktionen und ihrer Auswirkungen erfolgen und damit auch eine Unterscheidung von gelingenden Prozessen.

Die Ergebnisse stützen insofern zwar die Hypothese, dass das Verfahren eine gewisse Sensitivität besitzt, *unterschiedliche* Beurteilungen sichtbar zu machen. Dabei lassen sich jedoch hohe Werte bei der Beurteilung von Items zu Barrieren allein nicht als Expertise interpretieren. Eine Expertise hinsichtlich einer genauen und differenzierten Beobachtung und Beurteilung lässt sich in keiner der Gruppen feststellen, auch nicht in der Gruppe der Lehrkräfte. Das wurde im dritten Entwicklungsschritt erst durch die Verwendung von Items zu gelingenden Situationen

deutlich. Für den nächsten Entwicklungsschritt ist deshalb eine stärkere Betonung gelingender Prozesse und damit einer differenzierten Beurteilung zu überlegen.

11.5.9 Ergebnisse zur Kreuzvalidität

Frage 5: Lässt sich zeigen, dass die Datenstruktur und Befunde zu Gruppenunterschieden nicht zufällig oder stichprobengebunden sind, sondern sich replizieren lassen?

Hypothese 5.1: Die Datenstrukturen gleichen sich in den verschiedenen Stichproben.

Hypothese 5.2: Gruppenunterschiede lassen sich in verschiedenen Stichproben replizieren.

Im Rahmen der Kreuzvalidierung soll geprüft werden, ob sich die für die Gruppe der Hamburger Studierenden gewonnenen Erkenntnisse auch in anderen Stichproben replizieren lassen. Dies ist insbesondere in Bezug auf mögliche Erkenntnisse zur Beschaffenheit der PW von Bedeutung (Gold & Holodynski, 2017), so dass üblicherweise die Datenstruktur sowie die Gruppenunterschiede geprüft werden (Meschede, 2014). Um ausreichende Fallzahlen nutzen zu können, werden alle Studierenden der anderen Universitäten zusammengefasst. Das Vorgehen für die Faktorenanalyse entspricht dem der vorherigen Analysen. Das Maß der Stichprobeneignung nach dem Kaiser-Meyer-Olkin-Kriterium (.78) und die Signifikanz nach Bartlett ($p < .001$) zeigen, dass sich die Daten für eine Faktorenanalyse eignen.

Es ergeben sich drei Faktoren, wobei die Items zu den drei Vignetten jeweils einen Faktor bilden. Die Items zu gelingenden Situationen korrelieren in allen Vignetten negativ mit den übrigen Items (siehe Tabelle 31). Die drei Faktoren erklären rund 24 %, 13 % und 11 % der Gesamtvarianz (siehe Anhang D Tabelle D20). Die Struktur der Daten entspricht damit weitgehend den Ergebnissen der Hamburger Stichprobe. Die Komponenten korrelieren mit $r = .17$ bis $r = .33$ (siehe Anhang D Tabelle D21). Die Korrelationen der Items fallen für beide Stichproben für die Vignette 2 und 3 sehr ähnlich aus. Die Items zu *Allgegenwärtigkeit*, *Klarheit* und *Übergängen* zeigen jeweils die höchsten Korrelationen, während die Items zu *Raum* sowie *Regeln und Routinen* die geringsten Korrelationen zeigen. Für Vignette 4 unterscheiden sich die Korrelationen in den Stichproben etwas.

Tabelle 31*Ergebnisse der Faktorenanalyse*

Items	Faktorladungen		
	1	2	3
Faktor 1: Vignette 3			
V3 Klarheit Barriere	.80		
V3 Feedback gelingend	-.76		
V3 Regeln und Routinen Barriere	.54		
V3 Übergänge Barriere	.83		
V3 Raum gelingend	-.55		
V3 Allgegenwärtigkeit Barriere	.81		
Faktor 2: Vignette 2			
V2 Feedback Barriere		.61	
V2 Übergänge gelingend		-.69	
V2 Raum Barriere		.52	
V2 Allgegenwärtigkeit Barriere		.75	
V2 Klarheit Barriere		.68	
V2 Regeln und Routinen gelingend		-.66	
Faktor 3: Vignette 4			
V4 Raum Barriere			.78
V4 Übergänge Barriere			.70
V4 Klarheit Barriere			.51
V4 Feedback Barriere			.57
V4 Regeln und Routinen Barriere			.63
V4 Allgegenwärtigkeit gelingend			-.74

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4.

Gruppenunterschiede lassen sich aufgrund der Verteilung in den Stichproben nicht bzgl. der gleichen Teilgruppen wie in der Stichprobe der Universität Hamburg vergleichen. Studiengänge lassen sich nur für die Gruppe 4. bis 6. Bachelorsemester vergleichen. Ein Vergleich der Fachsemester ist nur in der Gruppe der Allgemeinpädagogikstudierenden möglich. Für den Vergleich zwischen Studierenden und Lehrkräften wird dieselbe Gruppe von Fortbildungslehrkräften genutzt.

Unterschiede zwischen den Fachsemestern lassen sich, geprüft mit dem Kruskal-Wallis-Test, nicht feststellen (siehe Anhang D Tabelle D22). Die Studienganggruppen im 4. bis 6. Bachelorsemester unterscheiden sich, geprüft mit dem Mann-Whitney-U-Test, nur bzgl. des 2. Faktors (Vignette 2) signifikant (siehe Anhang D Tabelle D23), $z = -1.920$, $p = .055$. Das entspricht mit $r = .20$ einem schwachen Effekt nach Cohen (1992) (siehe Anhang D Tabelle D24). Die differenzierte Betrachtung zeigt, dass dies auf die Items zu Barrieren zurückzuführen ist, $z = -2.525$, $p = .012$. Hier liegt mit $r = .26$ ein schwacher Effekt nach Cohen (1992) vor. Es zeigt sich, dass die Studierenden der Sonderpädagogik zutreffender beurteilen (siehe Anhang D Tabelle D25).

Studierende unterscheiden sich von Lehrkräften bzgl. der drei Faktoren nicht signifikant (siehe Anhang D Tabelle D26). Betrachtet man die Werte differenziert, so zeigen sich signifikante Unterschiede bei Items zu Barrieren in Vignette 2 (siehe Anhang D Tabelle D27), $z = 2.347$, $p = .019$ mit einem schwachen Effekt ($r = .18$) nach Cohen (1992) sowie bei Items zu gelingenden Situationen in Vignette 2 (siehe Anhang D Tabelle D28), $z = -2.741$, $p = .006$ mit einem schwachen Effekt ($r = .21$) nach Cohen (1992) sowie in Vignette 3 bei Items zu gelingenden Situationen (siehe Anhang D Tabelle D29), $z = -2.815$, $p = .005$. Auch hier liegt mit $r = .22$ ein schwacher Effekt nach Cohen (1992) vor. Des Weiteren kommt es in Vignette 4 bei Items zu Barrieren (siehe Anhang D Tabelle D30) zu einem signifikanten Unterschied, $z = 1.880$, $p = .060$. Es liegt mit $r = .15$ ein schwacher Effekt nach Cohen (1992) vor. In allen Fällen beurteilten Lehrkräfte die Barrieren zutreffender und Studierende die gelingenden Situationen.

11.5.10 Interpretation der Kreuzvalidität

Die Faktorenanalyse zeigt für die Stichprobe der Studierenden anderer Universitäten ein sehr ähnliches Bild wie für die Stichprobe der Universität Hamburg: Es bilden sich Testlets für die einzelnen Vignetten und es liegen negative Korrelationen zwischen Items zu Barrieren und Items zu gelingenden Situationen vor. Faktorladungen für die einzelnen Vignetten sind überwiegend hoch, während die Faktoren nur geringfügig miteinander korrelieren. Die Höhe der Ladungen der einzelnen Items auf den jeweiligen Faktoren verteilt sich in Vignette 2 und 3 sehr ähnlich. Lediglich die Items zu *Raum* und *Allgegenwärtigkeit* in Vignette 4 unterscheiden sich in beiden Faktorenanalysen.

Auch bezüglich der Kriteriumsvalidität wiederholen sich die Befunde insofern, als Studierende der Sonderpädagogik und Lehrkräfte Barrieren eindeutiger beurteilen, während gelingende Situationen weniger erkannt werden. Letzteres gelingt den Studierenden der Allgmein-pädagogik besser.

Die Aussagekraft ist wegen der Zusammensetzung der Stichprobe begrenzt. Die Ergebnisse sprechen aber dafür, dass die Befunde replizierbar und somit nicht zufällig sind.

11.6 Diskussion und Konsequenzen für die Weiterentwicklung

Nach dem aufwändigen Konstruktionsprozess zeigt sich, dass Schwierigkeitsgrade und die Verteilung der Items auf diesem Weg verbessert werden konnten. Die Vignetten 2, 3 und 4 zeigen akzeptable Trennschärfen und Reliabilitäten. Die Vignettenkonstruktion soll deshalb grundsätzlich beibehalten werden. Während in der klassischen Testentwicklung die erste Vignette sowie wenig trennscharfe Items verworfen werden müssten, ist es im Rahmen eines explorativen Vorgehens aufschlussreich, danach zu fragen, was die erste Vignette von den anderen unterscheidet und ob Veränderungen der Vignette zu besseren Itemkennwerten führen.

Die Faktorenstruktur zeigt über verschiedene Kohorten hinweg, dass die Vignetten in sich homogen sind, aber deutliche Testlets bilden. Dieser Effekt ist im nächsten Entwicklungsschritt genauer zu untersuchen. Der starke Einfluss der einzelnen Vignetten bringt die Frage mit sich, welche Faktoren in den Vignetten die Wahrnehmung beeinflussen. Die negativen Korrelationen zwischen Items zu Barrieren und Items zu gelingenden Situationen lassen vermuten, dass eine positive oder negative Gesamtwahrnehmung der Situation die Urteile stark beeinflusst, so dass diese Gesamtwahrnehmung erfasst werden sollte. Gleichzeitig ist zu überlegen, durch welche Veränderungen in den Vignetten und im methodischen Vorgehen eine differenziertere Wahrnehmung angeregt werden könnte bzw. genauer erfasst werden kann. Aus diagnostischer Sicht können passgenaue Interventionen nur dann erfolgen, wenn die Ausgangssituation möglichst genau beschrieben wird. Das erfordert eine *spezifische* Beobachtung und *differenzierte* Beurteilung von Unterricht, also auch gelingender Prozesse, während globale Beobachtungen und Pauschalurteile keine verwertbare Datenbasis liefern. Expertise im Erkennen wäre dann nicht nur an ein Erkennen von Barrieren zu knüpfen, sondern dieses Erkennen von Barrieren müsste differenziert und in Abgrenzung zu gelingenden Situationen erfolgen. Gelingende Situationen müssten also zu gleichen Anteilen in den Vignetten beschrieben sein.

Bezüglich der Gruppenvergleiche lassen sich wegen der kleinen Stichproben nur bedingt Aussagen treffen. Effekte zeigen sich zwischen den Gruppen nur punktuell und in eher geringem Ausmaß. Allerdings zeigen sich wiederholt gleiche Trends in den drei Entwicklungsschritten und auch in verschiedenen Stichproben, so dass diese als Hinweis gelten können, dass Unterschiede zwischen verschiedenen Gruppen erfasst werden können. Zudem macht es deutlich, dass es sinnvoll ist, mehrere unterschiedlich situierte Vignetten einzusetzen, weil diese von verschiedenen Personengruppen unterschiedlich eindeutig beurteilt werden.

Es wird allerdings auch deutlich, dass Unterschiede bzgl. der Beurteilung von Barrieren nicht unbedingt als höhere Expertise interpretiert werden können. Unterschiede werden eher hinsichtlich einer mehr oder weniger ausgeprägten negativen Beurteilung der Gesamtsituation

sichtbar. Das differenzierte Erkennen gelingender oder nicht gelingender Prozesse scheint eine Kompetenz zu sein, die nicht automatisch durch Berufserfahrung zunimmt. Es scheint eher im Gegenteil so zu sein, dass Lehrkräfte holistischer urteilen. Dies könnte damit zusammenhängen, dass es im realen Unterricht nötig ist, schnelle Urteile zu fällen. Zudem scheint eine unterschiedliche Fokussierung eine Rolle zu spielen, die dazu führt, dass Studierende der Allgemeinpädagogik gelingende Prozesse eher wahrnehmen und Studierende der Sonderpädagogik und Lehrkräfte eher misslingende Prozesse. Um dies zu überprüfen, soll die Stichprobe im nächsten Entwicklungsschritt Studierende unterschiedlicher Studiengänge und Lehrkräfte umfassen. Inwiefern Lerngelegenheiten eine Rolle spielen, lässt sich noch nicht beantworten. Zum Zeitpunkt der Datenerhebungen waren keine Seminare modular verankert, in denen sehr spezifisch das differenzierte Beobachten von Unterrichtsmerkmalen gelehrt wurde. Eine spezifische Lerngelegenheit für einen Teil der Stichprobe könnte diesbezüglich neue Erkenntnisse ermöglichen.

Weniger Einfluss auf das Erkennen scheinen die Skalengestaltung oder Itemmerkmale wie Polung oder Formulierung zu haben. Diese haben möglicherweise einen Einfluss auf den Urteilsprozess selbst und machen diesen anspruchsvoller. In Bezug auf das Erkennen von Lernbarrieren lässt sich aber kein systematischer Effekt erkennen.

Für den letzten Entwicklungsschritt ergeben sich daraus folgende Aspekte zur Gestaltung des Verfahrens: (1) die Beibehaltung des Prinzips, mehrere Vignetten zu nutzen, (2) eine weitere Prüfung und Qualifizierung der Vignetten, (3) die stärkere Betonung gelingender Prozesse, (4) die Erfassung des Gesamteindrucks, (5) die Berücksichtigung einer spezifischen Lerngelegenheit bei der Wahl der Stichprobe sowie (6) die Reduktion möglicher störender methodischer Einflüsse durch Distraktoren, unterschiedliche Polungen und Formulierungen.

12 Vierter Entwicklungsschritt

Im letzten Schritt der explorativen Testentwicklung sollen die Vignetten und Items durch eine weitere Expertinnen- und Expertenüberprüfung qualifiziert und in ausgewählten Gruppen erneut eingesetzt werden.

Der Fokus des letzten Entwicklungsschritts liegt auf der Frage, wovon das Erkennen von Lernbarrieren beeinflusst wird und welche Informationen zur Beurteilung in Abhängigkeit von der jeweiligen Vignette genutzt werden. Es stellt sich die Frage nach der Spezifität von Urteilen im Gegensatz zu Gesamteindrücken sowie dem Einfluss von Merkmalen der Lesenden in Form von Praxiserfahrung auf die Fähigkeit zum Erkennen von Lernbarrieren und gelingenden Situationen.

12.1 Überarbeitung der Vignetten und Items

Die Ergebnisse des dritten Entwicklungsschritts geben Hinweise darauf, dass die systematische Vignettenkonstruktion anhand validierter Beispiele ein erfolgversprechender Ansatz ist, so dass die grundsätzlichen Konstruktionskriterien beibehalten werden sollen. Die Zielsetzungen für die Weiterentwicklung ergeben sich unter anderem aus der Frage nach den Einflüssen auf die Beurteilung der Vignetten. Diese sollen durch verschiedene Ansätze erfasst werden:

(1) Bei der Überarbeitung der Vignetten sollen gelingende Situationen *insgesamt* in gleicher Anzahl wie Barrieren dargestellt werden, um die Fähigkeit einer differenzierten Beurteilung stärker in den Blick zu nehmen. Dabei ergibt sich die Möglichkeit, den Kontext *innerhalb* der Vignetten dahingehend zu variieren, wie der Anteil von Barrieren und gelingenden Situationen jeweils gestaltet wird. Vignetten mit mehr oder weniger Barrieren könnten Hinweise darauf geben, ob Urteile davon abhängen, wie die Situation insgesamt – gemessen an der Anzahl misslingender Prozesse – gelingt. Zudem soll insgesamt der Vignettenaufbau hinsichtlich der Verständlichkeit, Authentizität und Vorstellbarkeit geprüft werden.

(2) Neben der Anzahl von Barrieren und gelingenden Situationen pro Vignette könnten auch andere Faktoren die Gesamtwahrnehmung beeinflussen. Diese soll durch ein zusätzliches Item erhoben werden.

(3) Zu prüfen sind außerdem die Items. Mögliche Störfaktoren durch unterschiedliche Formulierungen, Polungen und Distraktoren sollen diskutiert werden. Zudem sollen die Itemformulierungen auch hinsichtlich der Benennung der Akteurinnen und Akteure geprüft werden sowie die Formulierungen der Instruktion und der Skala.

Methodisch wurde zur Qualifizierung der Vignetten und Items erneut eine Diskussionsrunde genutzt. Diese bestand aus fünf Expertinnen und Experten der Universität Hamburg, die sich im Rahmen eines Projekts zur Unterrichtsqualität intensiv mit der Beobachtung von KF und LU auseinandergesetzt hatten. Sie konnten sich im Vorfeld mit dem theoretischen Hintergrund des Verfahrens sowie mit den Vignetten, Items und der Instruktion vertraut machen. Die Vignetten wurden bzgl. der Verständlichkeit und Authentizität diskutiert und es wurde geprüft, ob die beschriebenen Handlungen nachvollziehbar die Merkmale der KF und LU und deren Ausprägung als Barriere oder gelingende Situation repräsentieren konnten. Items, Instruktion und Skala wurden bezüglich ihres Formats und der Formulierungen diskutiert. Maßgeblich für die Veränderungen war jeweils der Konsens der Diskussionsrunde.

12.1.1 Überarbeitung der Vignetten

Die Diskussion zu Vignette 1 ergab, dass diese von den Expertinnen und Experten als weniger authentisch beurteilt wurde. Die darin enthaltenen Barrieren wurden als sehr offensichtlich beurteilt. Die gelingende Situation dagegen wurde als weniger deutlich beurteilt. Die Lehrkraft wurde insgesamt als zu unsympathisch beschrieben. Außerdem wurde kritisch angemerkt, dass in den Vignetten variierte, ob zu Beginn Angaben zu Unterrichtsfach, Klassenstufe oder Thema gemacht wurden. Bei Beibehaltung der Situation wurden deshalb mehrere Formulierungen überarbeitet und die Angaben zur Situierung (Fach, Klassenstufe Thema) vereinheitlicht. Die Abschnitte zu den Merkmalen *Allgegenwärtigkeit* und *Feedback* wurden geändert und nicht mehr als Barriere dargestellt, sondern als gelingend. Das Verhältnis von Barrieren und gelingenden Situationen war damit in der überarbeiteten Vignette ausgewogen. Die Vignette lautet wie folgt:

Vignette 1 (Lösung siehe Anhang E1)

Unterrichtsbeginn Deutsch, 4. Klasse Grundschule, Thema Satzglieder. Die Kinder sitzen an Partnertischen. Der Lehrerin fällt auf, dass Murat noch Kaugummi kaut. Lehrerin: „Murat, spuck bitte dein Kaugummi aus!“ Murat steht auf, spuckt sein Kaugummi in den Müll, setzt sich wieder hin. Lehrerin: „Hanna, spuck dein Kaugummi aus!“ Hanna steht auf, spuckt ihr Kaugummi aus. In der Zwischenzeit fängt die Klasse an zu kichern – Murat packt ein neues Kaugummi aus. Lehrerin: „Murat, heb dir das bitte für die Pause auf. So, nun lasst uns endlich anfangen. Wir haben gestern mit den W-Fragen Satzglieder bestimmt. Das wollen wir jetzt noch mal so wie gestern wiederholen, damit ihr die Wer-oder-Was-, die Wem-oder-Was- und die Wen-oder-Was-Fragen noch sicherer könnt. Ich schreib drei Sätze an die Tafel und unterstreiche einige Wörter. Ihr schreibt die nicht ab. Anschließend sucht ihr euch einen Satz aus und überlegt allein für euch, mit welcher Frage ihr nach den unterstrichenen Wörtern

fragen könnt.“ (erklärt ein Beispiel) „Also ich schreibe an und ihr sucht dann einen Satz aus und findet eine Frage.“

Die Lehrerin schreibt an die Tafel. Als sich Luise zu Franzi umdreht, um ihr ihre neuen Sticker zu zeigen, gibt ihr die Lehrerin kurz ein Zeichen, aufmerksam zu bleiben. Anschließend fordert sie die Kinder auf: „Sucht einen Satz aus und notiert die Frage zu den unterstrichenen Wortarten. Ihr habt 5 Minuten Zeit.“

Emma stöhnt: „Oh Mann, Paul macht sich immer so breit, ich hab' gar keinen Platz!“ Lehrerin: „Versucht jetzt bitte mal, still zu arbeiten.“

Nach 5 Minuten beendet die Lehrerin die Stillarbeit und beginnt, Ergebnisse zu sammeln. Lehrerin: „Jetzt legen alle die Stifte hin und machen mit, sonst werden wir nicht fertig, ihr sollt das nachher ja auch noch geordnet aufschreiben.“ Paul: „Das ist doof, ich kann das irgendwie nicht.“ Lehrerin: „Du konntest doch letztes Mal schon gut die Wer-Frage stellen, versuch doch das mal zuerst.“

Vignette 2 wurde in der Diskussionsrunde als authentisch beurteilt. Lediglich der Abschnitt, der die Information zur Barriere *Raum* beinhalten sollte, wurde als schwer verständlich eingeschätzt und etwas umformuliert. Das Verhältnis von vier Barrieren und zwei gelingenden Situationen blieb damit bestehen. Vignette 2 lautet:

Vignette 2 (Lösung siehe Anhang E2)

Sachunterricht, 4. Klasse Grundschule, Thema Magnetismus. Die Kinder sitzen an Gruppentischen. Kati kritzelt an ihrem Platz auf einem Zettel rum, während der Lehrer an der Tafel den Ablauf für die Gruppenarbeit erklärt.

Lehrer: „Kati, jetzt dreh' dich doch bitte mal zur Tafel um, sonst kriegst du wieder nix mit.“ (erklärt weiter) „Also, ihr stellt erst mal Hypothesen auf und dann überprüft ihr die im Experiment und dokumentiert die Ergebnisse und macht euch vielleicht auch schon Notizen dazu, wenn euch was Bestimmtes auffällt, zum Beispiel, ähm, also einfach was euch auffällt. Die Aufteilung für die Gruppen kennt ihr, wer schon fertig ist, bleibt am Gruppentisch und kann wie immer Zeichnungen zu den Notizen machen. Los geht's!“

Die SuS beginnen sofort eifrig, sich in den Gruppen zu beschäftigen. Der Lehrer wendet sich einer Gruppe zu, in der es Diskussionen gibt, und redet längere Zeit mit der Gruppe. Mit der Zeit wird es immer unruhiger in der Klasse. Lehrer: „Jetzt seid doch bitte mal noch einen Moment leise, damit die letzte Gruppe auch die Chance hat, das fertig zu machen! Seid ihr überhaupt schon alle fertig?“

Die Gruppenarbeit wird nach einiger Zeit durch einen Gong beendet und die SuS kommen zügig im Kreis zusammen, um ausführlich die Ergebnisse zu besprechen. Eine Gruppe berichtet, dass Geldmünzen magnetisch sind (1 Cent), eine andere Gruppe hat das Gegenteil festgestellt (50 Cent). Die SuS sollen ihre Vermutungen zu den Gründen äußern.

Önder: „Vielleicht hat die andere Gruppe den Magneten falsch gehalten?“ Lehrer: „Was meinst du, Linn?“ Linn: „Das ist doch beides Metall. Vielleicht haben die den Magneten auch zu weit weg gehalten!“ Lehrer: „Der nächste, Louis.“ Louis: „Aber 50 Cent sind viel schwerer!“ Lehrer: „Kay, was meinst du?“ Kay: „Die sind doch aus unterschiedlichen Metallen, vielleicht hat das damit zu tun?“ Lehrer: „Ja, das ist richtig.“(...)

Auch in Vignette 3 wurde die Authentizität als gut bewertet. Die Beschreibung zur Barriere *Raum* wurde aber als zu schwer nachvollziehbar beurteilt. Zusätzlich wurde in dieser Vignette die Formulierung des Vignettenabschnitts zur Barriere in der *Klarheit* als uneindeutig eingeschätzt. In der Überarbeitung wurden deshalb beide Abschnitte umformuliert. Der Abschnitt zur *Klarheit* wurde als gelingende Situation beschrieben, so dass das Verhältnis von Barrieren und gelingenden Situationen in dieser Vignette ausgewogen war:

Vignette 3 (Lösung siehe Anhang E3)

Kunstunterricht, 5. Klasse Stadtteilschule, Thema meditatives Zeichnen. Lehrerin: „Wir haben ja letzte Woche mit dem Thema ‚Zentangle‘, dem meditativen Zeichnen, begonnen und Konturen gezeichnet, heute geht es wie angekündigt um die Muster. Ich zeig euch ein paar Beispiele am Smartboard, da seht ihr das am besten. Jonas, setz‘ dich heut ruhig kurz hier vorne hin.“ (Zeigt einige Muster am Smartboard). „Wer fertig ist, darf frei arbeiten und nächste Woche beginnen wir mit der Ausstellung der Bilder. Fangt schon mal an, ich geh noch mal ins Lehrerzimmer, da hab‘ ich noch Schritt-für-Schritt-Anleitungen für ein paar Muster.“

Die Lehrerin verlässt den Klassenraum, die meisten Kinder arbeiten, aber es ist wieder unruhig. Nach einer Weile kommt die Lehrerin zurück. Alina: „Ich bin schon ganz weit, Antoni hat noch nicht mal angefangen!“ Die Lehrerin beginnt daraufhin, Antoni noch mal den nächsten Arbeitsschritt zu erklären. Sie unterbricht: „Moment, Antoni.“ (Zur Klasse:) „Ihr habt euch eben so toll beruhigt, jetzt ist es schon wieder zu laut hier, bitte versucht mal, euch weiter auf eure Arbeit zu konzentrieren. Zentangle ist ja meditatives Zeichnen, da solltet ihr euch also drauf konzentrieren und das geht am besten, wenn ihr dabei nicht redet.“ (erklärt Antoni weiter)

Die Kinder zeichnen eine Weile. Alina und Carolin sind schnell fertig und fragen: „Dürfen wir jetzt Stempelbilder machen?“ Lehrerin: „Oh, ihr wart ja schnell! Ja, dann könnt ihr stempeln oder was ihr wollt.“ Alina und Carolin laufen los und wollen Stempel holen. Lehrerin: „Alina, Carolin, bitte räumt die Zeichenstifte weg.“ Alina und Carolin gehen zurück und räumen die Stifte weg und wollen dann stempeln. Lehrerin: „Bitte auch die Bilder in die Mappen tun!“ Alina und Carolin laufen wieder zurück und räumen die Bilder weg und laufen dann wieder zum Regal mit den Stempeln. Antoni: „Oh Mann, habt ihr’s bald?“

Vignette 4 dagegen wurde als zu fachspezifisch eingeschätzt. Es bestand in der Diskussionsrunde Einigkeit darüber, dass die Beurteilung der Vignette neben dem Erkennen gelingender oder misslingender Merkmale der KF und der (eher fachunabhängigen) LU auch fachliches Wissen erforderte. Zudem wurde sie als weniger authentisch beurteilt. Statt umfangreiche Änderungen vorzunehmen, wurde die Vignette verworfen und anhand von Unterrichtsbeispielen aus dem dritten Entwicklungsschritt neu konstruiert und diskutiert. Damit über alle vier Vignetten Barrieren und gelingende Situationen in gleicher Anzahl vertreten sind, wurde für die 4. Vignette ein Anteil von zwei Barrieren und vier gelingenden Situationen vorgegeben. Angaben zur Situation wurden entsprechend der anderen Vignetten gemacht. Außerdem wurden Geschlechter der Lehrpersonen zu gleichen Anteilen gewählt (zweimal Lehrerin und zweimal Lehrer). Vignette 4 lautet wie folgt:

Vignette 4 (Lösung siehe Anhang E4)

Mathematik, 2. Klasse Grundschule, Thema Addieren bis 100 mit Zehnerübergang. Der Lehrer kommt rein und begrüßt die Kinder. Es herrscht noch Gekicher. Der Lehrer sorgt schnell für Ruhe und kündigt eine Aufgabe an: „So, ihr sitzt ja schon wie besprochen in euren Arbeitsgruppen zusammen und sollt gleich auch in diesen Gruppen arbeiten. Lennet, wenn es eurer Gruppe wieder zu laut ist, dann könnt ihr in den Gruppenraum gehen, da liegt auch alles, was ihr braucht. Ihr sollt heute mal überlegen, welche schlaun Rechenwege ihr bei schwierigen Aufgaben nutzen könnt, also zum Beispiel sowas wie $38+27$. Fangt mal an, ich geh dabei rum.“

Die Kinder fangen zügig an und der Lehrer geht zur ersten Gruppe und gibt einige Anregungen, welches Material die Kinder nutzen könnten. Da eine andere Gruppe nicht zu arbeiten scheint, geht der Lehrer anschließend zu diesem Gruppentisch und stellt fest, dass jeder für sich rechnet. Lehrer: „Ihr sollt euch das aber auch gegenseitig vorstellen und erklären!“ Milla: „Wir wollten das jetzt erst aufschreiben und dann in der Gruppe besprechen. Oder sollen wir das nur mündlich machen?“ Lehrer: „Ach so, nee, so könnt ihr das auch machen.“

Nach einiger Zeit kündigt der Lehrer das Ende der Gruppenarbeit an. Die Kinder kommen tischweise zügig in den Stuhlkreis und setzen sich auf ihre Plätze. Lehrer: „So, dann zeigt mal der Reihe nach, wie ihr gerechnet habt. Yasmin: „Ich hab' das nicht aufgeschrieben.“ Lehrer: „Erzähl halt mal.“ Yasmin: „Wir haben 50 plus 40 gerechnet und erst die Null weggetan und dann wieder hingeschrieben.“ Lehrer: „Da hast du es dir aber leicht gemacht, das ist ja gar kein Zehnerübergang!“

Die Übersicht (Tabelle 32) zeigt die Systematik der vier Vignetten.

Tabelle 32

Aufteilung der Lernbarrieren auf die Vignetten 1–4 unter Zuordnung von gelingenden Situationen und Barrieren

Vignette	Über- gänge	Allgegen- wärtigkeit	Raum	Feedback	Klarheit	Regeln und Rou- tinen	Verhältnis Barrieren/ gelingend
Vignette 1	Barriere	gelingend	Barriere	gelingend	gelingend	Barriere	3/3
Vignette 2	gelingend	Barriere	Barriere	Barriere	Barriere	gelingend	4/2
Vignette 3	Barriere	Barriere	gelingend	gelingend	gelingend	Barriere	3/3
Vignette 4	gelingend	gelingend	gelingend	Barriere	Barriere	gelingend	2/4

12.1.2 Überarbeitung der Instruktion, Items und Skala

Die Instruktion und die Beurteilung der Items wurden von der Expertinnen- und Expertenrunde insgesamt als kompliziert beurteilt und es wurde eine Überarbeitung zugunsten einer Reduktion von potenziellen Störfaktoren vorgeschlagen. Instruktion, Skala und Itemformulierungen sollten vereinfacht und der Begrüßungstext überarbeitet werden.

Um dem Einwand Rechnung zu tragen, dass ein kontextfreier Unterrichtsausschnitt hinsichtlich *potenzieller* Barrieren beurteilt werden muss, wurde folgender Hinweis hinzugefügt:

Im Folgenden sind Ausschnitte aus verschiedenen Lernsituationen beschrieben. Lesen Sie die Beispiele aufmerksam durch und beurteilen Sie bitte im Anschluss an jeden Text die dazu formulierten Aussagen.

Beziehen Sie sich bei der Beurteilung der Items allein auf die Informationen aus dem Beispiel. Es geht darum, was Ihnen auf den ersten Blick aufgefallen ist und welche Schlüsse Sie ziehen, ohne dass Sie Kontextinformationen kennen.

Die Formulierung der Instruktion wurde vereinfacht und lautete: *Bitte beurteilen Sie die folgenden Aspekte, setzen Sie bitte jeweils EIN Kreuz IN ein Kästchen.*

Die vierstufige Skala wurde beibehalten, die Formulierungen *trifft wahrscheinlich zu* etc. wurden jedoch geändert, da der Aspekt der *potenziellen* Auswirkungen bereits im Begründungstext berücksichtigt wurde. Die Antwortmöglichkeiten der Skala lauten jetzt *trifft gar nicht zu*, *trifft eher nicht zu*, *trifft eher zu* und *trifft völlig zu*.

Die unterschiedliche Polung der Items wird aufgehoben und es werden alle Items als positive Aussagen formuliert. Unterschiede in der Antwortrichtung ergeben sich demnach nicht mehr aus der Polung der Items, sondern dadurch, dass Merkmale in den Vignetten unterschiedlich als Barriere oder gelingende Situation ausgeprägt sind. Distraktoren werden nicht mehr verwendet.

Bezüglich der Itemformulierungen hatten die Expertinnen und Experten die Vermutung, dass es einen Einfluss auf die Beurteilung haben könnte, ob die Lehrkraft als handelnde Person genannt wird oder nicht. Es wurde diskutiert, dass eine Beurteilung der Lehrkraft als Person zurückhaltender ausfallen könnte als eine Beurteilung der Situation, ohne direkt einer Person die Verantwortung dafür zuzuschreiben. Bei den Itemformulierungen wird deshalb möglichst auf die Nennung der Lehrkraft verzichtet.

Diskutiert wurden außerdem zusätzliche Items. Zum einen soll die Gesamtwahrnehmung der jeweiligen Vignetten erhoben werden, zum anderen soll sichergestellt werden, dass Leserinnen und Leser die Vignetten authentisch finden und sich die beschriebenen Situationen gut vorstellen können. Es entstanden so die folgenden neuen Items, die einheitlich für alle vier Vignetten genutzt werden:

<i>Einschätzung der Klassenatmosphäre</i>	Es herrscht insgesamt eine positive Klassenatmosphäre
<i>Übergänge</i>	Übergänge zwischen Unterrichtseinheiten gelingen reibungslos
<i>Allgegenwärtigkeit</i>	Die Lehrkraft hat die ganze Klasse im Blick.
<i>Raumbedingungen</i>	Räumliche Gegebenheiten werden berücksichtigt.
<i>Feedback</i>	Die Lehrkraft gibt konstruktives Feedback.
<i>Klarheit</i>	Unterrichtsinhalte und -ziele werden deutlich.
<i>Regeln und Routinen</i>	Der Unterrichtsverlauf lässt etablierte Regeln und Routinen erkennen.

Einschätzung der Vignette hinsichtlich der Authentizität Die beschriebene Situation ist authentisch.

Beurteilung der Vorstellbarkeit Ich kann mir die beschriebene Situation gut vorstellen.

Diskutiert wurde in der Expertinnen- und Expertenrunde außerdem, ob die Art der Instruktion dazu beiträgt, dass Urteile aufgrund der Gesamtwahrnehmung gefällt werden. Um zu prüfen, welche Informationen in den Vignetten genutzt werden, um die Items zu beurteilen, wurde deshalb eine zusätzliche Instruktion entwickelt. Diese lautet: *Unterstreichen Sie bei der Bearbeitung der ersten Vignetten die Stellen im Text, die Sie für die Beantwortung der Items genutzt haben. Notieren Sie jeweils, welchem Item Sie diese Information zugeordnet haben.*

Die Vignetten wurden in zufälliger Reihenfolge präsentiert, so dass alle vier Vignetten zu ungefähr gleichen Anteilen anhand dieser Instruktion bearbeitet wurden. Die Ergebnisse ermöglichen es zu prüfen, ob (1) die Informationen zur Beurteilung genutzt werden, die bei der Konstruktion dafür vorgesehen wurden, und (2) ob die Aufforderung zur Markierung einzelner Informationen zu einer differenzierteren Beurteilung führt.

12.2 Fragestellungen und Hypothesen

Für das überarbeitete Verfahren soll zunächst analog zu den ersten Entwicklungsschritten die Testgüte geprüft werden. Weitere Fragen beziehen sich auf mögliche Einflüsse auf Urteilsprozesse.

Authentizität und Vorstellbarkeit als Voraussetzung zur Beurteilung

Frage 1: Bei der Konstruktion der Vignetten wurden reale Unterrichtsbeispiele verwendet und durch den beschriebenen Validierungsprozess inkl. Diskussion von Expertinnen und Experten geprüft. Trägt dies dazu bei, dass die Vignetten als authentisch empfunden werden und sich die Lesenden die beschriebenen Situationen gut vorstellen können?

Hypothese 1: Alle vier Vignetten werden als authentisch beurteilt und die Lesenden können sich die Situationen gut vorstellen.

Mittlere Schwierigkeitsgrade

Frage 2: Die Darstellungen in den Vignetten wurden dahingehend überprüft, ob die Barrieren und gelingenden Situationen zwar gut erkennbar, aber nicht zu deutlich sind. Führen die

Überarbeitungen zu mittleren Schwierigkeitsgraden, die für eine Differenzierung unterschiedlicher Fähigkeitsausprägungen ausreichend sind?

Hypothese 2: Es zeigen sich für die Items nach der Überarbeitung mittlere Schwierigkeitsgrade und eine ausreichende Verteilung.

Trennschärpen und Reliabilität

Frage 3: Im vierten Entwicklungsschritt wurden gelingende Situationen und Barrieren über alle Vignetten in insgesamt gleicher Anzahl verteilt. Itemformate wurden vereinheitlicht und präzisiert. Führen diese Veränderungen zu höheren Trennschärpen und einer zufriedenstellenden Reliabilität?

Hypothese 3: Die Items sind trennscharf und reliabel.

Konstruktvalidität

Frage 4: Bei ansonsten konstanten Konstruktionskriterien wurden die Vignetten bzgl. der Anteile von Barrieren und gelingenden Situationen gezielt variiert, um unterschiedliche Stimuli zu erproben. Die Überarbeitungen sollten auch zu differenzierten Urteilen innerhalb der Vignetten führen. Zeigen sich diese Konstruktionsmerkmale in der Struktur der Daten und werden Barrieren und gelingende Situationen differenziert beurteilt?

Hypothese 4: Die Datenstruktur zeigt die Konstruktionsmerkmale der Vignetten. Barrieren und gelingende Situationen werden innerhalb der Vignetten differenziert beurteilt.

Mögliche Einflussfaktoren auf das Erkennen von Lernbarrieren

Merkmale der Lesenden: Einfluss der Expertise

Frage 5: Das Fachsemester erwies sich in der dritten Datenerhebung nicht als Prädiktor für die Beurteilung der Vignetten. Es wurde vermutet, dass dies auch damit zusammenhängen könnte, dass in den Studiengängen zum Zeitpunkt der Datenerhebung keine Lerngelegenheiten zum spezifischen Beobachten von Barrieren in Unterrichtsmerkmalen verankert waren. Wirken spezifische Lerngelegenheiten oder Berufserfahrung als Prädiktor für das Erkennen von Lernbarrieren und gelingenden Situationen?

Hypothese 5: Spezifische Lerngelegenheiten und Praxiserfahrung erweisen sich als Prädiktor für das Erkennen von Lernbarrieren und gelingenden Lehr-Lern-Prozessen.

Vignettengestaltung und Klassenatmosphäre: Das Gelingen der Situation insgesamt

Frage 6: Die Vignetten wurden so konstruiert, dass über alle vier Vignetten gleich viele Barrieren und gelingende Situationen beschrieben wurden, innerhalb der Vignetten variiert aber das Verhältnis. Ist die Anzahl der jeweils in den Vignetten beschriebenen Barrieren ein Prädiktor für die Beurteilung der Klassenatmosphäre? Und erfolgt die Beurteilung der Items differenziert und unabhängig von der Gesamtzahl der Barrieren und der Einschätzung der Klassenatmosphäre?

Hypothese 6.1: Die Anzahl der in der Vignette dargestellten Barrieren wirkt als Prädiktor für die Einschätzung der Klassenatmosphäre.

Hypothese 6.2: Die Beurteilung der Items erfolgt differenziert und ist von der jeweiligen Anzahl von Barrieren und gelingenden Situationen sowie von der Einschätzung der Klassenatmosphäre unabhängig.

Instruktion: Einfluss der Markierungsaufgabe

Frage 7: Die Instruktion wurde dahingehend variiert, dass bei jeweils einer Vignette von den Masterstudierenden unterstrichen und notiert werden sollte, welche Textinformationen jeweils zur Beurteilung der Items genutzt wurden. Dies sollte zum einen sichtbar machen, ob die Informationen wie bei der Konstruktion gedacht genutzt wurden. Führt diese Aufforderung auch dazu, dass Beurteilungen zutreffender und differenzierter werden?

Hypothese 7.1: Studierende nutzen die Textinformationen, die bei der Konstruktion der Vignetten gesetzt wurden.

Hypothese 7.2: Die Aufforderung zur Markierung führt im Vergleich zur Instruktion ohne Markierung nicht zu einer genaueren Beurteilung.

12.3 Stichprobe

Im vierten Entwicklungsschritt sollte unter anderem der Einfluss von Lerngelegenheiten und Berufserfahrung auf die Beurteilung von Textvignetten erfasst werden. Der neue Satz von vier Textvignetten wurde deshalb an der Universität Hamburg mit Studierenden sowie mit Lehrkräften im Rahmen einer Fortbildung erprobt. Die Stichprobe der Universität Hamburg setzte sich aus Studierenden der Sonderpädagogik im 4. Bachelorsemester und im 4. Mastersemester zusammen. Für die Bachelorstudierenden wurde die Praxiserfahrung erhoben. Studierende, die über mehr Praxiserfahrung verfügten als für Studierende übliche Erfahrung durch Tätigkeiten wie Nebenjobs oder Praktika, wurden ausgeschlossen. Die Masterstudierenden absolvierten zum Befragungszeitpunkt das zweite Kernpraktikum und befassten sich in diesem Rahmen mit

der Beobachtung und Analyse von problematischen Lehr-Lernsituationen. Sie erhielten vor der Bearbeitung der Vignetten außerdem eine spezifische Lerngelegenheit in Form eines *Microteachings* zur Beobachtung von Barrieren in Unterrichtsmerkmalen. Dieses bestand darin, dass Maßnahmen für die unterschiedlichen Merkmale der KF und LU erarbeitet wurden und die Operationalisierung für ErLe dargestellt wurde. Das *Microteaching* bestand aus einer Wissensvermittlung im Umfang einer 90-minütigen Seminarsitzung. Die Anwendung von Unterrichtsmerkmalen als Beobachtungskategorien wurde nicht geübt. Diese Gruppe von Studierenden bearbeitete als einzige die Aufgabe, bei der in der jeweils ersten Vignette (zufällige Reihenfolge der Vignetten) relevante Textstellen markiert werden sollten.

Die Lehrkräfte wurden im Rahmen eines Arbeitskreises zu einem anderen Thema angesprochen und um die Bearbeitung der Vignetten gebeten. Es handelte sich um Lehrkräfte an allgemeinbildenden Schulen mit mehr als fünf Jahren Berufserfahrung (siehe Tabelle 33).

Tabelle 33

Häufigkeiten für die Stichprobe des vierten Entwicklungsschritts

Gruppe	<i>n</i>
Studierende der Sonderpädagogik, 4. Bachelorsemester	109
Studierende der Sonderpädagogik, 4. Mastersemester	66
Lehrkräfte	18
Gesamt	193

12.4 Design und Durchführung

Die Beurteilungsaufgaben wurden an der Universität Hamburg während der Seminarzeiten bearbeitet. Die Bachelorstudierenden bearbeiteten eine Online-Version, die Masterstudierenden bearbeiteten wegen der Markierungsaufgabe eine Papier-und-Bleistift-Version. Die Lehrkräfte bearbeiteten die Online-Befragung zu Beginn einer Fortbildung. Um möglichst ähnliche Durchführungsbedingungen zu schaffen, bekamen alle Teilnehmenden neben den üblichen Datenschutzhinweisen vorab die gleiche Information:

Liebe Seminarteilnehmerinnen und Seminarteilnehmer (bzw. Liebe Lehrkräfte),
im Rahmen des Projekts „Professionelles Lehrerhandeln zur Förderung fachlichen Lernens unter sich verändernden gesellschaftlichen Bedingungen (ProfaLe)“ werden derzeit an der Universität Hamburg verschiedene Aufgaben entwickelt, mit denen Unterrichtssituationen beurteilt werden können.

Um die Aufgaben erproben und verbessern zu können, bitten wir um Ihre Mithilfe!

Um die Beurteilungssituation näher an den realen Unterricht zu rücken, in dem Informationen nur einmalig wahrgenommen werden können, war das Zurückblättern zum Vignettentext nicht möglich. Die Vignetten wurden in allen Gruppen in zufälliger Reihenfolge präsentiert.

12.4.1 Auswertung des Vignettentests

Zur Auswertung wurden alle Antworten in SPSS übertragen, in Werte übersetzt und so rekodiert, dass erkannte Barrieren bzw. erkannte gelingende Situationen jeweils den Wert 3 oder 4 erhielten. Die Werte 1 oder 2 stehen für nicht erkannte Situationen (siehe Anhang E1). Schwierigkeiten werden über die mittleren Schwierigkeitsgrade und die Verteilung ermittelt und anhand einer deskriptiven Statistik dargestellt. Ebenso wird zur Qualifizierung der Vignetten anhand von Mittelwerten dargestellt, ob die Studierenden und Lehrkräfte die Vignetten authentisch finden und sich die Situationen gut vorstellen können. Anschließend werden Trennschärfen dargestellt und die Reliabilität anhand von Cronbachs Alpha berechnet. Mit der Frage nach der Tauglichkeit der einzelnen Vignetten sowie der Vignetten insgesamt wird dies vignettenweise und für die Gesamtskala durchgeführt. Die Datenstruktur wird analog zu den vorherigen Entwicklungsschritten anhand einer Faktorenanalyse mit Promax-Rotation untersucht.

Um Erkenntnisse dazu zu gewinnen, welche Einflüsse bei der Beurteilung eine Rolle spielen, sollen zunächst Zusammenhänge der Beurteilungen mit Lerngelegenheiten und Praxiserfahrung erhoben werden. Die Darstellung von unterschiedlichen Fähigkeitsausprägungen je Gruppe erfolgt entsprechend der Faktorenstruktur. Um jeweils die Gewichtungen der Items zu berücksichtigen, die durch die Faktorenanalyse vorgegeben werden, werden für die Faktoren Faktorenwerte anhand gewichteter Mittelwerte (Itemrohwert mal Faktorladung) berechnet. Die Gruppen werden anhand non-parametrischer Verfahren verglichen.

Zusammenhänge zwischen der Vignettenkonstruktion (Anzahl der Barrieren pro Vignette) und der Einschätzung der Klassenatmosphäre werden mittels Friedman-Test geprüft. Dafür werden innerhalb der Teilstichproben die Mittelwerte für die Items zur Klassenatmosphäre verglichen.

Außerdem wird untersucht, ob die Anzahl der Barrieren pro Vignette Einfluss auf die Beurteilung der Items hat. Dies erfolgt, indem innerhalb der Teilstichproben die Mittelwerte für Items zu Barrieren und Items zu gelingenden Situationen untereinander und je Vignette mittels Friedman-Test verglichen werden.

Der Einfluss der subjektiv empfundenen Klassenatmosphäre wird untersucht, indem vignettenweise Spearman-Rho-Korrelationen zwischen der Einschätzung der Klassenatmosphäre und den Mittelwerten der Items zu Barrieren und gelingenden Situationen geprüft werden.

12.4.2 Auswertung der Markierungsaufgabe

Die Antworten der Masterstudierenden zur Markierungsaufgabe werden ausgewertet, indem (1) alle Markierungen je Vignette ausgezählt werden, um zu prüfen, wie viele Informationen die Studierenden je Vignette insgesamt finden. (2) Des Weiteren werden richtige Markierungen ausgezählt. Als richtig werden Markierungen gewertet, die Textstellen mit relevanten Informationen betreffen, die im Sinne der Konstruktion die Informationen zur Beurteilung der Items tragen. (3) Es wird ausgewertet, ob diese Markierungen der richtigen Stellen dann zu einer richtigen oder falschen Beurteilung des jeweiligen Items führen.

Um zu überprüfen, ob die Markierungsaufgabe zu differenzierteren oder zutreffenderen Urteilen von Barrieren und/oder gelingenden Merkmalen führte, werden Mittelwerte für Barrieren und gelingende Situationen je Vignette gebildet und es wird ein Gruppenvergleich vorgenommen (Gruppe 1: Vignette anhand der Markierungsaufgabe bearbeitet; Gruppe 2: Vignette ohne Markierungen bearbeitet). Da unmittelbare Einflüsse auf die Beurteilung der einzelnen Items sichtbar gemacht werden sollen, werden Mittelwerte für Barrieren und gelingende Situationen getrennt und ohne Berücksichtigung von Gewichtungen, die der Faktorenanalyse für die Gesamtgruppe entsprechen, dargestellt.

12.5 Ergebnisse und Interpretation

Im Folgenden werden die Ergebnisse in der Reihenfolge der Forschungsfragen dargestellt.

12.5.1 Ergebnisse zur Authentizität und Vorstellbarkeit

Frage 1: Trägt die Konstruktion und Prüfung dazu bei, dass die Vignetten als authentisch empfunden werden und sich die Lesenden die beschriebenen Situationen gut vorstellen können?

Hypothese 1: Alle vier Vignetten werden als authentisch beurteilt und die Lesenden können sich die Situationen gut vorstellen.

Um zunächst zur Qualifizierung der Vignetten die Authentizität und Vorstellbarkeit zu prüfen, werden für die vier Vignetten jeweils die Mittelwerte für die Teilstichproben dargestellt. Angestrebt werden dabei eine gute Authentizität und Vorstellbarkeit, also Wert 3 oder 4 und damit ein Mittelwert, der mindestens 3.00 beträgt. Dies trifft für fast alle Mittelwerte in allen Gruppen zu. Lediglich die Authentizität der Vignette 2 (V2) und Vignette 3 (V3) wird von den Studierenden des 4. Bachelorsemesters geringfügig schlechter bewertet ($M_{V2} = 2.96$, $SD = 0.69$

und $M_{V3} = 2.94$, $SD = 0.69$) (siehe Tabelle 34). Geprüft mit dem Friedman-Test, werden diese Unterschiede aber in den paarweisen Vergleichen nicht signifikant. Der Unterschied zwischen Vignette 1 und Vignette 3 verfehlt knapp eine Signifikanz ($p = .052$) (siehe Anhang E Tabelle E5).

Tabelle 34

Mittelwerte für die Items zu Authentizität und Vorstellbarkeit der Vignetten 1–4, unterteilt nach Teilstichproben

Teilstichprobe		Items							
		V1 A	V1 VB	V2 A	V2 VB	V3 A	V3 VB	V4 A	V4 VB
4. Bachelorsemester	<i>M</i>	3.21	3.34	2.96	3.24	2.94	3.22	3.03	3.28
	<i>SD</i>	0.75	0.64	0.69	0.62	0.69	0.70	0.62	0.67
	<i>n</i>	109	109	109	109	109	109	109	109
Masterstudierende	<i>M</i>	3.25	3.23	3.25	3.25	3.22	3.29	3.32	3.38
	<i>SD</i>	0.62	0.70	0.64	0.62	0.60	0.66	0.59	0.60
	<i>n</i>	64	66	64	64	64	65	65	66
Lehrkräfte	<i>M</i>	3.28	3.17	3.11	3.28	3.22	3.28	3.11	3.22
	<i>SD</i>	0.46	0.79	0.47	0.46	0.55	0.75	0.58	0.55
	<i>n</i>	18	18	18	18	18	18	18	18
Insgesamt	<i>M</i>	3.23	3.28	3.07	3.25	3.06	3.25	3.14	3.31
	<i>SD</i>	0.68	0.67	0.67	0.61	0.66	0.69	0.62	0.64
	<i>n</i>	191	193	191	191	191	192	192	193

Anmerkung. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4, A = Authentizität, VB = Vorstellbarkeit.

12.5.2 Interpretation zu Authentizität und Vorstellbarkeit

Die Ergebnisse für die Beurteilung der Authentizität und Vorstellbarkeit für die vier entworfenen Vignetten sind in allen Gruppen zufriedenstellend, so dass diese die erste Hypothese stützen. Die Vignetten können bzgl. Authentizität und Vorstellbarkeit als geeignet betrachtet werden.

12.5.3 Ergebnisse zu mittleren Schwierigkeitsgraden

Frage 2: Führen die Überarbeitungen zu Schwierigkeitsgraden, die für eine Differenzierung unterschiedlicher Fähigkeitsausprägungen ausreichend sind?

Hypothese 2: Es zeigen sich für die Items nach der Überarbeitung mittlere Schwierigkeitsgrade und eine ausreichende Verteilung.

Für alle Items wird die vierstufige Skala ausgenutzt, wobei die Extremwerte seltener vergeben werden (siehe Anhang E Tabelle E6). Die Mittelwerte liegen für die Items zu Barrieren bei einem Skalenmittel von 2.50 zwischen 2.59 ($SD = 0.82$) und 3.08 ($SD = 0.78$). Items zu gelingenden Situationen sind etwas schwerer. Die Mittelwerte liegen hier zwischen 2.23 ($SD = 0.74$) und 3.16 ($SD = 0.66$) (siehe Tabelle 35). Der Mittelwert aller Items zu Barrieren liegt bei 2.75 ($SD = 0.33$) und der Mittelwert aller Items zu gelingenden Situationen bei 2.69 ($SD = 0.35$) (siehe Tabelle 35).

Tabelle 35

Deskriptive Statistik zu Items zu Barrieren und Items zu gelingenden Situationen

Items	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>M</i>	<i>SD</i>	<i>Var</i>
V1 Übergänge (Barriere)	192	1	4	2.61	0.69	.48
V1 Allgegenwärtigkeit (gelingend)	193	1	4	2.86	0.78	.60
V1 Raum (Barriere)	187	1	4	2.61	0.70	.50
V1 Feedback (gelingend)	192	1	4	2.46	0.81	.66
V1 Klarheit (gelingend)	193	1	4	2.81	0.85	.79
V1 Regeln (Barriere)	193	1	4	2.59	0.82	.67
V2 Übergänge (gelingend)	193	1	4	2.54	0.66	.44
V2 Allgegenwärtigkeit (Barriere)	191	1	4	2.90	0.77	.60
V2 Raum (Barriere)	188	1	4	2.62	0.69	.47
V2 Feedback (Barriere)	191	1	4	3.08	0.78	.61
V2 Klarheit (Barriere)	190	1	4	2.61	0.74	.55
V2 Regeln (gelingend)	192	1	4	2.62	0.80	.65
V3 Übergänge (Barriere)	192	1	4	2.86	0.69	.47
V3 Allgegenwärtigkeit (Barriere)	191	1	4	2.83	0.81	.65
V3 Raum (gelingend)	189	1	4	2.35	0.70	.49
V3 Feedback (gelingend)	193	1	4	2.23	0.74	.54
V3 Klarheit (gelingend)	193	1	4	2.58	0.77	.60
V3 Regeln (Barriere)	193	1	4	2.90	0.71	.51
V4 Übergänge (gelingend)	193	1	4	3.04	0.76	.57
V4 Allgegenwärtigkeit (gelingend)	193	1	4	2.73	0.78	.62
V4 Raum (gelingend)	189	1	4	3.16	0.66	.43
V4 Feedback (Barriere)	190	1	4	2.81	0.75	.56
V4 Klarheit (Barriere)	193	1	4	2.64	0.84	.71
V4 Regeln (gelingend)	190	1	4	2.86	0.70	.49

Anmerkung. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4. Höchste und niedrigste Mittelwerte für Items zu Barrieren und Items zu gelingenden Situationen sind fett markiert.

12.5.4 Interpretation zu mittleren Schwierigkeitsgraden

Die mittleren Schwierigkeitsgrade der Items liegen um das Skalenmittel von 2.50 herum. Die Mittelwerte aller Items zu Barrieren und zu gelingenden Situationen liegen etwas über dem Skalenmittel. Die mittleren Schwierigkeitsgrade können damit als angemessen bewertet werden. Die Standardabweichungen und Varianzen sprechen für eine zufriedenstellende Streuung der Daten, so dass unterschiedliche Fähigkeitsausprägungen deutlich werden können.

12.5.5 Ergebnisse zu Trennschärfen und Reliabilität

Frage 3: Führen die Veränderungen zu höheren Trennschärfen und einer zufriedenstellenden Reliabilität?

Hypothese 3: Die Items sind trennscharf und reliabel.

Die Reliabilitätsanalyse für die Vignetten und die gesamte Skala zeigt durchweg negative bzw. sehr geringe Werte für Cronbachs Alpha (siehe Anhang E Tabellen E7 bis E11). Da sich bereits in der vorangegangenen Analyse gezeigt hat, dass Items zu Barrieren und Items zu gelingenden Situationen keine konsistente Skala bilden, wird die Analyse getrennt für Items zu gelingenden Situationen und Items zu Barrieren durchgeführt. Trennschärfen $r_{j(t-j)} > .30$ zeigen sich für Items zu gelingenden Situationen bei 9 von 12 Items (siehe Tabelle 36) und für 4 von 12 Items zu Barrieren (siehe Tabelle 37). Cronbachs Alpha beträgt für die Items zu gelingenden Situationen .68 und für die Items zu Barrieren .62.

Tabelle 36

Trennschärfen der Items zu gelingenden Situationen in den Vignetten 1-4

Items	Skalenmittelwert, wenn Item weglassen	Skalenvarianz, wenn Item weglassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weglassen
Vignette 1 Allgegenwärtigkeit	29.36	15.03	.41	.65
Vignette 1 Feedback	29.76	15.45	.32	.66
Vignette 1 Klarheit	29.41	14.68	.42	.64
Vignette 2 Übergänge	29.68	17.45	.05	.70
Vignette 2 Regeln	29.59	16.65	.13	.69
Vignette 3 Raum	29.86	16.43	.22	.68
Vignette 3 Feedback	29.98	15.28	.40	.65
Vignette 3 Klarheit	29.63	15.52	.33	.66

Vignette 4 Übergänge	29.18	15.77	.30	.66
Vignette 4 Allgegenwärtigkeit	29.48	15.02	.42	.64
Vignette 4 Raum	29.05	15.59	.41	.65
Vignette 4 Regeln	29.35	15.53	.39	.65

Tabelle 37

Trennschärfen der Items zu Barrieren in den Vignetten 1-4

Items	Skalenmittelwert, wenn Item wegge- lassen	Skalenvarianz, wenn Item weggelassen	Korrigierte I- tem-Skala-Kor- relation	Cronbachs Alpha, wenn Item wegge- lassen
Vignette 1 Übergänge	30.48	14.78	.13	.63
Vignette 1 Raum	30.48	14.08	.26	.60
Vignette 1 Regeln	30.52	13.47	.30	.59
Vignette 2 Allgegenwärtigkeit	30.21	13.99	.23	.61
Vignette 2 Raum	30.49	13.83	.32	.59
Vignette 2 Feedback	30.02	13.39	.33	.59
Vignette 2 Klarheit	30.50	13.60	.33	.59
Vignette 3 Übergänge	30.27	14.09	.27	.60
Vignette 3 Allgegenwärtigkeit	30.27	13.66	.27	.60
Vignette 3 Regeln	30.22	13.71	.31	.59
Vignette 4 Feedback	30.29	13.75	.29	.60
Vignette 4 Klarheit	30.47	13.97	.20	.61

12.5.6 Interpretation der Trennschärfen und Reliabilität

Das negative Cronbachs Alpha für die Gesamtskala zeigt, dass auch im vierten Entwicklungsschritt Items zu Barrieren und gelingenden Situationen keine konsistente Skala bilden. Betrachtet man die Items jeweils getrennt, kann – angesichts der Tatsache, dass mehrere Merkmale in unterschiedlichen Vignetten verwendet werden – das Cronbachs Alpha als akzeptabel bezeichnet werden (Eckes, 2015). Höhere Trennschärfen sind für die Gesamtskala dann nicht zu erwarten. Die Daten stützen Hypothese 3 demnach nicht für die Gesamtskala.

12.5.7 Ergebnisse zur Konstruktvalidität

Frage 4: Zeigen sich die Konstruktionsmerkmale für die Vignetten in der Struktur der Daten und werden Barrieren und gelingende Situationen differenziert beurteilt?

Hypothese 4: Die Datenstruktur zeigt die Konstruktionsmerkmale der Vignetten. Barrieren und gelingende Situationen werden innerhalb der Vignetten differenziert beurteilt.

Mit der Annahme korrelierender Faktoren wird eine Hauptkomponentenanalyse mit Promax-Rotation durchgeführt. Das Kaiser-Meyer-Olkin-Kriterium (.73) und die Signifikanz nach Bartlett ($p < .001$) weisen darauf hin, dass sich die Daten für eine Faktorenanalyse eignen. Die Anzahl der Faktoren wird auf Faktoren mit mehr als drei Ladungen $>.60$ (Guadagnoli & Velicer, 1988, zitiert nach Bortz, 1999) beschränkt. Es ergeben sich vier Faktoren, wobei jeweils die Items einer Vignette einen Faktor bilden. Die Items zu gelingenden Situationen und zu Barrieren korrelieren innerhalb der Faktoren jeweils negativ (siehe Tabelle 38). Die Faktoren erklären 18 %, 12 %, 10 % und 9 % der Gesamtvarianz (siehe Anhang E Tabelle E12) und korrelieren jeweils geringfügig mit $r = .004$ bis $r = -.27$ (siehe Anhang E Tabelle E13).

Tabelle 38*Ergebnisse der Faktorenanalyse*

Items	Faktorladungen			
	1	2	3	4
Faktor 1: Vignette 1				
V1 Übergänge Barriere	-.78			
V1 Allgegenwärtigkeit gelingend	.79			
V1 Raum Barriere	-.40			
V1 Feedback gelingend	.75			
V1 Klarheit gelingend	.71			
V1 Regeln Barriere	-.71			
Faktor 2: Vignette 4				
V4 Übergänge gelingend		.75		
V4 Allgegenwärtigkeit gelingend		.80		
V4 Raum gelingend		.58		
V4 Feedback Barriere		-.64		
V4 Klarheit Barriere		-.68		
V4 Regeln gelingend		.68		
Faktor 3: Vignette 2				
V2 Übergänge gelingend			-.72	
V2 Allgegenwärtigkeit Barriere			.64	
V2 Raum Barriere			.62	
V2 Feedback Barriere			.65	
V2 Klarheit Barriere			.70	
V2 Regeln gelingend			-.71	

 Faktor 4: Vignette 3

V3 Übergänge Barriere	.73
V3 Allgegenwärtigkeit Barriere	.64
V3 Raum gelingend	-.57
V3 Feedback gelingend	-.69
V3 Klarheit gelingend	-.61
V3 Regeln Barriere	.73

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4.

12.5.8 Interpretation zur Konstruktvalidität

Es zeigen sich Testlets für die Vignetten und innerhalb der Vignetten durchgehend negative Korrelationen zwischen Items zu Barrieren und gelingenden Situationen. Die Korrelationen sind – außer für das Item zur *Raumbarriere* in den Vignetten 1, 3 und 4 – mit $>.60$ recht hoch.

Insgesamt zeigen sich also die Konstruktionsmerkmale der Vignetten deutlich in der Datenstruktur, was Hypothese 4 stützen würde. Allerdings sollte durch eine gleichmäßige Verteilung von Barrieren und gelingenden Situationen eine *differenzierte* Beurteilung angeregt werden. Die durchweg negativen Korrelationen zwischen Items zu Barrieren und gelingenden Situationen zeigen jedoch, dass Beurteilungen nicht differenziert, sondern eher holistisch für die Vignetten erfolgen, was Hypothese 4 nicht stützt.

Kritisch muss angemerkt werden, dass die Instruktion möglicherweise anbahnt, einen Gesamteindruck wiederzugeben. Die Formulierung *Es geht darum, was Ihnen auf den ersten Blick aufgefallen ist und welche Schlüsse Sie ziehen, ohne dass Sie Kontextinformationen kennen* war dem Umstand geschuldet, dass in der Arbeit mit Vignetten immer wieder der Einwand geäußert wurde, dass für eine Beurteilung mehr Kontextinformationen nötig wären und sich Studierende teilweise nicht gut darauf fokussieren konnten oder wollten, die Bedingungen in nur diesem Ausschnitt zu beurteilen. Allerdings besteht auch in der Praxis die Herausforderung darin, schnelle Urteile zu bilden und relevante Situationen auf einen Blick zu erfassen, aber trotzdem misslingende von gelingenden Situationen zu unterscheiden.

Die Faktorenanalyse gibt Hinweise darauf, dass die Vignetten unabhängig von der Verteilung der Barrieren einseitig beurteilt werden. Welche Faktoren in welcher Weise auf die Beurteilungen wirken, soll in den nächsten Schritten genauer untersucht werden.

12.5.9 Mögliche Einflussfaktoren auf das Erkennen von Lernbarrieren

Die Daten sprechen dafür, dass mehr oder andere Faktoren als die Fähigkeitsausprägung zum Erkennen von Lernbarrieren die Beurteilungen beeinflussen. Neben Expertiseunterschieden durch z. B. Lerngelegenheiten oder Praxiserfahrung sollen Konstruktionsmerkmale der Vignetten, die Wahrnehmung der Klassenatmosphäre sowie die Instruktion als mögliche Einflussfaktoren untersucht werden.

12.5.9.1 Ergebnisse zu Merkmalen der Lesenden: Einfluss der Expertise.

Frage 5: Wirken spezifische Lerngelegenheiten oder Berufserfahrung als Prädiktor für das Erkennen von Lernbarrieren und gelingenden Situationen?

Hypothese 5: Spezifische Lerngelegenheiten und Praxiserfahrung erweisen sich als Prädiktor für das Erkennen von Lernbarrieren und gelingenden Lehr-Lern-Prozessen.

Die Gruppen der Stichprobe unterscheiden sich bzgl. der Lerngelegenheiten und Praxiserfahrung. Die Bachelorstudierenden hatten keine spezifische Lerngelegenheit zum Erkennen von Barrieren in Unterrichtsmerkmalen und keine Berufspraxis. Die Masterstudierenden bekamen ein *Microteaching* zu Unterrichtsmerkmalen und deren Operationalisierung und hatten erste Praxiserfahrungen durch die Kernpraktika. Die Lehrkräfte besaßen mindestens fünf Jahre Berufspraxis. Um die Gruppen zu vergleichen, wurden die Ergebnisse der Faktorenanalyse genutzt und gewichtete Mittelwerte (Itemrohwert mal Faktorwert) für alle Items berechnet. Die Analysen zeigen, dass sich Testlets pro Vignette bilden. Faktorwerte wurden deshalb berechnet, indem der Mittelwert aller gewichteten Items pro Faktor (also pro Vignette) gebildet wurde. Items zu Barrieren und gelingenden Situationen bilden unterschiedliche Skalen. Obwohl sich für die Vignetten Testlets bilden, sind die Reliabilitäten für die Gesamtskalen zu Barrieren und gelingenden Situationen aber akzeptabel (siehe Kapitel 12.5.5), so dass auch für die 12 Items zu Barrieren und für die 12 Items zu gelingenden Situationen gewichtete Mittelwerte gebildet werden. Geprüft wird mit non-parametrischen Verfahren und es wird ein Signifikanzniveau von 10 % festgelegt. Es wird für die Post-hoc-Tests jeweils die angepasste zweiseitige Signifikanz berichtet.

Der Kruskal-Wallis-Test zeigt, dass die Ergebnisse zu Faktor 2 (Vignette 4), $\chi^2(2) = 6.299$, $p = .043$, und Faktor 3 (Vignette 2), $\chi^2(2) = 5.983$, $p = .050$ von der Expertise beeinflusst werden (siehe Anhang E Tabelle E14). Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass der Unterschied in Vignette 2 zwischen Bachelorstudierenden und Masterstudierenden signifikant ist, $z = -2.335$, $p = .059$ (siehe Anhang E Tabelle E15). Es handelt sich mit $r = .18$ um einen schwachen Effekt nach Cohen (1992). Auch der Unterschied in Vignette 4 ist zwischen Bachelorstudierenden und Masterstudierenden signifikant, $z = -2.473$, $p = .040$ (siehe Anhang E Tabelle E16). Es handelt

sich mit $r = .19$ um einen schwachen Effekt nach Cohen (1992). Beide Vignetten werden von den Masterstudierenden insgesamt zutreffender beurteilt.

In Bezug auf das Erkennen von Barrieren zeigt der Kruskal-Wallis-Test signifikante Unterschiede, $\chi^2(2) = 17.766$, $p < .001$. Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass der Unterschied zwischen Bachelorstudierenden und Masterstudierenden signifikant ist, $z = -3.249$, $p = .003$. Es liegt mit $r = .25$ ein schwacher Effekt nach Cohen (1992) vor. Zudem zeigen die Post-hoc-Tests einen signifikanten Unterschied zwischen Bachelorstudierenden und Lehrkräften, $z = -3.364$, $p = .002$. Es liegt mit $r = .30$ ein mittlerer Effekt nach Cohen (1992) vor (siehe Anhang E Tabelle E17). Lehrkräfte beurteilen demnach Barrieren am zutreffendsten, gefolgt von Masterstudierenden und Bachelorstudierenden.

Auch bei der Beurteilung gelingender Momente zeigen sich signifikante Unterschiede im Kruskal-Wallis-Test, $\chi^2(2) = 6.094$, $p = .048$. Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass der Unterschied zwischen Lehrkräften und Bachelorstudierenden signifikant ist, $z = 2.393$, $p = .050$ mit einem schwachen Effekt ($r = .21$) nach Cohen (1992) sowie zwischen Lehrkräften und Masterstudierenden, $z = 2.311$, $p = .063$ mit einem schwachen Effekt ($r = .25$) nach Cohen (1992) (siehe Anhang E Tabelle E18). Studierende beurteilen demnach gelingende Situationen zutreffender als Lehrkräfte.

12.5.9.2 Interpretation zu Merkmalen der Lesenden: Einfluss der Expertise. Reliabilitäts- und Faktorenanalysen zeigen unterschiedliche Beurteilungen von Barrieren und gelingenden Situationen und für die Vignetten. Betrachtet man die Beurteilungen der jeweiligen Items für die verschiedenen Expertisegruppen, so zeigt sich, dass Lehrkräfte Barrieren sehr deutlich beurteilen, gelingende Situation jedoch weniger zutreffend als Studierende beurteilen. Bei Studierenden hängt es von den Vignetten ab, was zutreffender beurteilt wird. Sehr hohe Werte für Items zu Barrieren und sehr niedrige Werte für Items zu gelingenden Situationen führen in der Gruppe der Lehrkräfte insgesamt zu mittleren Werten für die Vignetten, so dass sich die Lehrkräfte hinsichtlich der Faktorenwerte pro Vignette nicht signifikant von anderen Gruppen unterscheiden. Hier zeigt sich, dass erst die differenzierte Auswertung von Barrieren und gelingenden Situationen Erkenntnisse zu Gruppenunterschieden ermöglicht.

Expertise im Beurteilen von Lehr-Lernprozessen sollte sich in einer differenzierten Beurteilung, also in einer Unterscheidung gelingender und nicht-gelingender Prozesse zeigen und somit in gleichermaßen hohen Werten für Items zu Barrieren und Items zu gelingenden Situationen in allen Vignetten. Diese Expertise zeigt sich allerdings in keiner der Gruppen und scheint weder durch die rein wissensbezogene spezifische Lerngelegenheit noch durch Berufspraxis in ausreichendem Maß zu entstehen. Lerngelegenheiten scheinen insofern eine Rolle zu spielen, als Masterstudierende

höhere Werte erzielen als Bachelorstudierende, allerdings ist dies abhängig von der Vignette. Eine differenzierte Beurteilung ist auch in der Gruppe der Masterstudierenden nicht zu erkennen.

Dass Unterschiede auch durch verschiedene Vignetten entstehen, zeigt, dass es sinnvoll ist, unterschiedliche Vignetten zu nutzen. Erstens entspricht das der ökologischen Validität – Barrieren treten situationsspezifisch in immer anderer Gestalt auf. Zweitens zeigt sich in den Daten keine universelle Fähigkeit zum Erkennen von Barrieren, sondern diese ist immer auch abhängig von der jeweiligen Situation.

12.5.9.3 Ergebnisse zum Einfluss durch die Vignettengestaltung und die Einschätzung der Klassenatmosphäre: Das Gelingen der Situation insgesamt.

Frage 6: Ist die Anzahl der jeweils in den Vignetten beschriebenen Barrieren ein Prädiktor für die Beurteilung der Klassenatmosphäre? Und erfolgt die Beurteilung der Items differenziert und unabhängig von der Gesamtzahl der Barrieren und der Einschätzung der Klassenatmosphäre?

Hypothese 6.1: Die Anzahl der in der Vignette dargestellten Barrieren wirkt als Prädiktor für die Einschätzung der Klassenatmosphäre.

Hypothese 6.2: Die Beurteilung der Items erfolgt differenziert und ist von der jeweiligen Anzahl von Barrieren und gelingenden Situationen sowie von der Einschätzung der Klassenatmosphäre unabhängig.

Bei der Konstruktion der Vignetten wurden Barrieren und gelingende Situationen ausgewählt in die Vignetten aufgenommen. Innerhalb der Vignetten variiert die Anzahl der Barrieren aber, um prüfen zu können, ob das Gelingen der Situation insgesamt – gemessen an der Gesamtzahl der Barrieren – einen Einfluss auf die Einschätzung der Klassenatmosphäre und auf die Beurteilungen hat. Dazu werden zuerst Mittelwerte für das Item zur Klassenatmosphäre für die einzelnen Gruppen dargestellt (siehe Tabelle 39). Es zeigt sich, dass die Klassenatmosphäre für Vignette 4, in der nur zwei Barrieren beschrieben werden, in allen Gruppen am positivsten eingeschätzt wird. Für die anderen Vignetten zeigen sich aber keine entsprechenden Zusammenhänge zwischen der Anzahl der Barrieren und der Einschätzung der Klassenatmosphäre – Vignette 2, die als einzige Vignette vier Barrieren enthält, wird am zweitpositivsten eingeschätzt.

Tabelle 39

Mittelwerte für die Einschätzungen der Klassenatmosphäre in den Vignetten 1–4, aufgeteilt nach Teilstichproben

Vignetten	Mittelwert Klassenatmosphäre		
	Bachelor	Master	Lehrkräfte
Vignette 1 (drei Barrieren)	2.39	2.53	2.50
Vignette 2 (vier Barrieren)	2.52	2.65	2.67
Vignette 3 (drei Barrieren)	2.41	2.65	2.61
Vignette 4 (zwei Barrieren)	2.94	3.06	3.22

Anmerkung. Bachelor $n = 109$, Master $n = 66$, Lehrkräfte $n = 18$.

Ob die Unterschiede zwischen den Beurteilungen der Klassenatmosphäre für die Vignetten innerhalb der Gruppen signifikant sind, wird mit dem Friedman-Test geprüft. Für die anschließend durchgeführten Post-hoc-Tests (Dunn-Bonferroni-Tests) wird jeweils die angepasste zweiseitige Signifikanz berichtet. Das Signifikanzniveau liegt bei 5 %.

In der Gruppe der Bachelorstudierenden zeigen sich signifikante Unterschiede zwischen den Beurteilungen der Klassenatmosphäre, $\chi^2(3) = 46.502$, $p < .001$, $n = 109$ (siehe Anhang Tabelle E19). Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass sich Vignette 4 jeweils signifikant von den anderen Vignetten unterscheidet (siehe Anhang E Tabelle E20). Die Klassenatmosphäre in Vignette 4 (zwei Barrieren) wird signifikant positiver beurteilt als in den anderen drei Vignetten:

Unterschied zu Vignette 1: $z = -0.853$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied zu Vignette 2: $z = -0.656$, $p = .001$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied zu Vignette 3: $z = -0.803$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .10$

Auch in der Gruppe der Masterstudierenden zeigen sich signifikante Unterschiede zwischen den Beurteilungen der Klassenatmosphäre, $\chi^2(3) = 31.081$, $p < .001$, $n = 63$ (siehe Anhang E Tabelle E21). Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass auch in dieser Gruppe die Unterschiede auf Vignette 4 zurückzuführen sind, die sich jeweils signifikant von den anderen Vignetten unterscheidet (siehe Anhang E Tabelle E22). Die Klassenatmosphäre in Vignette 4 (zwei Barrieren) wird signifikant positiver beurteilt als in den anderen drei Vignetten:

Unterschied zu Vignette 1: $z = -0.952$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .12$

Unterschied zu Vignette 2: $z = -0.722$, $p = .010$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied zu Vignette 3: $z = -0.738$, $p = .008$, schwacher Effekt nach Cohen (1992), $r = .10$

Signifikante Unterschiede in der Beurteilung der Klassenatmosphäre zeigen sich ebenfalls in der Gruppe der Lehrkräfte, $\chi^2(3) = 16.297, p < .001, n = 18$ (siehe Anhang E Tabelle E23). Post-hoc-Tests (Dunn-Bonferroni-Tests) zeigen, dass Vignette 4 in dieser Gruppe aber nur signifikant positiver als Vignette 1 (drei Barrieren) beurteilt wird, $z = -1.250, p = .022$ mit einem schwachen Effekt nach Cohen (1992), $r = .30$ (siehe Anhang E Tabelle E24).

Insgesamt zeigen sich also signifikante Unterschiede bzgl. der Einschätzung der Klassenatmosphäre nur für Vignette 4, in der nur zwei Barrieren beschrieben werden.

Geprüft wird auch, ob sich ein Zusammenhang zwischen der Anzahl der Barrieren einer Vignette mit der Beurteilung der Barrieren und gelingenden Situationen zeigt. Dies wird zunächst deskriptiv anhand der Mittelwerte dargestellt (siehe Tabelle 40). Es zeigt sich, dass alle Gruppen in Vignetten 4, in der nur zwei Barrieren, aber vier gelingende Situationen beschrieben sind, die Items zu gelingenden Situationen am zutreffendsten beurteilen. Allerdings werden in keiner Gruppe die Barrieren in Vignette 2 (vier Barrieren) am zutreffendsten beurteilt.

Tabelle 40

Mittelwerte für Items zu Barrieren und Items zu gelingenden Situationen in Abhängigkeit von der Anzahl der Barrieren pro Vignette, aufgeteilt nach Teilstichproben

Vignetten und Items	Mittelwerte		
	Bachelor	Master	Lehrkräfte
Vignette 1 Barrieren	2.57	2.58	2.96
Vignette 1 gelingend	2.75	2.79	2.22
Vignette 2 Barrieren	2.73	2.88	2.90
Vignette 2 gelingend	2.56	2.62	2.50
Vignette 3 Barrieren	2.78	2.96	2.98
Vignette 3 gelingend	2.45	2.30	2.37
Vignette 4 Barrieren	2.56	2.88	3.11
Vignette 4 gelingend	2.95	3.00	2.79

Anmerkung. Vignette 1: drei Barrieren, Vignetten 2: vier Barrieren, Vignette 3: drei Barrieren, Vignette 4: zwei Barrieren. Höchste Werte je Gruppe sind fett markiert.

Ob Unterschiede bzgl. der Beurteilungen von Barrieren und gelingenden Situationen innerhalb der Gruppen signifikant sind, wird mittels Friedman-Test geprüft. Für die anschließend durchgeführten Post-hoc-Tests (Dunn-Bonferroni-Tests) wird jeweils die angepasste zweiseitige Signifikanz berichtet. Das Signifikanzniveau liegt bei 5 %.

In der Gruppe der Bachelorstudierenden zeigen sich für die Beurteilungen der Barrieren signifikante Unterschiede, $\chi^2(3) = 14.463$, $p = .002$, $n = 109$ (siehe Anhang E Tabelle E25). Post-Hoc-Tests zeigen, dass sich die Beurteilungen in Vignette 1 und 3 signifikant voneinander unterscheiden, $z = -0.587$, $p = .005$ mit einem schwachen Effekt nach Cohen (1992), $r = 0.10$ (siehe Anhang E Tabelle E26). In beiden Vignetten ist das Verhältnis von Barrieren und gelingenden Situationen ausgewogen. Die Barrieren in Vignette 1 werden signifikant weniger zutreffend beurteilt.

In der Gruppe der Masterstudierenden zeigen sich für die Beurteilungen der Barrieren ebenfalls signifikante Unterschiede, $\chi^2(3) = 22.126$, $p < .001$, $n = 66$ (siehe Anhang E Tabelle E27). Post-hoc-Tests zeigen, dass sich die Beurteilungen in Vignette 1 jeweils von den anderen drei Vignetten signifikant unterscheiden. Für Vignette 1 (drei Barrieren) werden die Items zu Barrieren signifikant weniger zutreffend beurteilt als in den anderen Vignetten (siehe Anhang E Tabelle E28).

Unterschied Vignette 1 zu Vignette 2: $z = -0.674$, $p = .016$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied Vignette 1 zu Vignette 3: $z = -0.909$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .12$

Unterschied Vignette 1 zu Vignette 4: $z = -0.841$, $p = .001$, schwacher Effekt nach Cohen (1992), $r = .10$

In der Gruppe der Lehrkräfte finden sich keine signifikanten Unterschiede zwischen den Beurteilungen der Barrieren (siehe Anhang E Tabelle E29). Bzgl. der Beurteilung von Barrieren unterscheidet sich also nur bei den Studierenden Vignette 1 (drei Barrieren) signifikant. Barrieren werden in dieser Vignette von den Studierenden weniger gut erkannt.

Bzgl. der Items zu gelingenden Situationen zeigen sich in der Gruppe der Bachelorstudierenden signifikante Unterschiede, $\chi^2(3) = 44.468$, $p < .001$, $n = 109$ (siehe Anhang E Tabelle E30). Post-hoc-Tests zeigen, dass sich die Beurteilungen in Vignette 1 und 4 jeweils von Vignette 2 und 3 unterscheiden. Die Items zu gelingenden Situationen werden in Vignette 1 (drei Barrieren) und Vignette 4 (zwei Barrieren) jeweils signifikant zutreffender beurteilt (siehe Anhang E Tabelle E31).

Unterschied Vignette 1 zu Vignette 2: $z = 0.505$, $p = .023$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied Vignette 1 zu Vignette 3: $z = 0.693$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied Vignette 4 zu Vignette 2: $z = -0.821$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied Vignette 4 zu Vignette 3: $z = -1.009$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .10$

Auch in der Gruppe der Masterstudierenden werden Items zu gelingenden Situationen signifikant unterschiedlich beurteilt, $\chi^2(3) = 48.116$, $p < .001$, $n = 66$ (siehe Anhang E Tabelle E32). Post-hoc-Tests zeigen, dass sich Vignette 3 (drei Barrieren) signifikant von allen anderen Vignetten unterscheidet. Items zu gelingenden Situationen werden in dieser Vignette weniger zutreffend beurteilt als in den anderen Vignetten (siehe Anhang E Tabelle E33).

Unterschied zu Vignette 1: $z = 0.970$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .12$

Unterschied zu Vignette 2: $z = 0.758$, $p = .004$, schwacher Effekt nach Cohen (1992), $r = .10$

Unterschied zu Vignette 4: $z = -1.485$, $p < .001$, schwacher Effekt nach Cohen (1992), $r = .18$

Außerdem unterscheidet sich Vignette 4 signifikant von Vignette 2, $z = -.727$, $p = .007$ mit einem schwachen Effekt nach Cohen (1992), $r = .10$. Items zu gelingenden Situationen werden in Vignette 4 (zwei Barrieren) zutreffender als in Vignette 2 (vier Barrieren) beurteilt.

In der Gruppe der Lehrkräfte sind ebenfalls signifikante Unterschiede zu finden, $\chi^2(3) = 10.440$, $p = .015$, $n = 18$ (siehe Anhang E Tabelle E34). Post-hoc-Tests zeigen, dass ein signifikanter Unterschied zwischen Vignette 1 und 4 besteht, $z = -1.333$, $p = .012$. Das entspricht mit $r = .31$ einem mittleren Effekt nach Cohen (siehe Anhang E Tabelle E35). In Vignette 4 (zwei Barrieren) werden Items zu gelingenden Situationen am zutreffendsten beurteilt, in Vignette 1 (drei Barrieren) am wenigsten zutreffend.

Insgesamt werden Items zu gelingenden Situationen also vor allem in Vignette 4 von allen Gruppen zutreffend beurteilt. Die Urteile zu den anderen Vignetten unterscheiden sich aber.

Zuletzt soll geprüft werden, ob die Einschätzung der Klassenatmosphäre die Beurteilungen beeinflusst. Dazu wird innerhalb der Gruppen jeweils die Einschätzung zur Klassenatmosphäre nach Spearman-Rho mit den Mittelwerten der Items zu Barrieren und gelingenden Situationen je Vignette korreliert.

In der Gruppe der Bachelorstudierenden korrelieren in allen vier Vignetten die Items zu Klassenatmosphäre signifikant positiv mit den Items zu gelingenden Situationen und negativ mit den Items zu Barrieren (siehe Anhang E Tabellen E36 bis Tabelle E39).

In der Gruppe der Masterstudierenden sind Korrelationen zwischen der Einschätzung der Klassenatmosphäre und Items zu Barrieren und gelingenden Situationen fast durchgehend zu

finden. In Vignette 2 zeigt sich keine signifikante Korrelation der Klassenatmosphäre mit Items zu gelingenden Situationen und in Vignette 3 keine signifikante Korrelation mit Items zu Barrieren (siehe Anhang E Tabellen E40 bis E43).

In der Gruppe der Lehrkräfte zeigen sich fast keine signifikanten Korrelationen zur Klassenatmosphäre. Lediglich in Vignette 3 korreliert das Item zur Klassenatmosphäre mit den Items zu gelingenden Situationen (siehe Anhang E Tabellen E44 bis E47).

12.5.9.4 Interpretation zum Einfluss durch die Vignettengestaltung und die Einschätzung der Klassenatmosphäre: Das Gelingen der Situation insgesamt. Es finden sich bzgl. der Vignettengestaltung nur punktuell Zusammenhänge zwischen der Anzahl der Barrieren und den Beurteilungen. Lediglich für die Vignette 4, die am wenigsten Barrieren enthält, lässt sich feststellen, dass die Klassenatmosphäre besonders positiv beurteilt wird und Items zu gelingenden Situationen von allen Gruppen besonders zutreffend beurteilt werden. Andere Zusammenhänge zwischen der Anzahl der beschriebenen Barrieren und der Einschätzung der Klassenatmosphäre und der Items zu Barrieren und gelingenden Situationen lassen sich jedoch nicht finden. Das Konstruktionskriterium, wie viele Barrieren innerhalb einer Situation beschrieben werden, wirkt sich also überwiegend nicht auf die Einschätzung der Klassenatmosphäre aus und Hypothese 6.1. muss verworfen werden.

Für die Einschätzung der Situation insgesamt scheinen andere Informationen oder Eindrücke genutzt zu werden. Beispielsweise kommt es in Vignette 1 und 3, in denen die Klassenatmosphäre am schlechtesten eingeschätzt wird, zu Konflikten zwischen Kindern. In einem Fall beschwert sich ein Kind über den Sitznachbarn (*Emma stöhnt: „Oh Mann, Paul macht sich immer so breit, ich hab‘ gar keinen Platz!“*). In einem anderen Fall macht ein Kind darauf aufmerksam, dass ein anderes Kind nicht arbeitet (*Alina: „Ich bin schon ganz weit, Antoni hat noch nicht mal angefangen!“*). Möglicherweise spielen also Annahmen über Beziehungen innerhalb der Klasse eine Rolle dafür, wie die Klassenatmosphäre eingeschätzt wird.

Für die Beurteilung der Items zu Barrieren und gelingenden Situationen wurde angenommen, dass diese differenziert beurteilt werden. Die Konstruktionskriterien sollten dazu führen, dass Merkmale und ihre Ausprägung als Barriere oder gelingend gleich gut erkannt werden – unabhängig davon, wie oft sie in der Vignette insgesamt beschrieben werden. Zwar werden in Vignette 4, die vier gelingende Situationen enthält, Items zu gelingenden Situationen auch besonders zutreffend beurteilt. Ansonsten sind Beurteilungen aber davon unabhängig, wie Barrieren und gelingende Situationen innerhalb der Vignette aufgeteilt sind. Diese Ergebnisse stützen also die Hypothese 6.2.

Betrachtet man innerhalb der Vignetten aber die Korrelationen zwischen den Einschätzungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen, so zeigt sich, dass Beurteilungen nicht differenziert und unabhängig erfolgen. Bereits in der Faktorenanalyse zeigten sich negative Korrelationen zwischen Einschätzungen zu Barrieren und gelingenden Situationen. Betrachtet man dies differenziert für die Gruppen, so zeigt sich durchgehend, außer für die Beurteilungen der Lehrkräfte zu Vignette 3, dass Items zu Barrieren und gelingenden Situationen negativ korrelieren. Von allen Gruppen werden die Vignetten also eher holistisch insgesamt negativ oder insgesamt positiv und nicht differenziert beurteilt. Ob diese Urteile mit einem Gesamteindruck bzgl. der Klassenatmosphäre zusammenhängen, unterscheidet sich in den Gruppen. Während in der Gruppe der Bachelorstudierenden alle Urteile mit der Einschätzung der Klassenatmosphäre korrelieren und in der Gruppe der Masterstudierenden fast alle, zeigt sich in der Gruppe der Lehrkräfte lediglich ein Zusammenhang mit Items zu gelingenden Situationen in Vignette 3.

Studierende scheinen also nach einem Gesamteindruck zu urteilen, der sich in der Einschätzung der Klassenatmosphäre *und* der übrigens Items zeigt. Schätzen sie die Klassenatmosphäre als gut ein, beurteilen sie gelingende Situationen zutreffender und Barrieren weniger zutreffend. Lehrkräfte dagegen scheinen die Items zu Barrieren und gelingenden Situationen weniger abhängig von der Klassenatmosphäre einzuschätzen, fokussieren aber stark die misslingenden Prozesse und übersehen die gelingenden Situationen. Es scheinen also andere Einflüsse als Fähigkeitsausprägungen eine Rolle bei der Beurteilung der Vignetten zu spielen. Diese sind nicht durch Konstruktionsmerkmale der Vignetten zu erklären, sondern haben mit Gesamteindrücken oder Fokussierungen zu tun, die noch zu untersuchen sind.

12.5.9.5 Ergebnisse zur Instruktion: Einfluss der Markierungsaufgabe

Frage 7: Führt die Instruktion der Markierungsaufgabe dazu, dass Beurteilungen zutreffender und differenzierter werden?

Hypothese 7.1: Studierende nutzen die Textinformationen, die bei der Konstruktion der Vignetten gesetzt wurden.

Hypothese 7.2: Die Aufforderung zur Markierung führt im Vergleich zur Instruktion ohne Markierung nicht zu einer genaueren Beurteilung.

Für die Auswertungen wurden jeweils alle Markierungen in den Vignetten ausgezählt und mit den Masterratings verglichen. Dabei wurde ausgewertet, ob (1) die Textteile markiert wurden, die laut Masterrating die jeweils relevante Information tragen sollten (richtige Markierungen) und ob (2) die jeweiligen Items aufgrund der Markierung richtig beurteilt wurden (richtige

Markierungen mit richtiger Beurteilung). Dabei war es möglich, Hinweise auch in mehreren Sinnabschnitten der Vignette zu markieren – während einige Merkmale nur durch einzelne Handlungen repräsentiert wurden, gab es für andere Merkmale mehrere Hinweise (siehe Vignettenlösungen Anhang E E1 bis E4). Für alle Vignetten gilt, dass alle relevanten Textstellen zumindest von einem Teil der Studierenden genutzt wurden und keine der gesetzten Informationen unerkannt blieb. Zusätzlich wurden von den Studierenden noch andere Textstellen markiert, so dass insgesamt mehr Markierungen entstanden als bei der Konstruktion Informationen gesetzt wurden. Die meisten Markierungen wurden durchschnittlich in Vignette 4 vorgenommen und hier war auch der Anteil richtiger Markierungen mit richtiger Beurteilung am höchsten. Am geringsten waren die Anteile jeweils in Vignette 3, wobei hier mehr Markierungen vorgenommen wurden als in den Vignetten 1 und 2 (siehe Tabelle 41).

Tabelle 41

Anzahl markierter Stellen, Anteil richtiger Markierungen, Anteil richtiger Beurteilungen und Anteil falscher Beurteilungen in Bezug zum Mittelwert aller Items für die Vignetten 1-4

Vignette	<i>n</i>	Markierte Stellen gesamt	Im Durchschnitt markierte Stellen	Anteil richtiger Markierungen	Anteil richtiger Markierungen mit richtiger Beurteilung	Anteil falscher Beurteilungen richtig markierter Stellen	Mittelwert aller Items der Vignette
Vignette 1	15	91	6.5	53.9 %	38.5 %	15.4 %	2.68
Vignette 2	20	141	7.1	74.4 %	63.1 %	11.3 %	2.80
Vignette 3	17	124	7.3	51.6 %	33.9 %	17.7 %	2.63
Vignette 4	15	127	8.5	70.9 %	64.6 %	6.3 %	3.00

Anmerkung. *n* = Studierende, die die Vignette mit Markierungen versehen haben.

Es soll außerdem geprüft werden, ob die Instruktion, relevante Textstellen zu markieren, zu einer besseren und differenzierteren Beurteilung führt. Dazu werden je Vignette die Mittelwerte für Items zu Barrieren und Items zu gelingenden Situationen gebildet. Anhand dieser Mittelwerte werden die Studierenden, die Markierungen in der Vignette vorgenommen haben, mit den Masterstudierenden verglichen, die keine Markierungen in der jeweiligen Vignette vorgenommen haben.

Dabei zeigt sich, dass die Bearbeitung der Markierungsaufgabe bei den Vignetten 1 und 4 jeweils mit höheren Mittelwerten für die Items zu Barrieren einhergeht, aber niedrigeren Mittelwerten zu gelingenden Situationen. In Vignette 2 zeigen sich nur geringfügige Unterschiede zwischen den Mittelwerten der Gruppen. Für Vignette 3 erzielten jeweils die Studierenden, die

in der Vignette *keine* Markierungen vorgenommen hatten, höhere Mittelwerte für Items zu Barrieren (siehe Tabelle 42). Die Bearbeitung der Markierungsaufgabe wirkt sich also unterschiedlich aus. Geprüft mit dem Mann-Whitney-U-Test wird dies aber nur für Vignette 4 signifikant. Barrieren werden in Vignette 4 von Studierenden, die Markierungen vorgenommen haben, signifikant zutreffender beurteilt ($z = -2.301, p = .021$), gelingende Situationen signifikant weniger zutreffend ($z = -2.409, p = .016$). Das entspricht für die Barrieren einem schwachen Effekt nach Cohen mit $r = 0.28$ und für die gelingenden Situationen einem mittleren Effekt nach Cohen mit $r = 0.30$ (siehe Anhang E Tabellen E48 bis E51).

Tabelle 42

Mittelwerte der Items zu Barrieren und der Items zu gelingenden Situationen in Abhängigkeit von der Bearbeitung der Markierungsaufgabe

Vignette	Gruppe	n	M	
			Barrieren	gelingend
Vignette 1	Aufgabe bearbeitet	15	2.72	2.64
	Aufgabe nicht bearbeitet	51	2.54	2.82
Vignette 2	Aufgabe bearbeitet	20	2.89	2.68
	Aufgabe nicht bearbeitet	46	2.88	2.60
Vignette 3	Aufgabe bearbeitet	17	2.80	2.14
	Aufgabe nicht bearbeitet	49	3.02	2.35
Vignette 4	Aufgabe bearbeitet	15	3.21*	2.71*
	Aufgabe nicht bearbeitet	51	2.79*	3.07*

Anmerkung. Werte, die sich auf dem 5 %-Niveau signifikant voneinander unterscheiden sind mit * markiert.

12.5.9.6 Interpretation zur Instruktion: Einfluss der Markierungsaufgabe. Die Auswertung zeigt, dass die Studierenden bei der Beurteilung der Vignetten Bezug auf die Textstellen nehmen, die gemäß Konstruktion die relevanten Informationen enthalten. Dies spricht dafür, dass die Konstruktion der Vignetten geeignet ist, die gewünschten Beurteilungsprozesse auszulösen, was Hypothese 7.1 stützt. Es werden jedoch auch andere Informationen aus der Vignette genutzt, so dass auch die Situation insgesamt die Beurteilungen beeinflusst. Da keine isolierten Handlungen beschrieben werden, sondern Handlungen im Zusammenhang, ist dies jedoch naheliegend und nicht zu vermeiden. Wie viele Markierungen insgesamt vorgenommen werden, schwankt zwischen den Vignetten. Das heißt, dass die Vignetten über die intendierten Informa-

tionen hinaus unterschiedlich viel Informationsgehalt haben. Dies scheint aber keinen unmittelbaren Einfluss darauf zu haben, (1) wie viele richtige Markierungen erfolgen und (2) wie diese zur richtigen Beurteilung der Items beitragen. So zeigen sich beispielsweise für Vignette 2 nur die drittmeisten Markierungen insgesamt, der Anteil richtiger Markierungen ist aber am höchsten. Ob die richtig markierten Stellen dann auch richtig beurteilt werden, hängt mit der jeweiligen Schwierigkeit zusammen. Richtige Markierungen mit richtiger Interpretation sagen den Schwierigkeitsgrad der Vignette insgesamt voraus, der demnach durch die intendierten Informationen zustande kommt und nicht durch andere, zufällige Informationen in der Vignette.

Es zeigt sich auch, dass ein Bemerkten der relevanten Stellen allein nicht ausreicht. So werden zum Teil richtig markierte Stellen falsch beurteilt. Jedoch zeigt sich ein starker Zusammenhang: Werden relevante Stellen besonders häufig bemerkt (Vignette 2 und 4), dann werden sie auch häufiger richtig beurteilt. Ein geringes Erkennen von Barrieren und gelingenden Situationen hängt also auch damit zusammen, dass relevante Handlungen erst gar nicht bemerkt oder nicht als relevante Information bewertet werden.

Die Aufforderung, Textstellen zu markieren, die zur Beurteilung herangezogen werden, zeigt nur bei Vignette 4 einen signifikanten Zusammenhang mit den Beurteilungen. Barrieren werden besser gesehen und gelingende Situationen weniger zutreffend beurteilt. Dies ist insofern interessant, da diese Vignette nur zwei Barrieren enthält und die gelingenden Momente mit vier Merkmalen überwiegen. Möglicherweise fallen also die wenigen Barrieren stärker auf, wenn die entsprechenden Informationen genauer betrachtet werden. Ein Hinweis dazu könnte auch der höhere Mittelwert für gelingende Situationen in Vignette 2 sein, der von denjenigen Studierenden erzielt wurde, die die Vignette markiert haben. In dieser Vignette waren die gelingenden Situationen mit nur zwei Merkmalen in der Unterzahl. Allerdings ist dieser Effekt nicht signifikant. Die Hypothese 7.2 wird also gestützt, da in keiner Vignette die Variation der Instruktion zu einer differenzierten Beurteilung und damit zu höheren Mittelwerten für Barrieren und gelingende Situationen geführt hat. Trotz der Aufforderung zum genauen Hinsehen, die mit der Markierung verbunden ist, scheinen also Halo-Effekte zu wirken. Die Instruktion liefert somit auch keine Erklärung für die undifferenzierten Urteile.

Zusammengefasst werden relevante Textstellen gefunden und genutzt, so dass die Urteile mit der Vignettenkonstruktion zusammenhängen. Aber es wirken auch andere Faktoren auf das Erkennen der Barrieren und gelingenden Situationen, die noch zu bestimmen sind.

12.6 Diskussion

Durch den mehrschrittigen Konstruktions- und Überarbeitungsprozess sind vier Vignetten entstanden, die potenzielle Lernbarrieren sowie gelungene Lehr-Lernprozesse in unterschiedlichen Unterrichtsfächern und -situationen darstellen. Sie werden von Studierenden und Lehrkräften als authentisch und gut vorstellbar eingeschätzt und können die Fokussierung auf KF und LU sowie Beurteilungsprozesse hinsichtlich der mehr oder weniger gelungenen Umsetzung auslösen.

Die mittleren Schwierigkeitsgrade und Streuungen der Items sind ausreichend, um unterschiedliche Beurteilungen abbilden zu können. Es wird bereits durch die deskriptive Beschreibung der Items deutlich, dass Items zu Barrieren etwas leichter zu beantworten sind und Items zu gelingenden Situationen höhere Schwierigkeitsgrade aufweisen. Die Reliabilitätsprüfung zeigt, dass die jeweiligen Items keine gemeinsame Skala bilden. Bestätigt wird dies durch die durchgehend negativen Korrelationen zwischen Items zu Barrieren und gelingenden Situationen in der Faktorenanalyse. Zudem zeigen sich Testlets für die einzelnen Vignetten. Die Konstruktionsmerkmale der verschiedenen Vignetten und Ausprägungen hinsichtlich des Gelingens der Situationen spiegeln sich in der Beurteilung wider, so dass Urteile nicht zufällig oder unabhängig von den Vignetten sind, sondern durch diese ausgelöst werden. Die Auswertung der Markierungsaufgabe zeigt zudem, dass die gezielt gesetzten Informationen in den Vignetten auch für die Beurteilung genutzt werden, so dass die Konstruktion der Vignetten grundsätzlich als gelungen bezeichnet werden kann.

Gleichzeitig zeigt sich aber, dass gerade Studierende auch andere Informationen zur Beurteilung nutzen und diese zu Gesamteindrücken führen, die die Beurteilung beeinflussen. Lehrkräfte scheinen ihre Beurteilung der Lehr-Lernsituation zwar von einer mehr oder weniger positiven Einschätzung der Klassenatmosphäre lösen zu können, sehen aber vor allem misslingende Prozesse und beurteilen dann die Situation insgesamt als eher negativ.

Im Hinblick auf die Anforderungen in der Praxis, Lehr-Lernsituationen differenziert beurteilen zu können, muss Expertise an die Fähigkeit geknüpft sein, misslingende und gelingende Prozesse voneinander unterscheiden und jeweils zutreffend beurteilen zu können. Die Ergebnisse zeigen jedoch, dass diese Fähigkeit in keiner der Gruppen hoch ausgeprägt ist.

Dies weist darauf hin, dass die zu erfassende Fähigkeit sehr spezifisch und mit zunehmender Praxiserfahrung nicht automatisch vorhanden ist. Auch die spezifische Lerngelegenheit für die Masterstudierenden führt nicht zu einer gleichermaßen guten Beurteilung der Barrieren und gelingenden Situationen. Hier ist allerdings zu beachten, dass die Lerngelegenheit auf eine Seminarsitzung und nur auf Wissenserwerb begrenzt war. Dennoch lassen sich Unterschiede zwi-

schen Bachelor- und Masterstudierenden feststellen, die unterschiedliche Fähigkeitsausprägungen vermuten lassen. Insgesamt zeigt das Verfahren eine Sensitivität für unterschiedliche Fokussierungen und Beurteilungen, die sich noch nicht als Fähigkeitsausprägungen der modellierten Kompetenz interpretieren lassen. Es wird deutlich, dass Lehrkräfte stärker einseitig negativ beurteilen, während Bachelorstudierende gelingende Situationen etwas deutlicher erkennen. Dies ist auch abhängig von den einzelnen beschriebenen Situationen und die Beurteilungen fallen nicht über alle Vignetten hinweg gleich aus. Von Konstruktionskriterien wie der Anzahl beschriebener Barrieren pro Vignette scheint dies in allen Gruppen unabhängig zu sein. Für die Studierenden konnte gezeigt werden, dass die Einschätzung der Klassenatmosphäre insgesamt einen Zusammenhang mit den Urteilen zeigt. Bei den Lehrkräften ist das nicht der Fall. Welche Faktoren in den Vignetten zur Einschätzung der Klassenatmosphäre führen und bei den Studierenden dann auch die Urteile beeinflussen, lässt sich bislang noch nicht näher bestimmen.

Welche Bedeutung die Befunde der verschiedenen Entwicklungsschritte für die Entwicklung und den Einsatz textvignettenbasierter Verfahren für das Erkennen von Lernbarrieren mit sich bringen, soll im abschließenden Kapitel betrachtet werden.

13 Kritische Reflexion und Ausblick

In der abschließenden Betrachtung werden theoretische Grundlagen, Entwicklungsschritte und Ergebnisse der Arbeit kritisch reflektiert. Limitationen ergeben sich bzgl. der theoretischen Grundlagen aus dem verwendeten Barrierebegriff und der Modellierung der zu erfassenden Kompetenz. Weitere Limitationen ergeben sich in Bezug zur Verfahrensentwicklung aus der Vignettenkonstruktion, dem methodischen Vorgehen und der Wahl der Stichprobe. Trotz zu benennender Limitationen liefert die Arbeit Erkenntnisse, die für die Entwicklung vignettenbasierter Verfahren im Allgemeinen und für die Erfassung der Kompetenz *Erkennen von Lernbarrieren* im Besonderen bedeutsam sind. Diese werden als zentrale Ergebnisse vorgestellt. Der Ausblick skizziert Fragstellungen, die sich aus den zentralen Erkenntnissen dieser Arbeit ergeben.

13.1 Der Barrierebegriff

Kapitel 2 legt dar, dass Lernbarrieren bislang nicht systematisch und umfassend dargestellt wurden. Zwar besteht Einigkeit darüber, dass Bedingung von Lehr-Lernsituationen in einem komplexen Gefüge von Kind und Umfeld zu suchen sind. Barrieren werden jedoch meist an individuellen Bedingungen des Kindes festgemacht. Eine Systematisierung von Lernbarrieren im Sinne fehlender Gelingensbedingungen in der Unterrichtsgestaltung fehlte bislang.

Ausgehend von grundlegenden Studien, die Merkmale guten Unterrichts darstellen (siehe Kapitel 6), wird die hohe Relevanz gelingender KF und LU für die Unterrichtsqualität deutlich. Durch die Zusammenführung unterschiedlicher Konzeptualisierungen werden Merkmale der KF und LU benannt, anhand derer sich wesentliche Gelingensbedingungen qualitätvollen Unterrichts beschreiben lassen. Zur Beschreibung der KF werden in dieser Arbeit in Anlehnung an Holodynski et al. (2017) die folgenden Merkmale operationalisiert und genutzt: (1) die Allgegenwärtigkeit der Lehrkraft, die auch die passende Reaktion auf Störungen beinhaltet, (2) die Berücksichtigung von räumlichen Bedingungen, (3) die reibungslose Gestaltung von Übergängen sowie (4) die Etablierung und Einhaltung von Regeln und Routinen. Die LU wird repräsentiert durch die Merkmale (5) inhaltliche Klarheit, (6) Zielklarheit und (7) Veranschaulichung. Das Merkmal (8) Feedback lässt sich im Sinne positiver Präsenz der KF zuordnen, je nach inhaltlicher Gestaltung aber auch der LU. Da dies als generischer und nicht als fachdidaktischer Ansatz konzipiert wird, werden Merkmale der kognitiven Aktivierung nicht berücksichtigt, da diese ohne fachlichen Kontext schwer zu beschreiben sind.

Werden Maßnahmen, die zum Gelingen dieser Merkmale beitragen können, nicht oder nicht erfolgreich umgesetzt, so entstehen potenzielle Lernbarrieren, die den Lehr-Lernprozess für einzelne oder mehrere Kinder behindern können.

Die Operationalisierung der entworfenen Lernbarrieren repräsentiert eine Auswahl von Aspekten, deren Bedeutsamkeit für den Unterricht empirisch gut belegt ist. Die im Laufe der Verfahrensentwicklung geführten Expertinnen- und Expertendiskussion zeigen zudem, dass der verwendete Barrierebegriff gut nachvollziehbar ist. Er betont in jedem Fall die Notwendigkeit, eine kindzentrierte Vorgehensweise zu erweitern, ohne diese ganz auszuschließen.

Limitationen ergeben sich aus der Festlegung auf die verwendete Konzeptualisierung sowie die Auswahl und erfolgte Operationalisierung bestimmter Merkmale. Andere Konzeptualisierungen würden auch die Verwendung anderer bzw. weiterer Merkmale ermöglichen sowie die Berücksichtigung fachspezifischer Aspekte. Gerade aus sonderpädagogischer Perspektive ergeben sich aus individuellen Bedingungen besondere Bedürfnisse, die auch die Gestaltung der KF und LU betreffen. Zwar wird bei der Definition des Barrierebegriffs die Betonung darauf gelegt, dass Hindernisse in Interaktionen entstehen. Dies wird aber zunächst einseitig durch die Gestaltung von Unterrichtssituationen dargestellt, ohne solche individuellen Bedingungen zu beschreiben. Die erfolgte Operationalisierung von Lernbarrieren im Sinne mangelhaft umgesetzter KF und LU ist also ein notwendiger, aber nicht hinreichender Schritt, um problematische Lehr-Lernsituationen zu analysieren.

13.2 Die Kompetenz des Erkennens von Lernbarrieren

Das Kompetenzmodell von Blömeke et al. (2015a) liefert mit seinen einzelnen Komponenten einen geeigneten Rahmen, die Kompetenz des Erkennens von Lernbarrieren zu modellieren. Affektiv-motivationale Merkmale und professionelles Wissen nehmen als Dispositionen Einfluss auf den Wahrnehmungsprozess und die wissensgeleitete Interpretation der wahrgenommenen Aspekte. In Bezug auf das Erkennen von Lernbarrieren sind spezifische Wissensanteile und Annahmen über die Ursachen von Lernbarrieren notwendig sowie eine daraus resultierende Fokussierung auf Unterrichtsbedingungen. Die Kompetenz des Erkennens umfasst dann sowohl die Wahrnehmung relevanter Interaktionen als auch die Interpretation hinsichtlich der potenziell hindernden Auswirkung. Die hier modellierte Fähigkeit des Erkennens ist also anspruchsvoll und komplex, da sie mehrere Wissensanteile, eine spezifische Fokussierung sowie beide situationsspezifischen Teilkompetenzen (Wahrnehmen *und* Interpretieren) umfasst. Andere Studien zeigen, dass sich Expertiseunterschiede gerade im Schritt der wissensgeleiteten Interpretation zeigen und Novizen und Novizinnen zwar über Wissen verfügen und relevante Situationen bemerken, diese aber nur beschreiben und eine theoretische Fundierung noch fehlt (vgl. z. B. Barth, 2017; Blömeke et al., 2015b; Orschulik, 2021; Schäfer & Seidel, 2015). Eine Limitation besteht darin, dass die Erfassung der komplexen Fähigkeit des Erkennens eine solche Differenzierung nicht berücksichtigt.

Eine weitere Limitation ergibt sich daraus, dass spezifische Dispositionen, die für das Erkennen von Lernbarrieren notwendig sind, in der vorliegenden Studie nicht kontrolliert wurden. Zum einen kann das, was sichtbar ist, nicht vorausgesetzt werden, da die Wahrnehmung der Lernbarrieren auf einer jeweils individuellen Rekonstruktion der in der Vignette geschilderten Situationen beruht (Decristan et al., 2020; Fauth et al., 2020; König, 2015). Lesefähigkeit, die Fähigkeit zur Rekonstruktion der Situation oder auch die Erfahrung mit ähnlichen Aufgabenformaten zählen zu Dispositionen, die im Rahmen dieser Studie nicht erfasst wurden.

Die Fähigkeit des Erkennens wird in dieser Arbeit als *diagnostische* Kompetenz verstanden. Die dafür notwendige Teilkompetenz der Fokussierung auf Unterrichtsbedingungen und Interaktionen ist in zweierlei Hinsicht denkbar: (1) Spezifische Dispositionen können dazu führen, dass eine entsprechende Fokussierung spontan erfolgt, so dass Lernbarrieren in spontanen Beobachtungen von Unterrichtssituationen erkennbar sind. Dies entspricht der Erfassungssituation in der vorliegenden Arbeit. Probandinnen und Probanden werden dazu aufgefordert, die in den Vignetten beschriebenen Lehr-Lernsituationen spontan zu beurteilen. Welche spezifischen Dispositionen jeweils vorhanden sind und ob oder wie diese genutzt werden, wird in dieser Studie aber nicht erfasst. (2) Als geeignete diagnostische Methode zur Analyse von Lehr-Lernsituationen gilt vor allem die systematische Beobachtung. Diese erfordert eine Planung und Vorbereitung, die die Spezifizierung und gezielte Anwendung von Beobachtungskategorien umfasst (Ingenkamp & Lissmann, 2008; Karst, 2017; Kopmann, 2022; Ricken, 2017). Eine solche systematische Beobachtung wird durch die in dieser Arbeit verwendeten Impulse aber nicht initiiert und es finden keine Aktivierung im Hinblick auf mögliche spezifische Dispositionen statt. Aktuelle Erkenntnisse (Wenck et al., im Druck) zeigen, dass ein gutes Erkennen von Lernbarrieren aber gerade in der spezifischen und systematischen Vorbereitung begründet ist. Wie und welche Dispositionen Einfluss auf das Erkennen nehmen, kann durch die vorliegenden Ergebnisse noch nicht beantwortet werden. Der Aspekt der spezifischen Kompetenz wird im Folgenden auch in Bezug auf die Wahl der Stichprobe und die Formulierung der Instruktion diskutiert.

13.3 Die Vignettenkonstruktion

Für die Konstruktion von Vignetten wurden verschiedene Aspekte abgewogen. Diese betreffen u. a. das Format, die Komplexität und die Darstellung der Lernbarrieren. In vielen Studien (siehe Kapitel 4) werden Videos verwendet, um Authentizität und ökologische Validität zu erreichen. Die Komplexität und Synchronizität von Unterricht sollen auf diese Weise nah an der Realität abgebildet werden. Bei der Darstellung von Lernbarrieren spricht aber Vieles für die Verwendung von Textvignetten: Authentische Unterrichtsbeispiele können so abgewandelt und dargestellt werden, dass auch misslingende Unterrichtsszenen ohne Bloßstellung der agierenden

Personen möglich sind. Die Art, Anzahl und Deutlichkeit von Barrieren können dabei gezielt konstruiert und im Entwicklungsprozess variiert werden. Dies ist im Vergleich zu inszenierten Videos wesentlich ökonomischer möglich. Zudem sind Informationen, die durch die Textvignette gegeben werden, besser zu kontrollieren als in Videos, da in Videos eine Vielzahl von nonverbalen Informationen transportiert werden, die in Textvignetten keine Rolle spielen, wie z. B. Aussehen und Stimme von Personen, Körpersprache oder Informationen zur Raumgestaltung. Zum einen lassen sich dadurch ungewollte Einflüsse verringern, zum anderen entsteht auch eine geringere Anforderung an die kognitive Kapazität, weil sich Wahrnehmungsprozesse besser auf wesentliche Merkmale fokussieren lassen.

Gleichzeitig bringen diese Aspekte aber auch Limitationen mit sich. So könnte die Anforderung durch diese Reduktion ökologisch weniger valide sein, da die Wahrnehmung gleichzeitig auftretender Merkmale bei der Bearbeitung von Textvignetten nicht möglich und nicht notwendig ist, weil diese im Text linear dargestellt werden. Zudem müssen alle Merkmale mit Worten beschrieben werden, die im Video visuell wahrnehmbar wären. Beispielsweise sind räumliche Bedingungen wie Lichtverhältnisse oder der Sitzplätze im Video wahrnehmbar, im Text muss sprachlich darauf verwiesen werden. Zwar müssen diese Informationen dennoch als relevant erkannt und hinsichtlich ihrer möglichen Auswirkung interpretiert werden. Es kann jedoch auf diese Weise nicht erfasst werden, ob die Information auch neben einer Vielzahl anderer gleichzeitiger Informationen und Eindrücke ohne expliziten Verweis darauf wahrgenommen worden wäre.

Die Konstruktion der Vignetten erfolgte in Bezug auf die Verteilung der Lernbarrieren sehr systematisch. Auch die Situierung bzgl. Klassenstufe, Unterrichtsfach, Arbeitsphase und handelnden Personen wurde bedacht und bewusst variiert. Durch die gleichmäßige Verteilung von Unterrichtsmerkmalen bei gleichzeitiger Variation der Kontextinformationen sollte ermöglicht werden, eine vom Kontext unabhängige Fähigkeitsausprägung zur Erkennung von Lernbarrieren sichtbar zu machen. Ergebnisse dazu werden in Kapitel 13.6. diskutiert. Bzgl. der Vignettenkonstruktion ist kritisch anzumerken, dass – obwohl ein generischer Ansatz verfolgt werden sollte – dennoch Angaben zur fachlichen Situierung gemacht wurden. Andere Studien (vgl. z. B. Hoppe et al., 2020; Steffensky et al., 2015; Sunder et al., 2016) zeigen aber, dass die PW von der fachlichen und didaktisch-methodischen Situierung beeinflusst wird, so dass das Erkennen von Lernbarrieren von diesen Angaben möglicherweise beeinflusst wird, ohne dass dies in dieser Arbeit erfasst wird.

Andere Kontextinformationen zur Lernsituation wurden hingegen nicht zur Verfügung gestellt. Die in dieser Arbeit verwendeten Vignetten beschränken sich auf die Beschreibung von konkreten Lehr-Lernsituationen. Die Interpretation der beschriebenen Merkmale der KF und der

LU muss allein aufgrund der beschriebenen Informationen erfolgen, ohne beispielsweise individuelle Bedingungen, Beziehungen oder vorangegangene Unterrichtssituationen zu berücksichtigen. Helmke (2006) weist darauf hin, dass die Betrachtung einzelner Merkmale nicht hinreichend ist, um die Erreichung verschiedener Zielkriterien vorherzusagen. Vielmehr sind diese durch ein komplexes und wechselseitiges Angebots-Nutzungs-Modell begründet. Unterrichtsmerkmale sind außerdem nicht so zu verstehen, dass immer eine maximale Ausprägung aller Merkmale als optimal zu betrachten ist, vielmehr führt die richtige Passung zu qualitativem Unterricht und die positive Umsetzung einiger Merkmale kann fehlende Merkmale kompensieren (Helmke, 2000; Helmke, 2006; Lipowsky et al., 2005; Reusser & Pauli, 2010). Die Bearbeitung der Vignetten erfordert einerseits eine Konzentration auf beobachtbare Handlungen und die Fokussierung auf eben die Merkmale, deren Erkennung erfasst werden soll. Allerdings werden dabei solche komplexen Wirkmechanismen außer Acht gelassen.

Letztlich kann aber ein vignettenbasiertes Verfahren zur Erfassung der PW nur auf Ausschnitten beruhen und nicht alle Bedingungen berücksichtigen, die in realen Lehr-Lernprozessen eine Rolle spielen. Die dargestellten Lernbarrieren sind demnach (nur) als *potenzielle Barrieren* zu verstehen.

13.4 Der methodische Ansatz

Die methodischen Abwägungen galten unter anderem der Wahl zwischen einem qualitativen und einem quantitativen Ansatz. Qualitative Ansätze bieten vor allem die Möglichkeit, subjektive Theorien und Deutungen zu erfassen, ohne eine Interpretationsrichtung vorzugeben (Baer & Buholzer, 2005; Brovelli et al., 2013; van Es & Sherin, 2008; van Es & Sherin, 2010; Pisarek & Schilcher, 2015; Santaga et al., 2007; Schnell et al., 2018). So können verwendete Referenzsysteme sichtbar gemacht werden und bzgl. der PW Fokussierungen erfasst werden, die nicht durch das Aufgabenformat vorgegeben werden (Schnell et al., 2018). Für eine objektive und zeitökonomische Datenerfassung und -auswertung eignen sich allerdings quantitative Formate besser (Gold et al., 2013; Kersting, 2008; Rutsch et al., 2018a; Seidel & Stürmer, 2014; Tepner & Dollny, 2014).

Für die Entwicklung des Verfahrens ErLe wurden im ersten Schritt beide Ansätze erprobt. In der weiteren Entwicklung wurde der quantitative Ansatz verfolgt, weil sich in der Erprobung zeigte, dass Studierende im offenen Antwortformat kaum Lernbarrieren in Unterrichtsbedingungen fokussierten und die Inter-Rater-Reliabilität nicht zufriedenstellend war.

Die Formulierung von Items bringt das Problem mit sich, dass auf die zu erfassenden Merkmale sprachlich verwiesen werden muss. Der Schwerpunkt der erfassten Fähigkeiten liegt damit

weniger auf der spontanen Wahrnehmung relevanter Situationen, sondern es wird durch die direkte Abfrage eher erfasst, wie in der Vignette enthaltene Informationen zur Beurteilung genutzt werden (vgl. Barth, 2017; Brovelli et al., 2013). Mit dem Item *Die Lehrkraft gibt konstruktives Feedback* wird beispielsweise ein Urteil zum Feedback direkt abgefragt, ohne dass erfasst wird, ob das Feedback spontan wahrgenommen worden wäre.

Eine weitere Limitation ergibt sich daraus, dass in der Konstruktion der Vignette und der Items und bei deren Auswertung Merkmale genutzt werden, die theoretisch getrennt werden. In der Lehr-Lernsituation sind Merkmale der KF und LU jedoch nicht immer trennbar und somit auch nicht unbedingt getrennt beurteilbar (vgl. Gold & Holodynski, 2017; Piwowar, 2013). So greift beispielsweise ineinander, wenn durch eingeübte Regeln und Routinen ein Übergang gut gelingt oder wenn die Herstellung inhaltlicher Klarheit bzgl. einer Schüleräußerung unmittelbar mit einem Feedback verbunden ist. Beurteilungen, die sich in der Auswertung auf ein spezifisches Merkmal beziehen, wurden also möglicherweise aufgrund mehrerer Merkmale generiert. Eine Limitation ergibt sich daraus für die Aussagekraft der Interpretationen der Ergebnisse, und zwar immer dann, wenn sich Aussagen auf einzelne Merkmale beziehen, wie z. B. Schwierigkeitsgrade oder Rangfolgen für einzelne Merkmale.

Des Weiteren ist die Instruktion kritisch zu überdenken. In den ersten Entwicklungsschritten zeigte sich, dass Studierende Bedenken äußerten. Anhand eines kurzen Ausschnitts seien Beurteilungen nicht gut möglich und es seien mehr Informationen nötig. Die Instruktion wurde deshalb dahingehend geändert, dass sie zu einer spontanen (und insofern auch realitätsnahen) Beurteilung eines einzelnen Handlungsabschnitts motivieren sollte: *Beziehen Sie sich bei der Beurteilung der Items allein auf die Informationen aus dem Beispiel. Es geht darum, was Ihnen auf den ersten Blick aufgefallen ist und welche Schlüsse Sie ziehen, ohne dass Sie Kontextinformationen kennen.* Dies bewirkt aber möglicherweise, dass nicht differenziert beurteilt wird, sondern eher ein Eindruck gebildet wird. Eine solche Eindrucksbildung erfolgt nach Erkenntnissen der Wahrnehmungspsychologie aber zum einen holistisch und wird zum anderen durch markante Merkmale im Sinne eines Halo-Effekts sowie persistierende erste Eindrücke (*primacy effect*) geprägt, während andere Merkmale vernachlässigt werden (Stang & Urhahne, 2016; Timmermans et al., 2021; Werth et al., 2020).

Desweiteren wurde die Bewertung der Antworten mittels Punktevergabe so vorgenommen, dass jeweils von der Autorin auf Grundlage der Vignettenuvalidierung bestimmt wurde, was als richtig oder falsch zu werten war. Für die Antworten wurden Werte von 1 bis 4 vergeben, wobei jeweils die Extremwerte (*trifft völlig zu* bzw. *trifft gar nicht zu*) den höchsten Wert erhielten. Eindeutige Urteile führten also zu höheren Werten und damit auch zu höheren Testscores.

Dies beruhte auf der Annahme, dass zum einen Barrieren in den Vignetten deutlich ausgeprägt waren und zum anderen, dass ein sicheres Erkennen mit einem eindeutigen Urteil einhergeht. Daraus ergeben sich zwei Limitationen: (1) Die Vignetten wurden zwar in Expertinnen- und Expertendiskussionen validiert, aber nicht durch Expertinnen- und Expertenratings geprüft. Es bleibt also unklar, ob wirklich alle Merkmale in voller Ausprägung in den Vignetten enthalten sind. (2) Urteile am äußeren Rand der Likert-Skala wurden als sichere Urteile im Sinn von *Barriere bzw. gelingende Situation sicher erkannt* oder *Barriere bzw. gelingende Situation nicht erkannt* interpretiert. Dies lässt mögliche Urteilermerkmale wie z. B. eine Tendenz zur Mitte außer Acht und es bleibt unklar, ob und wie Differenzierungen, die durch die Likert-Skalen vorgegeben wurden, ausgenutzt wurden.

13.5 Die Stichprobe

Zur Erprobung des Verfahrens wurden als Stichprobe Studierende unterschiedlicher Fachsemester und Studiengänge gewählt sowie Lehrkräfte. Es wurde angenommen, dass sowohl unterschiedliche Schwerpunktsetzungen im Studium als auch mehr Lerngelegenheiten im Laufe des Studiums bzw. Praxiserfahrung im Unterricht unterschiedliche Expertise im Erkennen von Lernbarrieren bedingen würden. Diese Unterscheidungen sowie der geringe Rücklauf bei der Befragung der Lehrkräfte bedingten unterschiedlich große und zum Teil kleine Teilstichproben. Trotz der Verwendung non-parametrischer Verfahren begrenzt dies die Aussagekraft der Ergebnisse.

Zwar konnten Unterschiede zwischen den Teilstichproben gezeigt werden, diese waren aber nicht sehr ausgeprägt und nicht konsistent festzustellen. Bei der Modellierung des Kompetenzmodells wurde angenommen, dass für das Erkennen von Lernbarrieren spezifische Dispositionen sowie eine gezielte Fokussierung von Unterrichtsbedingungen notwendig sind. Eine besondere Bedeutung kommt der situationsspezifischen Fähigkeit des Erkennens zu – über Dispositionen hinaus ist also das Können relevant. Da die spezifische Beobachtung von Unterrichtsmerkmalen zum Zeitpunkt der Untersuchung an der Universität Hamburg nicht in den Modulen verankert war, ist diese Expertise möglicherweise aber nicht genug ausgeprägt und die Stichprobe deshalb nur mäßig geeignet gewesen. Jüngere Untersuchungen zeigen, dass ohne sehr spezifische Lerngelegenheiten und intensives Üben Barrieren in Unterrichtsmerkmalen sowohl von Studierenden als auch von Lehrkräften mit Berufserfahrung kaum fokussiert und erkannt werden (Tönnissen et al., 2023; Wenck et al., im Druck). Die Stichprobenwahl und die gewählte Abstufung der Expertise sind deshalb als Limitation zu betrachten.

13.6 Die zentralen Ergebnisse

Trotz der genannten Limitationen ergeben sich aus der Arbeit wichtige Erkenntnisse und Implikationen für die weitere Forschung. Diese betreffen die Akzeptanz und Bearbeitbarkeit des gewählten Formats, den gewählten Fokus auf Unterrichtsmerkmale, die Spezifität der Impulse durch Vignetten und Items, die Kombination mehrerer Vignetten und die Beurteilung sowohl misslingender als auch gelingender Situationen, die Sensitivität des Verfahrensentwurfs, die Kompetenz des Erkennens von Lernbarrieren sowie unterschiedliche Einflüsse auf die Beurteilung von Unterrichtssituationen.

Eine erste Voraussetzung für die Arbeit mit einem vignettenbasierten Verfahren ist die Akzeptanz seitens der Probandinnen und Probanden. In allen Entwicklungsschritten zeigt sich, dass die Bearbeitung der Vignetten als interessant, angemessen schwierig und verständlich empfunden wird. Auch die geringe Anzahl an Abbrüchen bei der Bearbeitung lässt auf eine gute Akzeptanz und Bearbeitbarkeit schließen. Die Vignetten werden als gut vorstellbar und authentisch beurteilt. *Textvignettenbasierte Verfahren mit geschlossenen Fragen sind demnach ein gut akzeptierter Ansatz.*

Der Vergleich der Ansätze mit unterschiedlichen Vignetten und offenen bzw. geschlossenen Frageformaten im ersten Entwicklungsschritt zeigt, dass sich Studierende zum einen inhaltlich spontan kaum auf Unterrichtsmerkmale beziehen, wenn sie eine problematische Unterrichtssituation beurteilen. Es werden vor allem Überlegungen zu den individuellen Bedingungen des Kindes geäußert. Dies steht im Einklang mit Befunden anderer Studien, die zeigen, dass diagnostische Ansätze eher kindzentriert erfolgen oder dass Bedingungen oder Handlungen isoliert betrachtet werden und weniger die Auswirkungen von Interaktionen (Kopmann, 2022; Thiel, 2021; Tönnissen et al., 2023). *Umso notwendiger erscheint eine spezifische Hinwendung zur Fokussierung von Lernbarrieren in Unterrichtsbedingungen.*

Zum anderen bleiben Kommentare oft auf einer eher nacherzählenden Ebene. So lässt sich zwar nachvollziehen, auf welche Handlungen Bezug genommen wird, also das Wahrnehmen bestimmter Handlungen. Eine theoriegeleitete Interpretation hinsichtlich der Auswirkung erfolgt spontan nur selten. Die Fähigkeit des Erkennens, die beide Schritte erfordert, ist deshalb auf diesem Weg zu wenig abbildbar. *Die spezifische Erfassung der Kompetenz Erkennen von Lernbarrieren erfordert also eine spezifische Abfrage, die durch die konstruierten Vignetten und Items erzielt werden konnte.*

Der im 2. Entwicklungsschritt verfolgte quantitative Ansatz bringt die Schwierigkeit mit sich, dass durch die Formulierung der Items auf relevante Aspekte verwiesen wird. Bei einer insgesamt eher misslingenden Unterrichtssituation ist die Beurteilung der Items als negativ deshalb

relativ einfach. Die Entwicklung weiterer Vignetten (3. Entwicklungsschritt) zeigt, dass zum einen der Umsetzung, wie Merkmale der KF und LU in der Vignette dargestellt werden, eine entscheidende Bedeutung zukommt. Das konkrete Auftreten einzelner Lernbarrieren in der jeweiligen Situation bestimmt letztlich den Schwierigkeitsgrad, so dass nur eine Vervielfältigung von Vignetten und Items ein repräsentatives Bild ergeben kann. *Die situationsspezifische Fähigkeit Lernbarrieren zu erkennen kann also nicht stellvertretend durch eine beispielhafte Vignette erfasst werden, sondern nur durch mehrere, unterschiedliche Vignetten.* Durch eine Erfassung mit einem Vignettenpool können mittlere Schwierigkeitsgrade und eine ausreichende Streuung der Items erreicht werden, so dass sich unterschiedliche Beurteilungen abbilden.

Zum anderen zeigt sich, dass das ausschließliche Auftreten von Barrieren zu einer eher einheitlich negativen Beurteilung führt. *Erst die Verwendung von gelingenden und misslingenden Situationen ermöglicht eine differenzierte Erfassung, welche Barrieren wirklich in der Unterscheidung zu gelingenden Situationen erkannt werden.*

Die Ergebnisse der 3. und 4. Datenerhebung machen deutlich, dass ein Verfahren mit mehreren Vignetten und sowohl Lernbarrieren als auch gelingenden Situationen unterschiedliche Beurteilungen erfassen kann. *Es ist also grundsätzlich von einer Sensitivität des Verfahrens für unterschiedliche Beurteilungen auszugehen.* Es zeigt sich aber auch, dass die Unterschiede nicht als unterschiedliche Expertise im Erkennen von Lernbarrieren interpretiert werden können, weil eben diese differenzierte Beurteilung nicht erfolgt. Die Reliabilitätsprüfungen und Faktorenanalysen zeigen, dass die jeweiligen Items zu Barrieren und gelingenden Situationen keine gemeinsame Skala bilden. Neben Erkenntnissen zur Verfahrenskonstruktion lassen sich also auch Einsichten in die Art und Weise der Beurteilungen von Lehr-Lernsituationen gewinnen. Diese scheinen innerhalb der Stichprobe eher holistisch zu erfolgen und zeigen, dass entweder insgesamt negativ oder positiv beurteilt wird. Bei der Urteilsbildung sind also vorwiegend Lernbarrieren oder vorwiegend gelingende Situationen ausschlaggebend für die Beurteilungen. Eine hohe Fähigkeitsausprägung im Erkennen von Lernbarrieren sollte dagegen differenzierte Beurteilungen über verschiedene Situationen hinweg mit sich bringen. *Dies scheint jedoch eine Kompetenz zu sein, die nicht durch allgemeine Lerngelegenheiten oder Berufserfahrung allein entsteht. Vielmehr scheinen Lehrkräfte besonders holistisch zu urteilen und insbesondere negative Situationen zu fokussieren, während Studierende der Allgemeinpädagogik und Bachelorstudierende eher gelingende Prozesse wahrnehmen.*

Dass die Beurteilungen nicht situationsunabhängig sind, zeigt sich auch in der Faktorenstruktur, in der die Items einer Vignette jeweils deutliche Testlets bilden. Das bedeutet, dass die Vignetten in sich konsistent, aber voneinander weitgehend unabhängig beurteilt werden. Für

die Beurteilungen scheinen Gruppenmerkmale eine Rolle zu spielen sowie Gesamteindrücke der Situationen, die auch von anderen Merkmalen als den Konstruktionskriterien abhängig sind. *Insgesamt scheint die Arbeit mit Vignetten also von zahlreichen Einflüssen auf der Seite der Urteilenden wie auch auf der Seite der Vignetten und damit verbundenen Items und Instruktionen begleitet zu werden.*

Zusammenfassend kann festgehalten werden, dass anhand der operationalisierten Lernbarrieren und Konstruktionskriterien Vignetten und Items entstanden sind, die dazu geeignet sind, Beurteilungsprozesse zum Gelingen oder Misslingen von Lehr-Lernprozessen auszulösen. Das Beurteilen von Lernbarrieren und das Beurteilen von gelingenden Situationen kann anhand dieser Vignetten mit akzeptabler Reliabilität erfasst werden. Die Vignetten selbst scheinen dabei einen deutlichen Einfluss auf die Beurteilungen zu haben. Zum einen wird der starke Einfluss der beschriebenen Unterrichtssituationen deutlich, so dass sich die Frage stellt, welche Einflüsse die Vignettengestaltung und Verfahrensmerkmale auf die Beurteilungen haben. Zum anderen rückt die Spezifität der Kompetenz *Erkennen von Lernbarrieren* in den Blick sowie die Frage danach, welche anderen Einflüsse auf der Seite der Urteilenden das Erkennen beeinflussen. Dies impliziert weitere mögliche Fragestellungen für die Weiterentwicklung vignettenbasierter Verfahren, die im letzten Kapitel skizziert werden sollen.

13.7 Ausblick

Die Wahrnehmung und Interpretation relevanter Unterrichtsereignisse ist eine wichtige Kompetenz für qualitätvollen Unterricht. In Bezug auf diagnostische Fragestellungen und potenzielle Lernbarrieren zeigen bisherige Ergebnisse, dass eine Fokussierung auf Unterrichtsmerkmale zur Analyse kritischer Lehr-Lernsituationen zu wenig erfolgt. Umso dringlicher scheint es, dass die Fertigkeit, Unterrichtsprozesse zu beobachten und Lernbarrieren zu erkennen, stärker ausgebildet wird. Damit stellt sich gleichzeitig die Frage nach der Erfassung dieser Kompetenz, um Lernstände und Fortschritte evaluieren zu können und besser zu verstehen, wie solche spezifischen Wahrnehmungsprozesse beschaffen sind und wovon sie beeinflusst werden. Im Folgenden werden mögliche weiterführende Forschungsansätze benannt:

- *Variation des Studiendesigns*: Eines der deutlichsten Ergebnisse ist, dass Beurteilungen von Lernbarrieren und gelingenden Situationen nicht differenziert erfolgen, sondern eher holistisch und dass die jeweiligen Vignetten diese Beurteilungen stark beeinflussen. Eine der Annahmen besteht darin, dass die spezifische Fähigkeit zur Beurteilung von Lernbarrieren in Unterrichtsmerkmalen in der Stichprobe nicht ausgebildet war und spezifisch erlernt werden muss (Wenck et al., im Druck). Eine Längsschnitterhebung mit dem entwickelten Verfahren

im Verlauf einer intensiven, spezifischen Lerngelegenheit könnte dazu Ergebnisse liefern. Dies wäre als Lernverlauf oder Kontrollgruppendesign denkbar.

- *Variation der Vignetten*: Eine weitere Annahme ist, dass die Komposition der Vignetten und die darin enthaltenen Kontextinformationen die Beurteilungen beeinflussen. Dies betrifft nicht nur gezielte Konstruktionskriterien wie die Anzahl von Barrieren pro Vignette. Offenbar haben auch andere, unbeabsichtigte Aspekte einen Einfluss auf die Gesamtwahrnehmung der Situationen. Solche Einflüsse auf die Wahrnehmung von Unterrichtssituationen könnten über die gezielte Variation von Merkmalen bei ansonsten gleichen Konstruktionskriterien erprobt werden. Denkbar wären beispielsweise Variationen bzgl. der Unterrichtsmerkmale in Anzahl, Reihenfolge und Ausprägung. Kontextinformationen könnten variiert werden bzgl. dargestellten individuellen Bedingungen von Kindern bzw. Jugendlichen, Namen, Unterrichtsfächern, Arbeitsphasen, Geschlechtern oder Eigenschaften der Lehrkraft wie Ausdrucksweise oder Freundlichkeit.
- *Variation der Instruktion*: Die Formulierung der Instruktion trägt möglicherweise zu einer holistischen Eindrucksbildung bei. Zwar lassen sich für die Variation der Instruktion im vierten Entwicklungsschritt (Markierungsaufgabe) keine relevanten Unterschiede zeigen. Ein Einfluss der Instruktion ist dennoch zu prüfen. Die Instruktion könnte durch variierte Formulierungen unterschiedlich stark differenzierte Beurteilungen von Unterrichtsmerkmalen initiieren.
- *Variation der Items*: Die Formulierungen der Items enthalten jeweils Umschreibungen, die die Unterrichtsmerkmale repräsentieren sollen (z. B. *Allgegenwärtigkeit: Die Lehrkraft hat die ganze Klasse im Blick*). Dabei bleibt unklar, ob Urteile tatsächlich im Hinblick auf Allgegenwärtigkeit gebildet werden oder ob die Formulierungen für Leserinnen und Leser auch andere Merkmale repräsentieren könnten. Wird das Verfahren in einer Stichprobe eingesetzt, in der Unterrichtsmerkmale bekannt sind, könnte auf diese direkt verwiesen werden, z. B. *Beurteilen Sie die Allgegenwärtigkeit der Lehrkraft*. Offen bleibt dabei, welche Informationen zur Beurteilung genutzt werden. Urteile zu den verwendeten Merkmalen würden aber explizit abgefragt.

Des Weiteren könnte die Vermutung überprüft werden, dass es in der Itemformulierung eine Rolle spielt, ob die Lehrkraft als handelnde Person genannt wird oder nicht und ob somit eher eine Person oder eher eine Situation beurteilt wird. Die Annahme besteht dabei darin, dass Personen möglicherweise zurückhaltender und weniger negativ beurteilt werden. Itemformulierungen, die diese Unterschiede messen könnten, sind z. B. *Es sind reibungslose Übergänge zwischen Unterrichtseinheiten erkennbar* vs. *Die Lehrkraft gestaltet Übergänge zwischen Unterrichtseinheiten reibungslos*.

- *Variation der Auswertungskriterien:* Die Auswertungsnorm wurde in der vorliegenden Arbeit von der Autorin bestimmt und durch ein Expertinnen- und Expertengespräch validiert. Es wurde angenommen, dass ein sicheres Erkennen der Situationen mit eindeutigen Urteilen einhergeht, also Wert 1 oder 4 auf der vierstufigen Likert-Skala. Unter der Annahme, dass nicht alle Merkmale gleich stark ausgeprägt sind und sich eine Expertise im Erkennen von Lernbarrieren auch darin zeigt, dass das Ausmaß möglicher hindernder Bedingungen erkannt wird, könnte eine Expertinnen- und Expertennorm weitere Erkenntnisse ermöglichen. Eine Bewertung mit 2 Punkten bei völliger Übereinstimmung mit der Expertinnen- und Expertennorm und mit 1 Punkt bei richtiger Tendenz könnte genauere Ergebnisse ermöglichen.
- *Variation des Ansatzes:* Für die Entwicklung des Verfahrens ErLe wurde ein quantitativer Ansatz gewählt, der das Erkennen von Lernbarrieren und gut umgesetzten Unterrichtsmerkmalen gezielt erfasst und eine ökonomische und objektive Auswertung erlaubt. In den Ergebnissen ergaben sich interessante Erkenntnisse hinsichtlich der Gruppenunterschiede, die weniger durch die eigentlich zu erfassenden Fähigkeitsausprägungen zu erklären sind und eher auf unterschiedliche Fokussierungen hinweisen. Dies wirft die Frage auf, ob und warum z. B. Lehrkräfte mehr die negativen Aspekte (Lernbarrieren) fokussieren oder Studierende der Allgemeinpädagogik eher positive Aspekte (gelingende Situationen). Solche unterschiedlichen Herangehensweisen in der Wahrnehmung von Unterricht können erfasst werden, indem der quantitative Ansatz durch eine qualitative Erhebung ergänzt wird. Möglich wäre z. B. die Auswertung der Frage *Was fällt Ihnen in dieser Situation besonders auf?* oder *Was halten Sie in dieser Situation für besonders relevant?*
- *Erfassen von Merkmalen der Urteilenden:* Als Teilkompetenz für das Erkennen von Lernbarrieren wurden spezifische Dispositionen benannt, aber noch nicht systematisch erfasst. Die Kombination mit Wissens- und Einstellungstests könnte neue Erkenntnisse zur Struktur und Genese der erfassten Kompetenz ermöglichen. Relevant wären beispielsweise Wissen zu KF und LU, Überzeugungen zu Gelingensbedingungen von Unterricht oder Art und Anzahl spezifischer Lerngelegenheiten.
- *Prüfung der prognostischen Validität und Kreuzvalidität:* Vor dem Hintergrund, dass Textvignetten als weniger realitätsnah gelten, ist zu fragen, ob sich Fähigkeitsausprägungen in einem textvignettenbasierten Verfahren auch in videobasierten Verfahren zeigen und ob diese auch prognostisch valide sind. Zu prüfen wäre dies durch die Kombination mit einem validierten videobasierten Verfahren und durch die Beobachtung von realem Unterricht.
- *Weitere Verfahrensentwicklung:* In der Studie wurde explorativ vorgegangen. Für die Items wurden Itemkennwerte bestimmt und es standen bei der weiteren Entwicklung Fragen im Vordergrund, welche Einflüsse Veränderungen an Items und Vignetten zur Folge haben und

welche Schlüsse sich daraus zu Beurteilungsprozessen ziehen lassen. Dementsprechend wurde mit einem kleinen und relativ gleichbleibenden Itempool gearbeitet. Zur Entwicklung eines validen und reliablen Verfahrens wäre es im weiteren Vorgehen nötig, mit einer größeren Menge an Items zu arbeiten, um diejenigen mit schlechten Itemkennwerten ausschließen zu können. Ein Vorgehen nach probabilistischer Testtheorie würde für die Entwicklung eines validen und reliablen Testverfahrens neue Erkenntnisse ermöglichen.

- *Andere Einsatzmöglichkeiten des entwickelten Materials:* Die entwickelten Vignetten wurden validiert, mit Experten und Expertinnen diskutiert und von Studierenden wie Lehrkräften als authentisch und gut vorstellbar beurteilt. Damit sind Unterrichtsbeispiele entstanden, die sich für die Auseinandersetzung mit realitätsnahen Problemsituationen eignen und Fallbeispiele für Lernbarrieren in Unterrichtsmerkmalen enthalten. Eine weitere Einsatzmöglichkeit des entstandenen Materials besteht somit in der Lehre, insbesondere zur Übung des systematischen Beobachtens in entsprechenden Lehrformaten oder Fortbildungen. Es kann den diagnostischen Blick erweitern und zur Sensibilisierung für Barrieren in Unterrichtsbedingungen beitragen.

Diesem Ausblick ließen sich sicherlich noch weitere Punkte hinzufügen. Die Entwicklung von vignettenbasierten Verfahren zur Erfassung der PW zeigt ebenso viel Potenzial wie Entwicklungsbedarf.

Es gibt kein größeres Hindernis des Fortgangs in den Wissenschaften als das Verlangen, den Erfolg davon zu früh verspüren zu wollen.

Georg Christoph Lichtenberg

Literaturverzeichnis

- Adamina, M., Möller, K., Steffensky, M., Sunder, C., & Wyssen, H.-P. (2017). *Maßnahmen der Lernunterstützung im naturwissenschaftlichen Sachunterricht – Kognitiv anregen und inhaltlich strukturieren*. Institut für Psychologie in Bildung und Erziehung der Westfälischen Wilhelms-Universität Münster. https://www.uni-muenster.de/imperia/md/content/ko-viu/szenen_lu/massnahmen-der-lernunterstuetzung-uebersicht_-viu.pdf.
- Agarwala, A. (2020). *Das Integrationsexperiment: Flüchtlinge an der Schule – eine Bilanz nach fünf Jahren*. Duden.
- von Aufschnaiter, C., Selter, C., & Michaelis, J. (2017). Nutzung von Vignetten zur Entwicklung von Diagnose- und Förderkompetenzen. Konzeptionelle Überlegungen und Beispiele aus der MINT-Lehrerbildung. In C. Selter, S. Hußmann, C. Hößle, C. Knipping, & K. Lengnink (Hrsg.), *Diagnose und Förderung heterogener Lerngruppen. – Theorien, Konzepte und Beispiele aus der MINT-Lehrerbildung* (S. 85-105). Waxmann.
- Auspurg, K., Hinz, T., & Liebig, S. (2009). Komplexität von Vignetten, Lerneffekte und Plausibilität im faktoriellen Survey. *Methoden, Daten, Analysen (mda)*, 3(1), 59–96. https://kops.uni-konstanz.de/bitstream/123456789/11746/1/04_Auspurg.pdf
- Avramidis, E., & Norwich, B. (2002). Teachers' Attitudes Towards Integration/Inclusion: A Review of the literature. *European Journal of Special Needs Education*, 17(2), 129–147. <https://doi.org/10.1080/08856250210129056>
- Baer, M. & Buholzer, A. (2005). Beiträge über Forschung und Evaluation in der Ausbildung von Lehrkräften: Analyse der Wirksamkeit der berufsfeldorientierten Ausbildung für den Erwerb von Unterrichts- und Diagnosekompetenz. *Beiträge zur Lehrerbildung*, 23 (2), 243-248.
- Baer, M., Dörr, G., Fraefel, U., Kocher, M., Küster, O., Larcher, S., Müller, P., Sempert, W., & Wyss, C. (2007). Werden angehende Lehrpersonen durch das Studium kompetenter? Kompetenzaufbau und Standarderreicherung in der berufswissenschaftlichen Ausbildung an drei pädagogischen Hochschulen in der Schweiz und in Deutschland. *Unterrichtswissenschaft*, 35(1), 15–47. <https://doi.org/10.25656/01:5485>
- Barrierefrei leben e.V. (o. D.). *Barrierefrei leben*. Abgerufen am 27.12.2023, von <https://www.barrierefrei-leben.de>.
- Barth, V. L. (2017). Professionelle Wahrnehmung von Störungen im Unterricht. In *Springer e-Books*. <https://doi.org/10.1007/978-3-658-16371-6>

- Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft*, 9 (4), 469-520. <https://doi.org/10.1007/s11618-006-0165-2>
- Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29-53). Waxmann. <https://doi.org/10.31244/9783830974338>
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M., & Tsai, Y. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180. <https://doi.org/10.3102/0002831209345157>
- Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderung (o. D.a). *Barrierefreiheit – von Beginn an*. Abgerufen am 27.12.2023, von <https://www.behindertenbeauftragter.de/DE/AS/schwerpunkte/barrierefreiheit/barrierefreiheit-node.html>.
- Beauftragter der Bundesregierung für die Belange von Menschen mit Behinderung (o. D.b). *Was ist Barrierefreiheit?* Abgerufen am 03.04.2023, von https://www.behindertenbeauftragter.de/DE/Themen/Barrierefreiheit/WasistBarrierefreiheit/WasistBarrierefreiheit_node.html.
- Behindertengleichstellungsgesetz [BGG] (2002). Gesetze im Internet. Abgerufen am 12.12.20023, von https://www.gesetze-im-internet.de/bgg/_4.html.
- Berliner, D. C. (1987). Der Experte im Lehrerberuf: Forschungsstrategien und Ergebnisse. *Unterrichtswissenschaft*, 15(3), 295-305. <https://doi.org/10.25656/01:5455>
- Berliner, D. C. (1994). The wonder of exemplary performances. In J. N. Mangieri, & C. C. Block (Hrsg.), *Creating powerful thinking in teachers and students. Diverse perspectives* (S. 1–37). Holt, Rinehart & Winston.
- Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, 56(3), 205–213. <https://doi.org/10.1177/0022487105275904>
- Beukelman, D. R., & Mirenda, P. (1998). *Augmentative & Alternative Communication: Supporting children & adults with complex communication needs*. Paul. H. Brookes.
- Beywl, W., & Schmid, P. (2014). Lehren als Profession. Kriterien guten Unterrichts - eine Synopse. *Weiterbildung*, 6, 10–14. <https://irf.fhnw.ch/handle/11654/30222?show=full>

Bilz, L., Sudeck, G., Bucksch, J., Klocke, A., Kolip, Melzer, W., Ravens-Sieberer, U., & Richter, M. (2016). *Schule und Gesundheit. Ergebnisse des WHO-Jugendgesundheitssurveys „Health Behaviour in School-aged Children“*. Beltz Juventa.

Blömeke, S., Felbrich, A., & Müller, C. (2008). Messung des erziehungswissenschaftlichen Wissens angehender Lehrkräfte. In S. Blömeke, G. Kaiser, & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung* (S. 171-193). Waxmann.

Blömeke, S., Gustafsson, J.E. & Shavelson, R. (2015a). Beyond Dichotomies: Competence Viewed as a Continuum. *Zeitschrift für Psychologie*, 223(1), 3-13.
<https://doi.org/10.1027/2151-2604/a000194>

Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., & König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learning and Instruction*, 79, 101600. <https://doi.org/10.1016/j.learninstruc.2022.101600>

Blömeke, S., König, J., Suhl, U., Hoth, J. & Döhrmann, M. (2015b). Wie situationsbezogen ist die Kompetenz von Lehrkräften? Zur Generalisierbarkeit von videobasierten Performanztests. *Zeitschrift für Pädagogik*, 61, 310-327.

Blömeke, S., & Suhl, U. (2010). Modellierung von Lehrerkompetenzen. *Zeitschrift für Erziehungswissenschaft*, 13 (3), 473-505. <https://doi.org/10.1007/s11618-010-0130-y>

Bloom, B. S. (1976). *Human characteristics and school learning*. McGraw-Hill.

Bönte, J., Lenske, G., Dicke, T. & Leutner, D. (2019). Inszenierte Unterrichtsvideovignetten zur Förderung des Wissens um Klassenführung von (angehenden) Lehrkräften. In H. Angenent, B. Heidkamp, & D. Kergel (Hrsg.), *Digital Diversity* (S. 241-257). Springer Fachmedien. https://doi.org/10.1007/978-3-658-26753-7_15

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, 33 (8), 3-15. <https://doi.org/10.3102/0013189X033008003>

Borko, H., & Putnam, R. T. (1996). Learning to teach. In D. C. Berliner, & R. C. Calfee (Hrsg.), *Handbook of educational psychology* (S. 673–708). Macmillan.

Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Springer.

- Bransford, J., Derry, S., Berliner, D., Hammerness, K., & Beckett, K. L. (2005). Theories of learning and their roles in teaching. In L. Darling-Hammond & J. Bransford (Hrsg.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (S. 40-87). Jossey-Bass.
- Bromme, R. (1992). *Der Lehrer als Experte: Zur Psychologie des professionellen Wissens*. Huber.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers. In F. E. Weinert (Hrsg.), *Psychologie des Unterrichts und der Schule* (S. 177–212). Hogrefe.
- Bromme, R. (2008). Lehrerexpertise. In W. Schneider, & M. Hasselhorn (Hrsg.), *Handbuch der Pädagogischen Psychologie*. (S. 159-167). Hogrefe.
- Brophy, J. E. (2006). History of research on classroom management. In C. M. Evertson, & C. S. Weinstein (Hrsg.), *Handbook of classroom management. Research, practice and contemporary issues* (S. 17–43). Lawrence Erlbaum Associates.
- Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Hrsg.), *Handbook of research on teaching. A project of the American Educational Research Association* (S. 328–375). MacMillan.
- Brovelli, D., Bölsterli, K., Rehm, M., & Wilhelm, M. (2013). Erfassen professioneller Kompetenzen für den naturwissenschaftlichen Unterricht: Ein Vignettentest mit authentisch komplexen Unterrichtssituationen und offenem Antwortformat. *Unterrichtswissenschaft*, 41 (4), 306-329.
- Brovelli, D., Bölsterli, K., Rehm, M., & Wilhelm, M. (2014). Using Vignette Testing to Measure Student Science Teachers' Professional Competencies. *American Journal of Educational Research*, 2(7), 555-558. <https://doi.org/10.12691/education-2-7-20>
- Brunner, E. (2018). Qualität von Mathematikunterricht: Eine Frage der Perspektive. *Journal für Mathematik Didaktik*, 39(2), 257-284. <https://doi.org/10.1007/s13138-017-0122-z>
- Bühner, M. (2006). Einführung in die Test- und Fragebogenkonstruktion. Pearson.
- Carroll, J. B. (1973). Ein Modell schulischen Lernens. Deutsch von G. und H. Tütgen, In W. Edelstein, & D. Hopf (Hrsg.): *Bedingungen des Bildungsprozesses*. Klett.
- Cipani, E. (2008). *Classroom management for all teachers: Plans for evidence-based practice* (3. Aufl.). Pearson Merrill/Prentice Hall.
- Clausen, M. (2002). *Unterrichtsqualität: Eine Frage der Perspektive?* Waxmann.

- Clausen, M., Reusser, K., & Klieme, E. (2003). Unterrichtsqualität auf der Basis hochinferenter Unterrichtsbeurteilungen. *Unterrichtswissenschaft*, 31(2), 122–141.
- Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta-analysis. *Journal of research in reading*, 42(2), 288-325. <https://doi.org/10.1111/1467-9817.12269>
- Cohen, J. (1992). A power primer. *Quant. Meth. Psychol.* 112, 155–159. <https://doi.org/10.1037//0033-2909.112.1.155>
- Colestock, A., & Sherin, M.G. (2009). Teachers' sense-making strategies while watching video of mathematics instruction. *Journal of Technology and Teacher Education*, 17 (1), 7-29.
- Decristan, J., Hess, M., Holzberger, D., & Praetorius, A.-K. (2020). Oberflächen- und Tiefenmerkmale: eine Reflexion zweier prominenter Begriffe der Unterrichtsforschung. *Zeitschrift für Pädagogik*, 66(1), 102-116.
- Denn, A. K. (2021). Interaktionen von Lehrpersonen mit Mädchen und Jungen im Mathematikunterricht der Grundschule. Springer. <https://doi.org/10.1007/978-3-658-32372-1>
- Denn, A., Lotz, M., Theurer, C., & Lipowsky, F. (2015). „Prima, Lisa. Richtig“ und „Psst, Max. Hör auf zu stören!“ Eine quantitative Studie zu Unterschieden im Feedbackverhalten von Lehrkräften gegenüber Mädchen und Jungen im Mathematikunterricht des zweiten Schuljahres. *Gender*, 7(1), 29–47. <https://doi.org/10.3224/gender.v7i1.18155>
- von Dewitz, N., Massumi, M., & Grießbach, J. (2016). *Neu zugewanderte Kinder, Jugendliche und junge Erwachsene: Entwicklungen im Jahr 2015*. Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache.
- Dicke, T., Elling, J., Schmeck, A., & Leutner, D. (2015). Reducing reality shock: The effects of classroom management skills training on beginning teachers. *Teaching and Teacher Education*, 48, 1-12. <https://doi.org/10.1016/j.tate.2015.01.013>
- Ditton, H. (2006). Unterrichtsqualität. In K.-H. Arnold, U. Sandfuchs, & J. Wiechmann (Hrsg.), *Handbuch Unterricht* (S. 235-243). Klinkhardt.
- Doyle, W. (1986). Classroom organization and management. In M. Wittrock (Hrsg.), *Handbook of research on teaching* (3. Aufl.) (S. 392-431). Macmillan.
- Doyle, W. (2006). Ecological approaches to classroom management. In C. M. Evertson, & C. S. Weinstein (Hrsg.), *Handbook of classroom management. Research, practice and contemporary issues* (S. 97–125). Lawrence Erlbaum Associates.

- Dreher, A., & Kuntze, S. (2015). Teachers' professional knowledge and noticing: The case of multiple representations in the mathematics classroom. *Educational Studies in Mathematics*, 88 (1), 89-114. <https://doi.org/10.1007/s10649-014-9577-8>
- Drollinger-Vetter, B., & Lipowsky, F. (2006). Fachdidaktische Qualität der Theoriephasen. In I. Hugener, C. Pauli, & K. Reusser (Hrsg.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“*. *Videoanalysen* (Materialien zur Bildungsforschung, Bd. 15, S. 189–205). GPPF.
- Dudenredaktion (o. D.a). *Duden online. Barriere*. Abgerufen am 3. April 2023, von <https://www.duden.de/rechtschreibung/Barriere>.
- Dudenredaktion (o. D.b). *Duden online. erkennen*. Abgerufen am 3. April 2023, von <https://www.duden.de/rechtschreibung/erkennen>.
- Eckes, T. (2015). Lokale Abhängigkeit von Items im TestDAF Leseverstehen. *Diagnostica*, 61(2), 93–106. <https://doi.org/10.1026/0012-1924/a000118>
- Einsiedler, W. (2001). *Ergebnisse und Probleme der Unterrichtsforschung im Primarbereich*. Institut für Grundschulforschung der Universität Erlangen-Nürnberg.
- Einsiedler, W. (2002). Das Konzept „Unterrichtsqualität“. *Unterrichtswissenschaft*, 30 (3), 194-196.
- Emmer, E. T., & Evertson, C. M. (2013). *Classroom management for middle and high school teachers* (9. Aufl.). Pearson.
- van Es, E. A., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and teacher education*, 24(2), 244-276. <https://doi.org/10.1016/j.tate.2006.11.005>
- van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education*, 13 (2), 155-176. <https://doi.org/10.1007/s10857-009-9130-3>
- Evolution of Reading in the Age of Digitisation [E-READ] (2019). *Stavanger Declaration*. <https://ereadcost.eu/wp-content/uploads/2019/01/StavangerDeclaration.pdf>.
- Ewerhardy, A., Kleickmann, T., & Möller, K. (2012). Fördert ein konstruktivistisch orientierter naturwissenschaftlicher Sachunterricht mit strukturierenden Anteilen das konzeptuelle Verständnis bei den Lernenden? *Zeitschrift für Grundschulforschung. Bildung im Elementar- und Primarbereich*, 5 (1), 76-88.

- Fauth, B., Decristan, J., Rieser, S., Klieme, E., & Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction, 29*, 1-9. <https://doi.org/10.1016/j.learninstruc.2013.07.001>
- Fauth, B., Göllner, R., Lenke, G., Praetorius, A.-K., & Wagner, W. (2020). Who sees what? Conceptual considerations on the measurement of teaching quality from different perspectives. *Zeitschrift Für Pädagogik, 66*, 138–155.
- Fischer, C., Kopmann, H., Rott, D., Veber, M., & Zeinz, H. (2014). Adaptive Lehrkompetenz und pädagogische Haltung. Lehrerbildung für eine inklusive Schule. In K. Zierer (Hrsg.), *Jahrbuch für Allgemeine Didaktik*. Vol. 4: Didaktik der inklusiven Schule (S. 1-19). Universität Tübingen.
- Frey, A. (2008). *Kompetenzstrukturen von Studierenden in der ersten und zweiten Phase der Lehrerbildung: eine nationale und internationale Standortbestimmung*. Empirische Pädagogik.
- Friesen, M., & Kuntze, S. (2016). Assessing pre-service teachers' competence of analysing learning support situations through a multi-format test instrument comprising of video, comic, and text vignettes. *Proceedings of ETC, 3*, 36-45.
- Friesen, M., Kuntze, S., & Vogel, M. (2018). Videos, Texte oder Comics? Die Rolle des Vignettenformats bei der Erhebung fachdidaktischer Analysekompetenz zum Umgang mit Darstellungen im Mathematikunterricht. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß, & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung - Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 153-177). Springer. https://doi.org/10.1007/978-3-658-20121-0_8
- Gabriel, K. (2014). *Videobasierte Erfassung von Unterrichtsqualität im Anfangsunterricht der Grundschule. Klassenführung und Unterrichtsklima in Deutsch und Mathematik*. Kassel University Press.
- Gebert, D., & Boerner, S. (1997). Mentale Lernbarrieren in Organisationen und Ansätze zu ihrer Überwindung. In Wieselhuber, N. (1997). *Handbuch lernende Organisation: Unternehmens- und Mitarbeiterpotentiale erfolgreich erschliessen* (S. 237-248). Gabler.
- Gippert, C., Gold, B., Seeger, D., Junker, R. & Holodynski, M. (2019). *Manual zur theoriegeleiteten Interpretation klassenführungsrelevanter Unterrichtereignisse*. Institut für Psychologie in Bildung und Erziehung der Westfälischen Wilhelms-Universität Münster. https://www.uni-muenster.de/imperia/md/content/koviu/viu_manual_analyse_kf.pdf.

- Gold, B., Förster, S., & Holodynski, M. (2013). Evaluation eines videobasierten Trainingsseminars zur Förderung der professionellen Wahrnehmung von Klassenführung im Grundschulunterricht. *Zeitschrift für Pädagogische Psychologie*, 27(3), 141–155.
<https://doi.org/10.1024/1010-0652/a000100>
- Gold, B., Hellermann, C., Burgula, K., & Holodynski, M. (2016a). Fallbasierte Unterrichtsanalyse. Effekte von video- und textbasierter Fallanalyse auf kognitive Belastung, aufgabenspezifisches Interesse und die professionelle Unterrichtswahrnehmung von Grundschullehrerstudierenden. *Unterrichtswissenschaft*, 44(4), 322–338.
- Gold, B., Hellermann, C. & Holodynski, M. (2016b). Professionelle Wahrnehmung von Klassenführung - Vergleich von zwei videobasierten Erfassungsmethoden. In K. Schwippert, & D. Prinz (Hrsg.), *Der Forschung - Der Lehre - Der Bildung: Aktuelle Entwicklungen der Empirischen Bildungsforschung* (S.103-118). Waxmann.
- Gold, B., & Holodynski, M. (2011). Klassenführung. In E. Kiel, & K. Zierer (Hrsg.), *Basiswissen Unterrichtsgestaltung: Unterrichtsgestaltung als Gegenstand der Praxis* (S. 133– 151). Schneider.
- Gold, B., & Holodynski, M. (2017). Using digital video to measure the professional vision of elementary classroom management: Test validation and methodological challenges. *Computers & Education*, 107, 13–30. <https://doi.org/10.1016/j.compedu.2016.12.012>
- Gold, B., Meschede, N., Fiebranz, A., Steffensky, M., Holodynski, M., & Möller, K. (2017). Professionelles Wissen über und Wahrnehmung von Klassenführung und Lernunterstützung im naturwissenschaftlichen Grundschulunterricht – Eine Zusammenhangsstudie aus generischer und naturwissenschaftsdidaktischer Perspektive. In H. Fischer, & E. Sumfleth (Hrsg.), *Professionelle Kompetenz von Lehrkräften der Chemie und Physik* (Studien zum Physik- und Chemielernen; Band 200). (S. 203-219). Logos.
- Goodwin, C. (1994). Professional Vision. *American Anthropologist, New Series*, 96 (3), 606–633.
<https://doi.org/10.1525/aa.1994.96.3.02a00100>
- Gruehn, S. (1995). Vereinbarkeit kognitiver und nicht-kognitiver Ziele im Unterricht. *Zeitschrift für Pädagogik*, 41, 531 – 553.
- Gruehn, S. (2000). *Unterricht und schulisches Lernen: Schüler als Quellen der Unterrichtsbeschreibung*. Waxmann.
- Gruschka, A. (2007). „Was ist guter Unterricht?“. Über neue Allgemein-Modellierungen aus dem Geiste der empirischen Unterrichtsforschung. *Pädagogische Korrespondenz* 36, 10–43.

- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component structure. *Psychological Bulletin*, 103, 265-275. <https://doi.org/10.1037/0033-2909.103.2.265>
- Haag, L. (2018). *Kernkompetenz Klassenführung*. Klinkhardt. <https://doi.org/10.36198/9783838549347>
- Haag, L., & Streber, D. (2012). *Klassenführung: Erfolgreich unterrichten mit Classroom Management*. Beltz.
- Hanisch, C., & Hennemann, T. (2020). Die Rolle der Schule in der kindlichen und jugendlichen Entwicklung. *Springer Reference Medizin*, 1–10. https://doi.org/10.1007/978-3-662-49289-5_44-1
- Hartig, J., & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 127-143). Springer. https://doi.org/10.1007/3-540-33020-8_9
- Hattie, J. (2008). *Visible learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. Routledge.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77 (1), 81-112. <https://doi.org/10.3102/003465430298487>
- Heinitz, B., & Nehring, A. (2020). Kriterien naturwissenschaftsdidaktischer Unterrichtsqualität – ein systematisches Review videobasierter Unterrichtsforschung. *Unterrichtswissenschaft* 48, 319–360. <https://doi.org/10.1007/s42010-020-00074-8>
- Hellermann, C., Gold, B., & Holodynski, M. (2015). Förderung von Klassenführungsfähigkeiten im Lehramtsstudium. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 47(2), 97–109. <https://doi.org/10.1026/0049-8637/a000129>
- Helmke, A. (2000). TIMSS und die Folgen: Der weite Weg von der externen Leistungsevaluation zur Verbesserung des Lehrens und Lernens. In U. P. Trier (Hrsg.), *Bildungswirksamkeit zwischen Forschung und Politik* (S. 135–164). Rüegger.
- Helmke, A. (2004). *Unterrichtsqualität: Erfassen, Bewerten, Verbessern* (3. Aufl.). Kallmeyersche Verlagsbuchhandlung.
- Helmke, A. (2006). Was wissen wir über guten Unterricht? *Pädagogik*, 2, 42-45.
- Helmke, A. (2007). *Unterrichtsqualität und Unterrichtsentwicklung: Wissenschaftliche Erkenntnisse zur Unterrichtsforschung und Konsequenzen für die Unterrichtsentwicklung*. Bertelsmann.

- Helmke, A. (2009). *Unterrichtsqualität und Lehrerprofessionalität. Diagnose, Evaluation und Verbesserung des Unterrichts*. Klett-Kallmeyer.
- Helmke, A. (2010). Unterrichtsqualität. In D. H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie* (4. überarb. u. erw. Aufl., S. 886-895). Beltz.
- Helmke, A. (2011). Forschung zur Lernwirksamkeit des Lehrerhandelns. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (630-643). Waxmann.
- Helmke, A., & Helmke, T. (2014). Wie wirksam ist gute Klassenführung? *Lernende Schule*, 65, 9-12.
- Helmke, A., Helmke, T., Heyne, N., Hosenfeld, A., Schrader, F.-W., & Wagner, W. (2010). Effiziente Klassenführung als Schlüsselmerkmal der Unterrichtsqualität – ein Untersuchungsbeispiel aus der Grundschule. In C. Spiel, R. Reimann, B. Schober, & P. Wagner (Hrsg.), *Bildungspsychologie* (S. 101-105). Hogrefe.
- Helmke, A., & Schrader, F.-W. (2008). Merkmale der Unterrichtsqualität: Potenzial, Reichweite und Grenzen. *SEMINAR – Lehrerbildung und Schule*, 3, 17-47.
- Heran-Dörr, E., Wiesner, H., & Kahlert, J. (2007). Schülerorientierung oder Orientierung an Schülervorstellungen? Wie Lehrkräfte vor und nach einer internetgestützten Fortbildungsmaßnahme über physikbezogenen Sachunterricht denken. *Zeitschrift für Didaktik der Naturwissenschaften*, 13, 161-180.
- Herbst, P., Aaron, W., & Erickson, A. (2013). How preservice teachers respond to representations of practice: A comparison of animations and video [Paper presented at the 2013 Annual Meeting of the AERA, San Francisco].
- Herbst, P., & Kosko, K. W. (2013). Using representations of practice to elicit mathematics teachers' tacit knowledge of practice: A comparison of responses to animations and videos. *Journal of Mathematics Teacher Education*, 17 (6), 515-537.
<https://doi.org/10.1007/s10857-013-9267-y>
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K., Hollingsworth, H., Jacobs, J., Chui, A. M., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E. A., Etterbeek, W., Manaster, C., & Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 Video Study*. National Center for Education Statistics.
<https://doi.org/10.1037/e610352011-003>

- Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F.K. Lester (Hrsg.), *Second handbook of research on mathematics teaching and learning. A project of the National Council of Teachers of Mathematics* (S. 371–404). Information Age.
- Hitchcock, A. (o. D.). „Um einen großartigen Film zu machen, brauchst du drei Dinge – das Drehbuch, das Drehbuch und das Drehbuch.“ Socrates. <https://www.socreate.it/de/blog-beiträge/drehbuchs schreiben/20-inspirierende-zitate-uber-das-drehbuchs schreiben>.
- Hogan, K., & Pressley, M. (1997). Scaffolding scientific competencies within classroom communities of inquiry. In K. Hogan, & M. Pressley (Hrsg.), *Scaffolding student learning: Instructional approaches and issues* (S. 74-107). Brookline.
- Hollenweger, J. (2021). ICF als gemeinsame konzeptuelle Grundlage. In A. Kunz, R. Luder & C. M. Bösch (Hrsg.), *Inklusive Pädagogik und Didaktik* (S. 33–56). Hep Verlag.
- Holodynski, M., Steffensky, M., Gold, B., Hellermann, C., Sunder, C., Fiebranz, A., Meschede, N., Glaser, O., Rauterberg, T., Todorova, M., Wolters, M., & Möller, K. (2017). Lernrelevante Situationen im Unterricht beschreiben und interpretieren: Videobasierte Erfassung professioneller Wahrnehmung von Klassenführung und Lernunterstützung im naturwissenschaftlichen Grundschulunterricht. In C. Gräsel & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals: Interdisziplinäre Betrachtungen; Befunde und Perspektiven* (S. 283-302). Springer. https://doi.org/10.1007/978-3-658-07274-2_15
- Hoppe, T., Renkl, A., & Rieß, W. (2020). Förderung von unterrichtsbegleitendem Diagnostizieren von Schülervorstellungen durch Video- und Textvignetten. *Unterrichtswissenschaft*, 48(4), 573–597. <https://doi.org/10.1007/s42010-020-00075-7>
- Ingenkamp, K.-H., & Lissmann, U. (2008). *Lehrbuch der pädagogischen Diagnostik*. Beltz.
- Jacobs, V.R. (2017). Complexities in measuring teacher noticing: Commentary. In E.O. Schack, M.H. Fisher, & J.A. Wilhelm (Hrsg.), *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks* (S. 273-279). Routledge. https://doi.org/10.1007/978-3-319-46753-5_16
- Jacobs, V.R., Lamb, L. L. C., & Philipp, R.A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 41 (2), 169- 202. <https://doi.org/10.5951/jresmetheduc.41.2.0169>
- Jahn, G., Prenzel, M., Stürmer, K., & Seidel, T. (2011). Varianten einer computergestützten Erhebung von Lehrerkompetenzen. Untersuchungen zu Anwendungen des Tools Observer. *Unterrichtswissenschaft*, 39 (2), 136-153.

- Janeczko, J., Junker, R., & Holodynski, M. (2022). *Digitales Lernmodul zur Förderung der professionellen Wahrnehmung von Klassenführung*. Institut für Psychologie in Bildung und Erziehung der Westfälische Wilhelms-Universität Münster.
- Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments. *Sociological Methods and Research* 34, 334-423. <https://doi.org/10.1177/0049124105283121>
- Jentsch, A., Schlesinger, L., Heinrichs, H., Kaiser, G., König, J. & Blömeke, S. (2020). Erfassung der fachspezifischen Qualität von Mathematikunterricht: Faktorenstruktur und Zusammenhänge zur professionellen Kompetenz von Mathematiklehrpersonen. *Journal für Mathematik-Didaktik*, 42(1), 97–121. <https://doi.org/10.1007/s13138-020-00168-x>
- Kaiser, G., Busse, A., Hoth, J., König, J., & Blömeke, S. (2015). About the complexities of video-based assessments: Theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *International Journal of Science and Mathematics Education*, 13, 369-387. <https://doi.org/10.1007/s10763-015-9616-7>
- Kaman, A., Ottová-Jordan, V., Bilz, L., Sudeck, G., Moor, I., & Ravens-Sieberer, U. (2020). Subjektive Gesundheit und Wohlbefinden von Kindern und Jugendlichen in Deutschland—Aktuelle Ergebnisse der HBSC-Studie 2017/18. *Journal of health monitoring* 5(3), 7-21.
- Karst, K. (2017). Akkurate Urteile – die Ansätze von Schrader (1989) und McElvany et al. (2009). In A. Südkamp, & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz von Lehrkräften* (S. 21–24). Waxmann.
- Katelhoen, P., & Moroni, M. C. (2018). Inszenierungen direkter Rede in mündlichen Interaktionen. *Studi Germanici*, 1(1), 179–208. <https://air.unimi.it/handle/2434/574532>
- Keller, E., Schmitt, M., & Glaser, C. (2018). Lehrerprofessionalisierung im Bereich der Schreibdidaktik – Konzeption eines vignettenbasierten Testverfahrens. In *Springer eBooks* (S. 93–114). https://doi.org/10.1007/978-3-658-20121-0_5
- Keppens, K., Consuegra, E., Goossens, M., Maeyer, S. de, & Vanderlinde, R. (2019). Measuring preservice teachers' professional vision of inclusive classrooms: A video-based comparative judgement instrument. *Teaching and Teacher Education*, 78, 1–14. <https://doi.org/10.1016/j.tate.2018.10.007>
- Kersting, N. (2008). Using video clips of mathematics classroom instruction as item prompts to measure teachers' knowledge of teaching mathematics. *Educational and Psychological Measurement*, 68(5), 845–861. <https://doi.org/10.1177/0013164407313369>

- Kersting, N.B., Givvin, K.B., Thompson, B., J., Santagata, R., & Stigler, J.W. (2012). Measuring usable knowledge: Teachers' analyses of mathematics classroom videos predict teaching quality and student learning. *American Educational Research Journal*, 49 (3), 568-589. <https://doi.org/10.3102/0002831212437853>
- Kim, L. E., & Klassen, R. M. (2018). Teachers' cognitive processing of complex school-based scenarios: Differences across experience levels. *Teaching and Teacher Education*, 73, 215-226. <https://doi.org/10.1016/j.tate.2018.04.006>
- Kleickmann, T., Großschedl, J., Harms, Z., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V., & Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer – Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft*, 42(3), 280-288.
- Kleickmann, T., Praetorius, A. K., & Steffensky, M. (2018). Qualität naturwissenschaftlichen Unterrichts in der Grundschule: Mehr als drei Basisdimensionen. *Vortrag bei der Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF)*. Basel.
- Kleickmann, T., Steffensky, M. & Praetorius, A. (2020). Quality of teaching in science education: more than three basic dimensions? *Zeitschrift für Pädagogik*, 66(1), 37–53. <https://doi.org/10.3262/zpb2001037>
- Klieme, E. (2013). Qualitätsbeurteilung von Schule und Unterricht: Möglichkeiten und Grenzen einer begriffsanalytischen Reflexion – ein Kommentar zu Helmut Heid. *Zeitschrift für Erziehungswissenschaften*, 16, 433–441. <https://doi.org/10.1007/s11618-013-0356-6>
- Klieme E., & Hartig J. (2007). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin, & H. Krüger (Hrsg.), *Kompetenzdiagnostik*. ZfE, Sonderheft 8, 11-33. https://doi.org/10.1007/978-3-531-90865-6_2
- Klieme, E., Lipowsky, F., Rakoczy, K., & Ratzka, N. (2006). Qualitätsdimensionen und Wirksamkeit von Mathematikunterricht. Theoretische Grundlagen und ausgewählte Ergebnisse des Projekts „Pythagoras“. In M. Prenzel, & L. Allolio-Näcke (Hrsg.), *Untersuchungen zur Bildungsqualität von Schule. Abschlussbericht des DFG-Schwerpunktprogramms* (S. 127–146). Waxmann.
- Klieme, E., & Rakoczy, K. (2008). Empirische Unterrichtsforschung und Fachdidaktik. Outcomeorientierte Messung und Prozessqualität des Unterrichts. *Zeitschrift für Pädagogik*, 54, 222–237.

- Kniewel, I., Lindmeier, A. M., & Heinze, A. (2015). Beyond knowledge: Measuring primary teachers' subject-specific competences in and for teaching mathematics with items based on video vignettes. *International Journal of Science and Mathematics Education*, 13 (2), 309-329. <https://doi.org/10.1007/s10763-014-9608-z>
- Kompetenzzentrum für ein barrierefreies Hamburg (o. D.). *Kompetent barrierefrei*. Abgerufen am 27.12.2023, von <https://kompetent-barrierefrei.de>.
- König, J. (2015). Measuring Classroom Management Expertise (CME) of Teachers: A Video-Based Assessment Approach and Statistical Results. *Cogent Education*, 2(1), 1-15. <https://doi.org/10.1080/2331186X.2014.991178>
- König, J., & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen (PUW). Dokumentation der Kurzfassung des TEDS-M-Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerbildung*. Humboldt-Universität.
- König, J., Blömeke, S. & Kaiser, G. (2015). Early career mathematics teachers' general pedagogical knowledge and skills: Do teacher education, teaching experience, and working conditions make a difference? *International Journal of Science and Mathematics Education*, 13, 331-350. <https://doi.org/10.1007/s10763-015-9618-5>
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A., & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations? A video-based assessment approach. *Teaching and Teacher Education*, 38, 76-88. <https://doi.org/10.1016/j.tate.2013.11.004>
- König, J., & Kramer, C. (2016). Teacher professional knowledge and classroom management: on the relation of general pedagogical knowledge (GPK) and classroom management expertise (CME). *ZDM – The International Journal on Mathematics Education*, 48(1), 139-151. <https://doi.org/10.1007/s11858-015-0705-4>
- König, J., & Lebens, M. (2012). Classroom Management Expertise (CME) von Lehrkräften messen: Überlegungen zur Testung mithilfe von Videovignetten und erste empirische Befunde. *Lehrerbildung auf dem Prüfstand*, 5(1), 3-28.
- König, J., & Seifert, A. (Hrsg.) (2012). *Lehramtsstudierende erwerben pädagogisches Professionswissen: Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrer ausbildung*. Waxmann.
- Kopmann, H. (2022). Zwischen Anspruch und Alltäglichkeit: Konzeptuelle und unterrichtspraktische Perspektiven auf eine inklusive Diagnostik. *Qfi – Qualifizierung für Inklusion*, 3. <https://doi.org/10.21248/qfi.75>

- Kounin, J. S. (2006). *Techniken der Klassenführung*. Waxmann.
- Kramer, C., König, J., Kaiser, G., Ligtoet, R., & Blömeke, S. (2017). Der Einsatz von Unterrichtsvideos in der universitären Ausbildung: Zur Wirksamkeit video- und transkriptgestützter Seminare zur Klassenführung auf pädagogisches Wissen und situationsspezifische Fähigkeiten angehender Lehrkräfte. *Zeitschrift für Erziehungswissenschaft*, 20(1), 137-164.
<https://doi.org/10.1007/s11618-017-0732-8>
- Krammer, K. (2014). Fallbasiertes Lernen mit Unterrichtsvideos in der Lehrerinnen- und Lehrerbildung. *Beiträge zur Lehrerinnen- und Lehrerbildung*, 32(2), 164–175.
<https://doi.org/10.36950/bzl.32.2014.9610>
- Krammer, K., Lipowsky, F., Pauli, C., Schnetzler, C. L., & Reusser, K. (2012). Unterrichtsvideos als Medium zur Professionalisierung und als Instrument der Kompetenzerfassung von Lehrpersonen. In M. Kobarg, C. Fischer, I. M. Dalehefte, F. Trepke, & M. Menk (Hrsg.), *Lehrerprofessionalisierung wissenschaftlich begleiten – Strategien und Methoden* (S. 69–86). Waxmann.
- Krosanke, N. (2021). *Entwicklung der professionellen Kompetenz von Mathematiklehramtsstudierenden zur Bedeutung von Sprache: Eine qualitative Studie zur professionellen Unterrichtswahrnehmung und der Kompetenz zur Analyse von Textaufgaben*. Springer.
<https://doi.org/10.1007/978-3-658-33505-2>
- Kufner, S. (2014). Was ist adaptives Lehren und wie lässt sich dessen Qualität empirisch erfassen? *Paradigma* (6), 55–67.
- Kultusministerkonferenz (2004). *Standards für die Lehrerbildung*. Abgerufen am 28.2.2022, von https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf.
- Kunina-Habenicht, O., Schulze-Stocker, F., Kunter, M., Baumert, J., Leutner, D., Förster, D., Lohse-Bossenz, H., & Terhart, E. (2013). Die Bedeutung der Lerngelegenheiten im Lehramtsstudium und deren individuelle Nutzung für den Aufbau des bildungswissenschaftlichen Wissens. *Zeitschrift für Pädagogik*, 59 (1), 1–23.
- Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction*, 17, 494–509.
<https://doi.org/10.1016/j.learninstruc.2007.09.002>

- Kunter, M., & Ewald, S. (2016). Bedingungen und Effekte von Unterricht: Aktuelle Forschungsperspektiven aus der pädagogischen Psychologie. In N. McElvany, W. Bos, H. G. Holtappels, M. M. Gebauer, & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 9–31). Waxmann.
- Kunter, M., Kleickmann, T., Klusmann, U. & Richter, D. (2011). Die Entwicklung professioneller Kompetenz von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 55–68). Waxmann.
<https://doi.org/10.31244/9783830974338>
- Kunter, M., & Klusmann, U. (2010). Kompetenzmessung bei Lehrkräften – Methodische Herausforderungen. *Unterrichtswissenschaft*, 38(1), 68–86.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, 105, 805–820. <https://doi.org/10.1037/a0032583>
- Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. UTB.
<https://doi.org/10.36198/9783838538952>
- Kunter, M., & Voss, T. (2011). Das Modell der Unterrichtsqualität in COACTIV: Eine multikriteriale Analyse. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 85–113). Waxmann. <https://doi.org/10.31244/9783830974338>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33 (1), 159-174. <https://doi.org/10.2307/2529310>
- Langner, A., & Jugel, D. (2019). Ohne Verstehen kein pädagogisches Handeln – Diagnostik im Kontext von Inklusion. In A. Langner, A., M. Ritter, J. Steffens, & D. Jugel (Hrsg.), *Inklusive Bildung forschend entdecken*. Springer.
https://doi.org/10.1007/978-3-658-25515-2_6
- Lauth, G., Brunstein, J., & Grünke, M. (2014). Lernstörungen im Überblick: Arten, Klassifikation, Verbreitung und Entwicklungsperspektiven. Interventionen bei Lernstörungen: Förderung, Training und Therapie in der Praxis. In G. W. Lauth, M. Grünke, & J. C. Brunstein (Hrsg.), *Interventionen bei Lernstörungen: Förderung, Training und Therapie in der Praxis* (2. Überarbeitete und erweiterte Auflage) Hogrefe.

- Lenske, G., Wagner, W., Wirth, J., Thillmann, H., Cauet, E., Liepertz, S., & Leutner, D. (2016). Die Bedeutung des pädagogisch-psychologischen Wissens für die Qualität der Klassenführung und den Lernzuwachs der Schüler/innen im Physikunterricht. *Zeitschrift für Erziehungswissenschaft*, 19(1), 211–233. <https://doi.org/10.1007/s11618-015-0659-x>
- Lersch, R. (2006). Lehrerbildung im Urteil der Auszubildenden. Eine empirische Studie zu beiden Phasen der Lehrerausbildung. In C. Allemann-Ghionda, & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern* (51. Beiheft der Zeitschrift für Pädagogik, 164-181). Beltz.
- Leuders, T., & Prediger, S. (2016). *Flexibel differenzieren und fokussiert fördern im Mathematikunterricht*. Cornelsen Scriptor.
- Lichtenberg, G. C. (o. D.). „Es gibt kein größeres Hindernis des Fortgangs in den Wissenschaften als das Verlangen, den Erfolg davon zu früh verspüren zu wollen.“ Gut zitiert. https://www.gutzitiert.de/zitat_autor_georg_christoph_lichtenberg_thema_wissenschaft_zitat_22389.html.
- Lienert, G. A., & Raatz, U. (2011). Testaufbau und Testanalyse [Setup and analysis of tests]. Beltz.
- Lindmeier, A. (2011). *Modeling and measuring knowledge and competencies of teachers*. Waxmann.
- Lindmeier, A. (2013). Videovignettenbasierte standardisierte Erhebung von Lehrerkognitionen. In U. Riegel, & K. Macha (Hrsg.), *Videobasierte Kompetenzforschung in den Fachdidaktiken* (S. 45–62). Waxmann.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler. In C. Allemann-Ghionda, & E. Terhart (Hrsg.), *Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern. Ausbildung und Beruf* (S. 47–70). Beltz.
- Lipowsky, F. (2009). Unterricht. In E. Wild, & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 74–101). Springer. https://doi.org/10.1007/978-3-540-88573-3_4
- Lipowsky, F. (2010). Lernen im Beruf – Empirische Befunde zur Wirksamkeit von Lehrerfortbildung. In F. Müller, A. Eichenberger, M. Lüders, & J. Mayr (Hrsg.), *Lehrerinnen und Lehrer lernen – Konzepte und Befunde zur Lehrerfortbildung* (S. 51-72). Waxmann.

- Lipowsky, F., & Bleck, V. (2019). Was wissen wir über guten Unterricht? – Ein Update. In U. Steffens, & R. Messner (Hrsg.), *Unterrichtsqualität. Konzepte und Bilanzen gelingenden Lehrens und Lernens. Grundlagen der Qualität von Schule* (Bd. 3, S. 219 – 249). Waxmann.
- Lipowsky, F., Drollinger-Vetter, B., Klieme, E., Pauli, C., & Reusser, K. (2018). Generische und fachdidaktische Dimensionen von Unterrichtsqualität – Zwei Seiten einer Medaille? In M. Martens, K. Rabenstein, K. Bräu, M. Fetzer, H. Gresch, I. Hardy, & C. Schelle (Hrsg.), *Konstruktionen von Fachlichkeit: Ansätze, Erträge und Diskussionen in der empirischen Unterrichtsforschung* (S. 183 – 202). Klinkhardt.
- Lipowsky F., Rakoczy, F., Klieme, E., Reusser, K. & Pauli, C. (2005). Unterrichtsqualität im Schnittpunkt unterschiedlicher Perspektiven. Rahmenkonzept und erste Ergebnisse einer binationalen Studie zum Mathematikunterricht in der Sekundarstufe I. In H. Holtappels, & K. Höhmann (Hrsg.), *Schulentwicklung und Schulwirksamkeit* (S. 223–238). Juventa.
- Lütje-Klose, B., & Neumann, P. (2018). Professionalisierung für eine inklusive Schule. *Basiswissen Lehrerbildung: Inklusion in Schule und Unterricht. Grundlagen in der Sonderpädagogik*, 129-151.
- Maag Merki, K., & Grob, U. (2005). Überfachliche Kompetenzen: Zur Validierung eines Indikatorensystems. In A. Frey (Hrsg.), *Kompetenzdiagnostik. Theorien und Methoden zur Erfassung und Bewertung von beruflichen Kompetenzen* (S. 7–30). Empirische Pädagogik e.V..
- Maag Merki, K., & Werner, S. (2011). Erfassung und Bewertung professioneller Kompetenz von Lehrpersonen. In E. Terhart, H. Bennewitz, & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (S. 573-591). Waxmann.
- Martschinke, S. (2015). Facetten adaptiven Unterrichts aus der Sicht der Unterrichtsforschung. In K. Liebers, B. Landwehr, A. Marquardt, & K. Schlotter (Hrsg.), *Lernprozessbegleitung und adaptives Lernen in der Grundschule. Forschungsbezogene Beiträge, Jahrbuch Grundschulforschung Bd. 19* (S. 15– 32. Springer. https://doi.org/10.1007/978-3-658-11346-9_2
- Marzano, R. (2000). *A new era of school reform: Going where the research takes us*. McREL.
- Mayer, A. (2016). Sprachliche Lernbarrieren beim Erwerb mathematischer Kompetenzen. In U. Stitzinger, S. Sallat, & U. Lüdtke, Ulrike (Hrsg.), *Sprache und Inklusion als Chance?!* (S. 269-278). Schulz-Kirchner.
- Mayring, P. (2008). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (10. Aufl.). Beltz. https://doi.org/10.1007/978-3-8349-9441-7_42

- Merkens, H. (2010). *Unterricht: Eine Einführung*. Verlag für Sozialwissenschaften.
<https://doi.org/10.1007/978-3-531-91910-2>
- Meschede, N. (2014). *Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. Theoretische Beschreibung und empirische Erfassung*. Logos.
- Meschede, N., & Steffensky, M. (2018). Methodologische Perspektive: Audiovisuelle Daten als Lerngelegenheiten in der Lehrer/innenbildung. In M. Sonnleitner, S. Prock, A. Rank, & P. Kirchhoff (Hrsg.), *Video- und Audiografie von Unterricht in der LehrerInnenbildung* (S. 21-36). Budrich.
- Meschede, N., Steffensky, M., Wolters, M., & Möller, K. (2015). Professionelle Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht. Theoretische Beschreibung und empirische Erfassung. *Unterrichtswissenschaft*, 43(4), 317-335.
- Meyer, H. (2004). *Was ist guter Unterricht?* Cornelsen.
- Möller, K. (2016). Bedingungen und Effekte qualitätvollen Unterrichts - ein Beitrag aus fachdidaktischer Perspektive. In N. McElvany, W. Bos, H. Holtappels, M. Gebauer, & F. Schwabe (Hrsg.), *Bedingungen und Effekte guten Unterrichts* (S. 43-64). Waxmann.
- Moje, E. B., & Wade, S. E. (1997). What case discussions reveal about teacher thinking. *Teaching and Teacher Education*, 13 (7), 691-712. [https://doi.org/10.1016/S0742-051X\(97\)00015-2](https://doi.org/10.1016/S0742-051X(97)00015-2)
- Moosbrugger, H., & Kelava, A. (2007). *Testtheorie und Fragebogenkonstruktion*. Springer.
<https://doi.org/10.1007/978-3-540-71635-8>
- Morris-Lange, S., Wagner, K., & Altinay, L. (2016). *Lehrerbildung in der Einwanderungsgesellschaft. Qualifizierung für den Normalfall Vielfalt*. Forschungsbereich beim Sachverständigenrat deutscher Stiftungen für Integration und Migration (SVR).
- Müller, M. M., & Gold, B. (2023). Videobasierte Erfassung wissensbasierten Verarbeitens als Teilprozess der professionellen Unterrichtswahrnehmung-Analyse eines geschlossenen und offenen Verfahrens. *Zeitschrift für Erziehungswissenschaft*, 26, 7–29.
<https://doi.org/10.1007/s11618-022-01128-6>
- Organisation für wirtschaftliche Zusammenarbeit und Entwicklung [OECD] (2016). *Low-performing students. Why they fall behind and how to help them succeed (PISA)*. OECD Publishing. <https://doi.org/10.1787/9789264250246-en>

- Oliver, R. M., Wehby, J. H., & Reschly, D. J. (2011). Teacher Classroom Management Practices: Effects on Disruptive or Aggressive Student Behavior. Society for Research on Educational Effectiveness. <https://doi.org/10.4073/csr.2011.4>
- Ollesch, J., Dörfler, T., & Vogel, M. (2018). Die inhaltliche Validierung von Unterrichtsvignetten durch eine mehrstufige Expertenbefragung. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß, & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 129-152). Springer. https://doi.org/10.1007/978-3-658-20121-0_7
- Ophardt, D., & Thiel, F. (2013). *Klassenmanagement: Ein Arbeitsbuch für die Schule*. Kohlhammer. <https://doi.org/10.17433/978-3-17-023845-9>
- Ophardt, D., & Thiel, F. (2017). Klassenmanagement als Basisdimension der Unterrichtsqualität. In M. K. W. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion* (S. 245–266). Springer. https://doi.org/10.1007/978-3-658-15083-9_11
- Orschulik, A. B. (2021). *Entwicklung der Professionellen Unterrichtswahrnehmung: Eine Studie zur Entwicklung Studierender in universitären Praxisphasen*. Springer. <https://doi.org/10.1007/978-3-658-33931-9>
- Oser, F., & Baeriswyl, F. J. (2001). Choreographies of teaching: Bridging instruction to learning. In V. Richardson (Hrsg.), *Handbook of research on teaching* (S. 1031 – 1065). American Educational Research Association.
- Oser, F., Dick, A., & Patry, J.-L. (Hrsg.) (1992). *Effective and responsible teaching. The new synthesis*. Jossey-Bass.
- Oser, F., Heinzer, S., & Salzmann, P. (2010). Die Messung der Qualität von professionellen Kompetenzprofilen von Lehrpersonen mit Hilfe der Einschätzung von Filmvignetten. Chancen und Grenzen des advokatorischen Ansatzes. *Unterrichtswissenschaft*, 38, 5–28.
- Otto, J., Migas, K., Austermann, N., & Bos, W. (2016). Integration neu zugewanderter Kinder und Jugendlicher ohne Deutschkenntnisse. Möglichkeiten, Herausforderungen und Perspektiven. Waxmann.
- Pauli, C. (2012). Merkmale guter Unterrichtsqualität im mathematisch-naturwissenschaftlichen Unterricht aus der Perspektive von Lernenden und Lehrpersonen. In R. Lazarides, & A. Ittel (Hrsg.), *Differenzierung im mathematisch-naturwissenschaftlichen Unterricht. Implikationen für Theorie und Praxis* (S. 13–34). Klinkhardt.

- Pea, R. D. (2004). The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. *Journal of the Learning Sciences*, 13 (3), 423-451. https://doi.org/10.1207/s15327809jls1303_6
- Pissarek, M., & Schilcher, A. (2015). Fachspezifische Lehrerkompetenzen im Fach Deutsch messen? Modellierung und Konstruktvalidierung eines Erhebungsinstrumentes im Rahmen der Projektgruppe FALKO Regensburg. In C. Bräuer, & D. Wieser (Hrsg.), *Lehrende im Blick. Empirische Lehrerforschung in der Deutschdidaktik* (S. 321-340). Springer. https://doi.org/10.1007/978-3-658-09734-9_15
- Piwowar, V. (2013). *Konzeptualisierung, Erfassung und Entwicklung von Kompetenzen im Klassenmanagement. Dissertation zur Erlangung des akademischen Grades Doktor der Philosophie (Dr. phil.)*, Freie Universität Berlin.
- Praetorius, A.-K., & Charalambous, C. Y. (2018). Classroom observation frameworks for studying instructional quality: looking back and looking forward. *ZDM*, 50(3), 535–553. <https://doi.org/10.1007/s11858-018-0946-0>
- Praetorius, A. K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM*, 50, 407 – 426. <https://doi.org/10.1007/s11858-018-0918-4>
- Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K. & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2–12. <https://doi.org/10.1016/j.learninstruc.2013.12.002>
- Prenzel, M., Gogolin, I., & Krüger, H.-H. (2007). Editorial zum Themenheft Kompetenzdiagnostik. *Zeitschrift für Erziehungswissenschaft*, 10(8), 5–8.
- Przibilla, B., Lauterbach, A., Boshold, F., Linderkamp, F., & Krezmien, M. (2016). Entwicklung und Validierung eines Online-Surveys zur Erhebung von Kompetenzen und Einstellungen von Lehrkräften bezüglich der Inklusion. *Empirische Sonderpädagogik* 1, 36-53.
- Rakoczy, K. (2008). *Motivationsunterstützung im Mathematikunterricht. Unterricht aus der Perspektive von Lernenden und Beobachtern*. Waxmann.
- Rakoczy, K., & Pauli, C. (2006). Hoch inferentes Rating: Beurteilung der Qualität unterrichtlicher Prozesse. In I. Hugener, C. Pauli, & K. Reusser (Hrsg.), *Dokumentation der Erhebungs- und Auswertungsinstrumente zur schweizerisch-deutschen Videostudie „Unterrichtsqualität, Lernverhalten und mathematisches Verständnis“*. *Videoanalysen* (Materialien zur Bildungsforschung, Bd. 15, S. 206–233). GPPF.

- Renkl, A. (2011). Aktives Lernen: Von sinnvollen und weniger sinnvollen theoretischen Perspektiven zu einem schillernden Konstrukt. *Unterrichtswissenschaft, 39*, 197–212.
- Reusser, K., & Pauli, C. (2010). Unterrichtsgestaltung und Unterrichtsqualität - Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht: Einleitung und Überblick. In K. Reusser, C. Pauli, & M. Waldis (Hrsg.), *Unterrichtsgestaltung und Unterrichtsqualität – Ergebnisse einer internationalen und schweizerischen Videostudie zum Mathematikunterricht* (S. 9-32). Waxmann.
- Reyer, T. (2004). *Oberflächenmerkmale und Tiefenstrukturen im Unterricht*. Logos.
- Ricken, G. (2017). Kompetent sein für Inklusive Schulen heisst auch Diagnostizieren lernen. In M. Gercke, S. Opalinski, & T. Thonagel (Hrsg.), *Inklusive Bildung und gesellschaftliche Exklusion* (S. 187–199). Springer. https://doi.org/10.1007/978-3-658-17084-4_12
- Ricken, G., & Degenhardt, (2016). Sonderpädagogik in allgemeinen und sonderpädagogischen Lehramtsstudiengängen? Die Gretchenfrage nach den Basics und dem Spezifischen einer Profession. In K.-P. Becker (Hrsg.), *Bewährtes bewahren - Neues gestalten: Wissenschaftliches Symposium aus Anlass des 90. Geburtstages von Prof. em. Dr. paed. habil. Klaus-Peter Becker* (S. 85–97). Stiftung Rehabilitationszentrum Berlin-Ost.
- Ricken, G., & Schuck, K. D. (2011). Pädagogische Diagnostik und Lernen. *Behinderung, Bildung, Partizipation, 4*, 110-119.
- Riese, J., & Reinhold, P. (2012). Die professionelle Kompetenz angehender Physiklehrkräfte in verschiedenen Ausbildungsformen. *Zeitschrift für Erziehungswissenschaft, 15*(1), 111–143. <https://doi.org/10.1007/s11618-012-0259-y>
- Rössler, P. (2017). Inhaltsanalyse (Vol. 2671). UTB. <https://doi.org/10.36198/9783838547060>
- Rogers, B., & Kühn, M. (2013). *Classroom Management: Das Praxisbuch*. Beltz.
- Rosenberger, K. (2016). Fall-Vignetten. Ein methodisches Instrument in der Bildungsforschung. https://www.researchgate.net/publication/341372392_Fall-Vignetten_Ein_methodisches_Instrument_in_der_Bildungsforschung
- Rostampour, P. (1998). Die Zustimmung- und Ablehnungstendenzen im Zusammenhang mit der Itemrichtung: positiv vs. negativ formulierte Items. *ZA-Information/Zentralarchiv für Empirische Sozialforschung, 42*, 148-169.
- Roth, K. J., Garnier, H., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Video-based lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching, 48* (2), 117-148. <https://doi.org/10.1002/tea.20408>

- Rutsch, J., Seidenfuß, M., Vogel, M., Dörfler, T., & Rehm, M. (2017). Fachdidaktische Unterrichtsvignetten in Forschung und Lehre: Überblick über Forschungsarbeiten und Einsatzmöglichkeiten. *Beiträge zur Lehrerinnen- und Lehrerbildung* 35(3), 487-505. <https://doi.org/10.36950/bzl.35.3.2017.9461>
- Rutsch, J., Vogel, M., Rehm, M., & Dörfler, T. (2018a). Modellierung der Testletstruktur bei vignettenbasierten Testverfahren mit geschlossenem Antwortformat. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 27-46). Springer. https://doi.org/10.1007/978-3-658-20121-0_2
- Rutsch, J., Vogel, M., Seidenfuß, M., Dörfler, T., & Rehm, M. (2018b). Professionalisierungsprozesse angehender Lehrkräfte untersuchen. In J. Rutsch, M. Rehm, M. Vogel, M. Seidenfuß, & T. Dörfler (Hrsg.), *Effektive Kompetenzdiagnose in der Lehrerbildung. Professionalisierungsprozesse angehender Lehrkräfte untersuchen* (S. 9-26). Springer. https://doi.org/10.1007/978-3-658-20121-0_1
- Sabers, D., Cushing, K., & Berliner, D. (1991). Differences Among Teachers in a Task Characterized by Simultaneity, Multidimensional, and Immediacy. *American Educational Research Journal*, 28(1), 63–88. <https://doi.org/10.3102/00028312028001063>
- Santagata, R., & Angelici, G. (2010). Studying the impact of the lesson analysis framework on preservice teachers' abilities to reflect on videos of classroom teaching. *Journal of Teacher Education*, 61 (4), 339-349. <https://doi.org/10.1177/0022487110369555>
- Santagata, R., & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM - Mathematics Education*, 43(1), 133–145. <https://doi.org/10.1007/s11858-010-0292-3>
- Santagata, R., & Yeh, C. (2013). Learning to teach mathematics and to analyse teaching effectiveness. Evidence from a video- and practice-based approach. *Journal of Mathematics Teacher Education*, 17(6), 491-514. <https://doi.org/10.1007/s10857-013-9263-2>
- Santagata, R., & Yeh, C. (2016). The role of perception, interpretation, and decision making in the development of beginning teachers' competence. *ZDM - Mathematics Education*, 48(1), 153–165. <https://doi.org/10.1007/s11858-015-0737-9>
- Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: An empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education*, 10(2), 123–140. <https://doi.org/10.1007/s10857-007-9029-9>

- Schäfer, S., & Seidel, T. (2015). Noticing and reasoning of teaching and learning components by preservice teachers. *Journal for Educational Research Online*, 7(2), 34–58.
- Schneider, J., Bohl, T., Kleinknecht, M., Rehm, M., Kuntze, S., & Syring, M. (2016). Unterricht analysieren und reflektieren mit unterschiedlichen Fallmedien: Ist Video wirklich besser als Text? *Unterrichtswissenschaft*, 44 (4), 474-490.
- Schnell, R., Hill, P., & Esser, E. (2018). *Methoden der empirischen Sozialforschung*. De Gruyter.
- Schnurr, S. (2003). Vignetten in quantitativen und qualitativen Forschungsdesigns. In H.-U. Otto, G. Oelerich, & H.-G. Micheel (Hrsg.), *Empirische Forschung und Soziale Arbeit. Ein Lehr- und Arbeitsbuch* (S. 393–400). Luchterhand.
- Scholten, N., Höttecke, D., & Sprenger, S. (2018). Conceptualizing Geography Teachers` Subject-Specific Noticing during Instruction. *European Journal of Geography*, 9(3), 80–97.
- Scholten, N., Höttecke, D., & Sprenger, S. (2019). How do geography teachers notice critical incidents during instruction? *International Research in Geographical and Environmental Education*, 3(2), 1–15. <https://doi.org/10.1080/10382046.2019.1670915>
- Schrader, F.-W., Helmke, A., & Dotzler, H. (1997). Zielkonflikte in der Grundschule. In F. E. Weinert & A. Helmke (Hrsg.), *Entwicklung im Grundschulalter*. (S. 299-316). Psychologie Verlags Union.
- Schüppel, J. (1996). *Wissensmanagement: Organisatorisches Lernen im Spannungsfeld von Wissens- und Lernbarrieren*. Gabler.
- Schwindt, K. (2008). *Lehrpersonen betrachten Unterricht*. Waxmann.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed). Sage Publications.
- Seidel, T. (2003a). Sichtstrukturen – Organisation unterrichtlicher Aktivitäten. In T. Seidel, M. Prenzel, R. Duit, & M. Lehrke (Hrsg.), *Technischer Bericht zur Videostudie „Lehr-Lern-Prozesse im Physikunterricht“* (S. 113–127). Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN).
- Seidel, T. (2003b). *Lehr-Lernskripts im Unterricht*. Waxmann.
- Seidel, T. (2015). Klassenführung. In: E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 107-119). Springer. https://doi.org/10.1007/978-3-642-41291-2_5

- Seidel, T., Blomberg, G., & Stürmer, K. (2010). "Observer" – Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. Projekt OBSERVE. In E. Klieme, D. Leutner, & M. Kenk (Hrsg.), *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes*. 56. Beiheft der Zeitschrift für Pädagogik (S. 296-306). Beltz.
- Seidel, T. & Prenzel, M. (2007). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen – Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen. *Zeitschrift für Erziehungswissenschaft*, 10 (Sonderheft 8), 201-216. https://doi.org/10.1007/978-3-531-90865-6_12
- Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77(4), 454–499. <https://doi.org/10.3102/0034654307310317>
- Seidel, T., & Stürmer, K. (2014). Modeling and measuring the structure of professional vision in preservice teachers. *American Educational Research Journal*, 51(4), 739- 771. <https://doi.org/10.3102/0002831214531321>
- Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M. & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, 27(2), 259–267. <https://doi.org/10.1016/j.tate.2010.08.009>
- Seifert, A., Hilligus, A., & Schaper, N. (2009). Entwicklung und psychometrische Überprüfung eines Messinstruments zur Erfassung pädagogischer Kompetenzen in der universitären Lehrerbildung. *Lehrerbildung auf dem Prüfstand*, 2(1), 82–103.
- Seiffert, H. (2016). Lernbarrieren beim Fachwortlernen – zum Beispiel Mathematik. In U. Stitzinger, S. Sallat, & U. Lüdtke (Hrsg.), *Sprache und Inklusion als Chance?!* (S. 279-285). Schulz-Kirchner.
- Seifried, J., & Wuttke, E. (2017). Der Einsatz von Videovignetten in der wirtschaftspädagogischen Forschung: Messung und Förderung von fachwissenschaftlichen und fachdidaktischen Kompetenzen angehender Lehrpersonen. In C. Gräsel, & K. Trempler (Hrsg.), *Entwicklung von Professionalität pädagogischen Personals. Interdisziplinäre Betrachtungen, Befunde und Perspektiven* (S. 303-322). Springer. https://doi.org/10.1007/978-3-658-07274-2_16

- Sherin, M. G. (2001). Developing a professional vision of classroom events. In T. L. Wood, B. S. Nelson, & J. Warfield (Hrsg.), *Beyond classical pedagogy. Teaching elementary school mathematics* (S. 75-93). Lawrence Erlbaum.
- Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron, & S. Derry (Hrsg.), *Video research in the learning sciences* (S. 383-395). Erlbaum.
- Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, 60, 20-37.
<https://doi.org/10.1177/0022487108328155>
- Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011a). Situating the study of teacher noticing. In M. G. Sherin, V.R. Jacobs, & R.A. Philipp (Hrsg.), *Mathematics Teacher Noticing. Seeing Through Teachers' Eyes* (S. 3-13). Routledge. <https://doi.org/10.4324/9780203832714>
- Sherin, M. G., Russ R. S., & Colestock, A. A. (2011b). Accessing Mathematics Teachers' In-the-Moment Noticing. In M. G. Sherin, V. R. Jacobs & R. A. Philipp (Hrsg.), *Mathematics Teacher Noticing. Seeing Through Teachers' Eyes* (S.79-94). Routledge.
<https://doi.org/10.4324/9780203832714>
- Sherin, B., & Star, J. R. (2011). Reflections on the study of teacher noticing. In M. G. Sherin, V.R. Jacobs, & R.A. Philipp (Hrsg.), *Mathematics Teacher Noticing. Seeing Through Teachers' Eyes* (S. 66-78). Routledge. <https://doi.org/10.4324/9780203832714>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4-14. <https://doi.org/10.2307/1175860>
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Spiro, R. J., Feltovich, P. J., Jacobson, M. J., & Coulson, R. L. (1991). Cognitive flexibility, constructivism, and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains. *Educational Technology*, 31 (5), 24-33.
- Spitzer, S. M., Phelps, C. M., Beyers, J. E., Johnson, D. Y., & Sieminski, E. M. (2011). Developing prospective elementary teachers' abilities to identify evidence of student mathematical achievement. *Journal of Mathematics Teacher Education*, 14, 67-87.
<https://doi.org/10.1007/s10857-010-9141-0>
- Spoto, D. (2010). *The art of Alfred Hitchcock: Fifty years of his motion pictures*. Anchor.

- Stang, J., & Urhahne, D. (2016). Wie gut schätzen Lehrkräfte Leistung, Konzentration, Arbeits- und Sozialverhalten ihrer Schülerinnen und Schüler ein? Ein Beitrag zur diagnostischen Kompetenz von Lehrkräften. *Psychologie in Erziehung und Unterricht*, 63, 204-219. <https://doi.org/10.2378/peu2016.art18d>
- Star, J. R. & Strickland, S. K. (2008). Learning to observe: using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education*, 11(2), 107–125. <https://doi.org/10.1007/s10857-007-9063-7>
- Steffensky, M., Gold, B., Holodynski, M., & Möller, K. (2015). Professional vision of classroom management and learning support in science classrooms - does professional vision differ across general and content-specific classroom interactions? *International Journal of Science and Mathematics Education*, 13(2), 351–368. <https://doi.org/10.1007/s10763-014-9607-0>
- Stiehler, S., Fritsche, C., & Reutlinger, C. (2012). Der Einsatz von Fall-Vignetten. Potential für sozialräumliche Fragestellungen. *sozialraum.de (4)1*. URL: <https://www.sozialraum.de/der-einsatz-von-fall-vignetten.php>, Datum des Zugriffs: 03.01.2024
- Stockero, S. L., & Rupnow, R. L. (2017). Measuring noticing within complex mathematics classroom interactions. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Hrsg.), *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks* (S. 281- 301). Springer. https://doi.org/10.1007/978-3-319-46753-5_17
- Streiner, D. L. (2003). Starting at the beginning: An introduction to coefficient alpha and internal consistency. *Journal of Personality Assessment*, 80, 99–103. https://doi.org/10.1207/S15327752JPA8001_18
- Stürmer, K., & Seidel, T. (2015). Assessing Professional Vision in Teacher Candidates. *Zeitschrift für Psychologie*, 223(1), 54–63. <https://doi.org/10.1027/2151-2604/a000200>
- Stürmer, K., & Seidel, T. (2017). Connecting generic pedagogical knowledge with practice. In S. Guerriero (Hrsg.), *Pedagogical knowledge and the changing nature of the teaching profession* (S. 137-149). OECD. <https://doi.org/10.1787/9789264270695-8-en>
- Sunder, C., Todorova, M., & Möller, K. (2016). Kann die professionelle Unterrichtswahrnehmung von Sachunterrichtsstudierenden trainiert werden? Konzeption und Erprobung einer Intervention mit Videos aus dem naturwissenschaftlichen Grundschulunterricht. *Zeitschrift für Didaktik der Naturwissenschaften*, 22(1), 1–12. <https://doi.org/10.1007/s40573-015-0037-5>

- Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction*, 12(3), 185–233. https://doi.org/10.1207/s1532690xci1203_1
- Syring, M., Bohl, T., Kleinknecht, M., Kuntze, S., Rehm, M., & Schneider, J. (2015). Video or text in case-based teacher education? An examination of the effects of different media on cognitive load and motivational-emotional processes in case-based learning. *Zeitschrift für Erziehungswissenschaft*, 18(4), 667–685. <https://doi.org/10.1007/s11618-015-0631-9>
- Taubert, R. T. (2007). *Classroom management: Sound theory and effective practice* (4. Aufl.). Praeger.
- Templer, K. J., & Lange, S. R. (2008). Internet testing: Equivalence between proctored lab and unproctored field conditions. *Computers in Human Behavior*, 24, 1216–1228. <https://doi.org/10.1016/j.chb.2007.04.006>
- Tepner, O., & Dollny, S. (2014). Entwicklung eines Testverfahrens zur Analyse fachdidaktischen Wissens. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 311-323). Springer Spektrum. https://doi.org/10.1007/978-3-642-37827-0_25
- Terhart, E. (2007). Erfassung und Beurteilung der beruflichen Kompetenzen von Lehrkräften. In M. Lüders, & J. Wissinger (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* (S. 37-62). Waxmann.
- Terhart, E., Schulze, F., Kunina-Habenicht, O., Dicke, T., Förster, D., Lohse-Bossenz, H., Gößling, J., Kunter, M., Baumert, J., & Leutner, D. (2012). Bildungswissenschaftliches Wissen und der Erwerb professioneller Kompetenz in der Lehramtsausbildung – Eine Kurzdarstellung des BilWiss-Projekts. *Lehrerbildung auf dem Prüfstand*, 5(1), 96–106.
- Teuscher, D., Leatham, K. R., & Peterson, B. E. (2017). From a framework to a lens: Learning to notice student mathematical thinking. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Hrsg.), *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks* (S. 31-48). Springer. https://doi.org/10.1007/978-3-319-46753-5_3
- Thiel, F. (2021). Herausforderungen des Klassenmanagements – Störungsprävention und -intervention im Unterricht. *Schulverwaltung – Fachzeitschrift für Schulentwicklung und Schulmanagement*, 9, 67–70.
- Thiel, F., Richter, S. G., & Ophardt, D. (2012). Steuerung von Übergängen im Unterricht. Eine Experten-Novizen-Studie zum Klassenmanagement. *Zeitschrift für Erziehungswissenschaft*, 15, 727–752. <https://doi.org/10.1007/s11618-012-0325-5>

- Timmermans, A. C., Rubie-Davies, C. M., & Wang, S. (2021). Adjusting expectations or maintaining first impressions? *Learning and Instruction*, 75. <https://doi.org/10.1016/j.learninstruc.2021.101483>
- Tönnissen, L., Ricken, G., Wenck, S., & Hövel, D. C. (2023). Beobachten lernen in Hamburg und Zürich. Erprobung von Lehrkonzepten in unterschiedlichen Seminarkontexten. *Emotionale und soziale Entwicklung in der Pädagogik der Erziehungshilfe und bei Verhaltensstörungen: ESE*, 5(5), 184-196. <https://doi.org/10.35468/6021-13>
- Vieluf, S., Praetorius, A-K., Rackoczy, K., Kleinknecht, M., & Pietsch, M. (2020). Angebots-Nutzungs-Modelle der Wirkweise des Unterrichts: Ein kritischer Vergleich verschiedener Modellvarianten. *Zeitschrift für Pädagogik, Beiheft*, (66), 63-80.
- Vock, M., & Gronastaj, A. (2017). *Umgang mit Heterogenität in Schule und Unterricht*. Friedrich-Ebert-Stiftung.
- Vogelsang, C. (2014). *Validierung eines Instruments zur Erfassung der professionellen Handlungskompetenz von (angehenden) Physiklehrkräften. Zusammenhangsanalysen zwischen Lehrerkompetenz und Lehrerperformanz*. Logos.
- Vogler, J. (2019). *Professionelle Entscheidungen im Sportunterricht. Eine empirische Untersuchung zum fachdidaktischen Wissen von Lehrpersonen*. Springer. <https://doi.org/10.1007/978-3-658-28691-0>
- Vogt, J. (2013). *Aspekte erzählender Prosa: eine Einführung in Erzähltechnik und Romantheorie*. Springer. <https://doi.org/10.36198/9783838540566>
- Vogt, F., & Rogalla, M. (2009). Developing Adaptive Teaching Competency through Coaching. *Teaching and Teacher Education*, 25, S. 1051–1060. <https://doi.org/10.1016/j.tate.2009.04.002>
- Voss, T., Kunina-Habenicht, O., Hoehne, V., & Kunter, M. (2015). Stichwort pädagogisches Wissen von Lehrkräften: Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, 18 (2), 187-223. <https://doi.org/10.1007/s11618-015-0626-6>
- Voss, T., & Kunter, M. (2011). Pädagogisch-psychologisches Wissen von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV* (S. 193-214). Waxmann. <https://doi.org/10.31244/9783830974338>

- Voss, T., Kunter, M., & Baumert, J. (2011). Assessing teacher candidates' general pedagogical and psychological knowledge: Test construction and validation. *Journal of Educational Psychology, 103*(4), 952-969. <https://doi.org/10.1037/a0025125>
- Voss, T., Kunter, M., Seiz, J., Hoehne, V., & Baumert, J. (2014). Die Bedeutung des pädagogisch-psychologischen Wissens von angehenden Lehrkräften für die Unterrichtsqualität. *Zeitschrift für Pädagogik, 60*(2), 184-201.
- Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993). Toward a knowledge base for school learning. *Review of Educational Research, 63*(3), 249 – 294. <https://doi.org/10.3102/00346543063003249>
- Weber, K. E., Prilop, C. N., Viehoff, S., Gold, B., & Kleinknecht, M. (2020). Fördert eine videobasierte Intervention im Praktikum die professionelle Wahrnehmung von Klassenführung? – Eine quantitativ-inhaltsanalytische Messung von Subprozessen professioneller Wahrnehmung. *Zeitschrift für Erziehungswissenschaft, 23*(2), 343-365. <https://doi.org/10.1007/s11618-020-00939-9>
- Weinert, F. E. (2001a). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Hrsg.), *Defining and selecting key competencies* (S. 45–65). Hogrefe.
- Weinert, F. E. (2001b). Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 17-32). Beltz.
- Weinert, F. E., Schrader, F.-W., & Helmke, A. (1989). Quality of instruction and achievement outcomes. *International Journal of Educational Research, 13*, 895-914. [https://doi.org/10.1016/0883-0355\(89\)90072-4](https://doi.org/10.1016/0883-0355(89)90072-4)
- Wenck, S., Fischer-Böhm, A., & Ricken, G. (im Druck). Systematisches Beobachten von Unterrichtsprozessen als Basis für die Erkennung von Lernbarrieren. Entwicklung, Erprobung und Evaluation eines Seminarkonzepts. In G. Kaiser, E. Arnold, & J. Doll (Hrsg.), *Innovative Ansätze zur Veränderung der Lehrkräfteausbildung*. Waxmann.
- Werth, L., Denzler, M., & Mayer, J. (2020). Soziale Wahrnehmung. In L. Werth, B. Seibt, & J. Mayer (Hrsg.), *Sozialpsychologie – Das Individuum im sozialen Kontext* (S. 137-186). Springer. https://doi.org/10.1007/978-3-662-53897-5_5
- Wiedmann, M. (2015). *Teacher's professional vision of student interaction in collaborative learning settings: development and evaluation of an assessment tool*. Dissertation, Albert-Ludwigs-Universität Freiburg.

- Wiedmann M., Leuders T. (2019). Entwicklung eines computergestützten Messinstruments für die professionelle Wahrnehmung von Schülerinteraktionen. In T. Leuders, M. Nückles, S. Mikelskis-Seifert, & K. Philipp (Hrsg.) *Pädagogische Professionalität in Mathematik und Naturwissenschaften*. Springer. <https://doi.org/10.1007/978-3-658-08644-2>
- Wöhlke, C. (2020). Entwicklung und Validierung eines Instruments zur Erfassung der professionellen Unterrichtswahrnehmung angehender Physiklehrkräfte. (Volume 298). Logos. <https://doi.org/10.30819/5149>
- Wolff, C. E., Jarodzka, H., & Boshuizen, H. P. A. (2017). See and tell: Differences between expert and novice teachers' interpretations of problematic classroom management events. *Teaching and Teacher Education*, 66, 295-308. <https://doi.org/10.1016/j.tate.2017.04.015>
- Wolters, M. (2014). *Wie kompetent sind (angehende) Lehrkräfte in der professionellen Wahrnehmung kognitiv anregender Situationen im naturwissenschaftlichen Grundschulunterricht?* Inaugural-Dissertation, Westfälische Wilhelms-Universität, Münster.
- World Health Organization (2001). *International Classification of Functioning, Disability and Health: ICF*. World Health Organization.
- Wygotski, L. (1987). *Ausgewählte Schriften. Band 2: Arbeiten zur psychischen Entwicklung der Persönlichkeit*. Pahl-Rugenstein.
- Yadav, A., Philips, M. M., Lundeberg, M. A., Koehler, M. J., Hilden, K., & Dirkin, K. H. (2011). If a picture is worth a thousand words is video worth a million? Differences in affective and cognitive processing of video and text cases. *Journal of Computing in Higher Education*, 23 (1), 15-37. <https://doi.org/10.1007/s12528-011-9042-y>
- Zimbardo, P. G. (1988). *Psychologie*. Springer.
- Zucker, V. (2019). Erkennen und Beschreiben von formativem Assessment im naturwissenschaftlichen Grundschulunterricht: Entwicklung eines Instruments zur Erfassung von Teilfähigkeiten der professionellen Wahrnehmung von Lehramtsstudierenden (Vol. 281). Logos.

Tabelle B12	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studiengangsgruppen für die gewichteten Mittelwerte für Faktor 2, getrennt nach Studiendauer.....	263
Anhang C: Zweiter Entwicklungsschritt	264
C1: Bepunktung der Items in Vignette Luise	264
Tabelle C2	Häufigkeiten für die Beantwortungen der Items zu Barrieren in Vignette Luise	265
Tabelle C3	Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der zweiten Datenerhebung	266
Tabelle C4	Komponentenkorrelationsmatrix zur Faktorenanalyse der zweiten Datenerhebung	266
Tabelle C5	Teststatistiken des Kruskal-Wallis-Tests bei unabhängigen Stichproben zum Einfluss der Studiendauer auf die gewichteten Mittelwerte für Faktor 1	266
Tabelle C6	Paarweise Vergleiche der Fachsemester zum Einfluss der Studiendauer auf die gewichteten Mittelwerte für Faktor 1.....	267
Anhang D: Dritter Entwicklungsschritt	268
D1: Lösung Vignette 1	268
D2: Lösung Vignette 2	270
D3: Lösung Vignette 3	272
D4: Lösung Vignette 4	275
Tabelle D5	Häufigkeiten für die Beantwortungen der Items zu Barrieren und der Items zu gelingenden Bedingungen in der dritten Datenerhebung	277
Tabelle D6	Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1	278
Tabelle D7	Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2	278
Tabelle D8	Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3	279
Tabelle D9	Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4	279

Tabelle D10	Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der dritten Datenerhebung (Studierende der Universität Hamburg).....	280
Tabelle D11	Komponentenkorrelationsmatrix zur Faktorenanalyse der dritten Datenerhebung.....	281
Tabelle D12	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studienganggruppen im 1. und 2. Bachelorsemester anhand gewichteter Mittelwerte je Vignette (Studierende der Universität Hamburg).....	282
Tabelle D13	Teststatistiken des Kruskal-Wallis-Test bei unabhängigen Stichproben zum Einfluss der Studiendauer auf die gewichteten Mittelwerte je Vignette (Studierende der Universität Hamburg im Studiengang Sonderpädagogik)	282
Tabelle D14	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte je Vignette	283
Tabelle D15	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 2	284
Tabelle D16	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 2	284
Tabelle D17	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 3	285

Tabelle D18	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 4	285
Tabelle D19	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 4	286
Tabelle D20	Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der dritten Datenerhebung (Studierende anderer Universitäten)	287
Tabelle D21	Komponentenkorrelationsmatrix zur Faktorenanalyse der dritten Datenerhebung (Studierende anderer Universitäten)	288
Tabelle D22	Teststatistiken des Kruskal-Wallis-Tests bei unabhängigen Stichproben zum Einfluss der Studiendauer auf die gewichteten Mittelwerte je Vignette (Studierende der Allgemeinpädagogik, andere Universitäten)	289
Tabelle D23	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studienganggruppen im 4. bis 6. Bachelorsemester anhand gewichteter Mittelwerte je Vignette (Studierende anderer Universitäten)	289
Tabelle D24	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studienganggruppen im 4. bis 6. Bachelorsemester anhand gewichteter Mittelwerte für Vignette 2 (Studierende anderer Universitäten)	290
Tabelle D25	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studienganggruppen im 4. bis 6. Bachelorsemester anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 2 (Studierende anderer Universitäten)	290

Tabelle D26	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte je Vignette	291
Tabelle D27	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 2	291
Tabelle D28	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 2	292
Tabelle D29	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 3	292
Tabelle D30	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 4	293
Anhang E: Vierter Entwicklungsschritt		294
E1: Lösung und Werte Vignette 1		294
E2: Lösung und Werte Vignette 2		297
E3: Lösung und Werte Vignette 3		300
E4: Lösung und Werte Vignette 4		303
Tabelle E5	Paarweise Vergleiche des Friedman-Tests zur Einschätzung der Authentizität der Vignetten (Bachelorstudierende)	306
Tabelle E6	Häufigkeiten für die Beantwortungen der Items zu Barrieren und der Items zu gelingenden Bedingungen in der vierten Datenerhebung	307
Tabelle E7	Reliabilitätsstatistik zur Gesamtskala der vierten Datenerhebung	308

Tabelle E8	Reliabilitätsstatistik zu Vignette 1 der vierten Datenerhebung.....	308
Tabelle E9	Reliabilitätsstatistik zu Vignette 2 der vierten Datenerhebung.....	308
Tabelle E10	Reliabilitätsstatistik zu Vignette 3 der vierten Datenerhebung.....	308
Tabelle E11	Reliabilitätsstatistik zu Vignette 4 der vierten Datenerhebung.....	309
Tabelle E12	Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der vierten Datenerhebung.....	310
Tabelle E13	Komponentenkorrelationsmatrix zur Faktorenanalyse der vierten Datenerhebung	311
Tabelle E14	Teststatistiken des Kruskal-Wallis-Tests bei unabhängigen Stichproben zum Einfluss der Expertise auf die gewichteten Mittelwerte je Vignetten und auf die Items zu Barrieren und zu gelingenden Situationen	311
Tabelle E15	Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für	311
Tabelle E16	Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für Vignette 4	312
Tabelle E17	Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für Items zu	312
Tabelle E18	Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für Items zu gelingenden Situationen	312
Tabelle E19	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen der Klassenatmosphäre (Bachelorstudierende).....	313
Tabelle E20	Paarweise Vergleiche der Einschätzungen der Klassenatmosphäre (Bachelorstudierende).....	313
Tabelle E21	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen der Klassenatmosphäre (Masterstudierende).....	314

Tabelle E22	Paarweise Vergleiche der Einschätzung der Klassenatmosphäre (Masterstudierende)	314
Tabelle E23	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen der Klassenatmosphäre (Lehrkräfte).....	315
Tabelle E24	Paarweise Vergleiche der Einschätzung der Klassenatmosphäre (Lehrkräfte)	315
Tabelle E25	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu Barrieren je Vignette (Bachelorstudierende)	316
Tabelle E26	Paarweise Vergleiche der Einschätzungen von Items zu Barrieren (Bachelorstudierende)	316
Tabelle E27	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu Barrieren je Vignette (Masterstudierende)	317
Tabelle E28	Paarweise Vergleiche der Einschätzungen von Items zu Barrieren (Masterstudierende).....	317
Tabelle E29	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu Barrieren je Vignette (Lehrkräfte)	318
Tabelle E30	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu gelingenden Situationen je Vignette (Bachelorstudierende)	318
Tabelle E31	Paarweise Vergleiche der Einschätzungen von Items zu gelingenden Situationen (Bachelorstudierende).....	319
Tabelle E32	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu gelingenden Situationen je Vignette (Masterstudierende).....	319

Tabelle E33	Paarweise Vergleiche der Einschätzungen von Items zu gelingenden Situationen (Masterstudierende)	320
Tabelle E34	Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu gelingenden Situationen je Vignette (Lehrkräfte)	320
Tabelle E35	Paarweise Vergleiche der Einschätzungen von Items zu gelingenden Situationen (Lehrkräfte).....	321
Tabelle E36	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1 (Bachelorstudierende)	322
Tabelle E37	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2 (Bachelorstudierende)	323
Tabelle E38	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3 (Bachelorstudierende)	324
Tabelle E39	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4 (Bachelorstudierende)	325
Tabelle E40	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1 (Masterstudierende)	326
Tabelle E41	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2 (Masterstudierende)	327

Tabelle E42	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3 (Masterstudierende)	328
Tabelle E43	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4 (Masterstudierende)	329
Tabelle E44	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1 (Lehrkräfte)	330
Tabelle E45	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2 (Lehrkräfte)	331
Tabelle E46	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3 (Lehrkräfte)	332
Tabelle E47	Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4 (Lehrkräfte)	333
Tabelle E48	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 1	333

Tabelle E49	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 2	334
Tabelle E50	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 3	334
Tabelle E51	Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 4	334

Anhang A: Vignettenentwicklung

A1: Masterlösung Vignette *Philipp*

Vignettenabschnitt	Zuordnung	Kommentar
<p>Philipp sitzt während des Matheunterrichts in seiner Klasse an seinem Gruppentisch. Vor ihm liegt ein Arbeitsblatt mit einer Textaufgabe. Die anderen Kinder an seinem Tisch haben bereits in Gruppenarbeit mit der Bearbeitung der Aufgabe begonnen und diskutieren lebhaft. Philipp soll die Aufgabe mit Hilfe seiner Lehrerin lösen.</p>	KF, Monitoring: Raum/Arbeitsituation	Philipp sitzt am Gruppentisch, nimmt aber an keiner Gruppenarbeit teil. Die Zusammenarbeit mit der Lehrkraft könnte unter diesen Bedingungen und bei dem Geräuschpegel („diskutieren lebhaft“) schwierig sein.
<p>Lehrerin: „Wir beide gucken jetzt mal zusammen, wie du die Textaufgabe lösen kannst. Hier steht: ‚Für die Musik-Aufführung in der Schulturnhalle stellt der Hausmeister 280 Stühle auf. Außerdem benötigt der Elternbeirat noch 20 Stühle für den Verkauf. Im Orchester spielen 24 Kinder, die während der Aufführung stehen. Die Stühle sind in 6er Stapeln im Material-Lager gestapelt.‘</p> <p>Okay? (wartet kurz, Philipp nickt)</p>	LU, inhaltliche Strukturierung: Inhaltliche Klarheit	Die Lehrkraft strukturiert die Aufgabe nicht, sondern liest nur noch mal vor. Der Input ist damit rein sprachlich. Die Nachfrage „Okay“ ist keine ausreichende Verständnissicherung.
<p>Dann wollen wir jetzt mal überlegen, wie du das angehen kannst. Ich würde ganz gut finden, wenn du laut denkst, damit ich schauen kann, welche Rechenstrategien du dir überlegst und wie wir die nutzen können. Und dann kann ich dir immer dann helfen, wenn du allein nicht weiterkommst.</p>	LU, inhaltliche Strukturierung: Zielklarheit	Die Lehrkraft äußert zuerst, dass sie gemeinsam überlegen wollen, gibt die Schritte dann aber vor. Sie nennt mehrere Aspekte (laut denken, Rechenstrategien, Hilfestellung), verdeutlicht aber nicht das Vorgehen. Es bleibt unklar, ob Philipp nachdenken oder erklären oder etwas aufschreiben soll.
<p>Was haben wir denn da jetzt alles für Informationen? Vielleicht fängst du mal mit der Frage an?“</p> <p>Die Lehrerin bleibt neben Philipp stehen und wartet, Philipp schaut auf den Text und bleibt still.</p>	KF, prozessuale Strukturierung: Übergänge	Sie bietet Hilfestellung an und erwartet dann aber eine selbstständige Bearbeitung, weshalb der Übergang in die Arbeitsphase nicht gelingt. Es ist unklar, ob die Lehrkraft die Frage zur Textaufgabe oder Philipps Fragen meint. Sie leitet nicht in die Arbeitsphase über, sondern bleibt nur stehen

<p>Dann rempelt Philipp seinen Sitznachbarn an und schimpft: „Mann, jetzt rück doch mal ein Stück weg!“</p> <p>Lehrerin: „Na, na, Philipp, du hast so viel Platz wie immer. Nun fang mal an!“</p>	<p>KF, Monitoring: Allgegenwärtigkeit, Reaktion auf Störung</p>	<p>und wartet, so dass die Situation hier ins Stocken gerät.</p> <p>Die Lehrkraft hat nicht die Klasse und die Arbeitssituation im Blick, während sie mit Philipp beschäftigt ist. Sie zieht nicht in Betracht, dass die Arbeitssituation für Philipp ungünstig sein könnte, z. B. auch aufgrund der Umgebungsgereusche und der anderen arbeitenden Kinder am Gruppentisch. Sie reagiert nicht angemessen auf die Störung (erst zu spät als Philipp deshalb schimpft und sie unterbindet diese nicht).</p>
<p>Philipp: „Und wie?“</p> <p>Lehrerin: „Wie gesagt, vielleicht mit der Frage. Versuch mal.“</p> <p>Philipp: „Wie geht die denn?“</p> <p>Lehrerin: „Das wollen wir ja jetzt rausfinden. Wie könnten wir vorgehen?“</p> <p>Philipp schaut seine Lehrerin erwartungsvoll an, diese wartet geduldig.</p>	<p>LU, inhaltliche Strukturierung: Maßnahmen zur Veranschaulichung</p>	<p>Die Lehrkraft bietet keine Lernunterstützung. Sie klärt nicht auf, wo das Verständnisproblem von Philipp liegt, und ergreift keine Maßnahmen zur Veranschaulichung.</p>
<p>Ein Kind vom Nachbartisch ruft derweil: „Wir sind schon fertig!“</p> <p>Philipp knallt sauer den Stift auf den Tisch: „Oh Mann, ich hab keine Lust! Ich mach das nicht!“</p>	<p>KF, Regeln und Routinen</p>	<p>Es gibt keine Regelung für unterschiedliche Arbeitstempi bei der Gruppenarbeit und keine Regelung, wie sich Kinder verhalten, wenn sie früher fertig sind, so dass das Reinrufen zur Frustration von Philipp führen könnte.</p>
<p>Lehrerin: „Na, komm. Wenn du dich schon wieder von vornherein weigerst, dann kommen wir nie weiter...“</p>	<p>KF/LU, Feedback</p>	<p>Die Lehrkraft gibt kein lernbezogenes, sondern ein persönliches Feedback. Sie unterstellt ein Nicht-Wollen und impliziert, dass es in Philipps Verantwortung liegt, ob er Fortschritte macht.</p>

Anhang B: Erster Entwicklungsschritt

B1: Codierleitfaden Vignette *Philipp*

Bedeutung der Werte

Wert 0: Es wird kein Bezug auf die Kategorie oder Textstelle genommen.

Wert 1: Es wird Bezug auf eine Textstelle genommen, aber eher nacherzählend oder beschreibend.

Wert 2: Es wird Bezug auf eine Textstelle genommen und bewertet oder beurteilt, die Äußerung lässt erkennen, dass die Barriere richtig erkannt wurde.

Wert 3: Die Barriere wurde erkannt und es wird ein Verbesserungsvorschlag gemacht (die Qualität des Verbesserungsvorschlags fließt nicht in die Wertevergabe ein).

Wert 4: Es wird zwar Bezug auf eine Textstelle genommen, aber etwas unterstellt, was dem Text nicht entnehmbar ist, oder etwas formuliert, das erkennen lässt, dass die Barriere nicht erkannt wurde.

Kategorie 1:

Klassenführung, Monitoring, Raum/Arbeitssituation: Philipp sitzt am Gruppentisch, nimmt aber nicht an der Gruppenarbeit teil, Hinweise auf hohen Umgebungslärm, Lehrerin *steht* neben ihm.

Beispiele für Wert 1:

1. Philipp sitzt am Gruppentisch.
2. Philipp löst die Aufgabe mit der Lehrerin.
3. Lehrerin sitzt/steht bei Philipp.
4. Philipp macht nicht bei der Gruppenarbeit mit.

Beispiele für Wert 2:

- Es besteht ein Widerspruch zwischen Sitzordnung und Arbeitsform.
- Philipp sitzt mit am Gruppentisch, nimmt aber an Gruppenarbeit nicht teil.
- Andere arbeiten schon, Philip nicht, dadurch entsteht ein ungleiches Arbeitstempo/Ablenkung/Lärm.

Beispiele für Wert 3:

- Philipp sollte auch allein sitzen, wenn er allein arbeitet.

- Philipp bräuchte eine ruhigere Arbeitsumgebung.
- Die Lehrerin sollte besser mit Philipp am Tisch *sitzen*.

Beispiele für Wert 4:

1. Philipp arbeitet nicht in einer Gruppe, weil er unbeliebt ist.
2. Im Gegensatz zu allen anderen am Tisch verweigert Phillip die Arbeit.
3. Die anderen Kinder finden es nicht gut, dass Philipp eine Sonderbehandlung bekommt.

Kategorie 2:

Lernunterstützung, inhaltliche Strukturierung, inhaltliche Klarheit: Die Lehrerin liest die Aufgabe nur noch mal vor, ohne etwas zu erklären, der Input ist damit rein sprachlich und zu komplex. Die Aufgabe hat eine eher hohe sprachliche Anforderung und eine Frage ist in der Aufgabe nicht formuliert. Die Nachfrage „Okay?“ stellt keine hinreichende Verständnissicherung dar. (Fokus auf Aufgabe)

Beispiele für Wert 1:

- Die Lehrerin liest vor.
- Die Lehrerin fragt, ob Philipp die Aufgabe verstanden hat.
- Philipp nickt.

Beispiele für Wert 2:

- Durch einmaliges Vorlesen kann sich Phillip nicht alle Informationen merken.
- Die Aufgabe ist zu komplex/sprachlich zu anspruchsvoll.
- Philipp hat die Aufgabe nicht verstanden und die Lehrerin merkt das nicht, Nicken ist keine echte Bestätigung für Verständnis.
- Die Lehrerin erklärt die Aufgabe nicht, so dass unklar ist, ob Philipps Leseschwäche das Problem ist oder er die Aufgabe mathematisch nicht versteht.

Beispiele für Wert 3:

- Lehrerin könnte Philipp unterstützen, den Text selbst zu erlesen und zu verstehen.
- Philipp bräuchte eine textfreie Rechenaufgabe oder eine Visualisierung der Aufgabe.
- Philip müsste Schritt für Schritt bei der Aufgabe vorgehen.
- Die Lehrerin müsste die Aufgabe für Philipp strukturieren.

Beispiel für Wert 4:

1. Philipp versteht die Aufgabe nicht, obwohl sie erklärt wurde.

2. Philipp hat nicht zugehört und deshalb nichts verstanden
3. Philipp verweigert die Arbeit.

Kategorie 3:

Lernunterstützung, inhaltliche Strukturierung, Zielklarheit: Die Arbeitsinstruktion ist unklar und zu komplex. Die Lehrerin verdeutlicht nicht, wie Philipp vorgehen soll. Es bleibt unklar, ob Philipp nur mündlich arbeiten soll oder was er aufschreiben soll. (Fokus auf Arbeitsinstruktion)

Beispiele für Wert 1:

- Lehrerin fordert Philipp auf anzufangen.
- Die Lehrerin macht einen Vorschlag.
- Philipp weiß nicht, was er tun soll (ohne Begründung, woran das liegt).
- Philipp nickt, fängt aber nicht an.

Beispiele für Wert 2:

- Die Lehrerin gibt zu unklare Instruktionen, so dass Philipp nicht weiß, wie er anfangen soll.
- Die Lehrerin stellt zu viele Fragen auf einmal, so dass der Arbeitsauftrag nicht klar wird.
- Philipp kennt vielleicht „lautes Denken“ nicht und ihm ist das Ziel davon nicht klar.

Beispiele für Wert 3:

- Die Lehrerin müsste Teilschritte erklären.
- Die Lehrerin müsste eine genauere Arbeitsinstruktion geben.
- Die Lehrerin müsste sich einfacherer Sprache bedienen, damit Philipp klar wird, wie er vorgehen soll.

Beispiel für Wert 4:

- Die Lehrerin gibt Philipp genau vor, was er tun soll, aber Philipp verweigert die Mitarbeit.
- Die Lehrerin gibt sich viel Mühe bei der Erklärung.
- Die Lehrerin erklärt die Aufgabe wie sonst auch immer.

Kategorie 4:

Klassenführung, Monitoring, Allgegenwärtigkeit, Reaktion auf Störung: Die Lehrkraft hat nicht den Rest der Klasse im Blick, während sie mit Philipp arbeitet, sie hat nicht im Blick, ob Philipp sich vielleicht zurecht eingeengt fühlt oder es ihm zu laut ist. Die Reaktion kommt möglicherweise zu spät (erst als Philipp wütend wird) und sie beseitigt die Störung auch nicht.

Beispiele für Wert 1:

- Philipp stößt andere weg.
- Philipp reagiert frustriert.
- Die Lehrerin tadelt Philipp.

Beispiele für Wert 2:

- Lehrerin geht zu spät/nicht angemessen darauf ein, dass Philipp sich unwohl fühlt.
- Die Lehrerin konzentriert sich nur auf Philipp und bekommt nichts anderes mit.
- Die Ermahnung der Lehrerin ist falsch/kontraproduktiv, weil sie sich nur an Philipp richtet und nicht die anderen im Blick hat.

Beispiele für Wert 3:

- Die Lehrerin könnte nach Anrempeln anders/unterstützend statt ermahmend auf Philipp eingehen.
- Die Lehrerin müsste auch im Blick haben, was die anderen machen und für Ruhe sorgen.
- Die Lehrerin hätte vorher dafür sorgen müssen, dass es Philipp erst gar nicht zu laut oder zu eng wird.

Beispiel für Wert 4:

- Die Mitschüler mobben Philipp.
- Philipp sucht aktiv/absichtlich Ablenkung, weil er die Aufgabe nicht machen will.
- Das zeigt, dass sich Philipp grundsätzlich leicht ablenken lässt.
- Philipp hat ein schlechtes Verhältnis zu Mitschülern/ist aggressiv gegenüber anderen.

Kategorie 5:

Klassenführung, Prozessuale Strukturierung, Übergänge: Die Lehrerin bietet Hilfestellung an, die sie dann aber nicht gibt, so dass der Übergang ins Arbeiten stockt. Es ist unklar, ob sich die Aufforderung eine Frage zu formulieren auf die Frage der Textaufgabe oder Philipps Fragen bezieht. Die gibt kein eindeutiges Signal, sondern wartet nur, so dass Philipp nicht anfängt. (Fokus auf Übergang in selbständige Arbeit, Arbeitsfluss)

Beispiele für Wert 1:

- Philipp fängt nicht an/wartet.
- Die Lehrerin wartet.
- Philipp stellt die Frage nicht.

Beispiele für Wert 2:

- Philipp fängt nicht selbstständig an zu arbeiten, weil er keine passende Hilfestellung hat.
- Philipp fragt nach Hilfe, bekommt diese aber nicht, so dass er nicht selbstständig weiterarbeiten kann.
- Lehrerin spricht von „wir“/verspricht Hilfe, erwartet aber dass Philipp allein weitermacht/keine Aufforderung zur selbständigen Arbeit.

Beispiele für Wert 3:

1. Statt zu sagen: „Wir gucken, wie *du* das lösen kannst“, sollte die Lehrerin besser eine Hilfestellung zum selbstständigen Anfangen geben.
2. Die Lehrerin sollte Bedingungen schaffen, damit Philipp anfangen kann zu arbeiten.
3. Die Lehrerin sollte ein klares Signal geben und nicht nur warten.

Beispiel für Wert 4:

- Philipp beginnt nicht, weil er keine Eigenverantwortung übernimmt.
- Philipp möchte, dass die Lehrerin ihm vorsagt.
- Philipp kann grundsätzlich nicht selbstständig arbeiten.

Kategorie 6:

Lernunterstützung, inhaltliche Strukturierung, Maßnahmen zur Veranschaulichung: Die Lehrerin ergründet Philipps Verständnisproblem nicht und bietet keine passende Unterstützung durch Veranschaulichung der Aufgabe an.

Beispiele für Wert 1:

- Philipp weiß nicht, wie es geht.
- Philipp braucht Hilfe.
- Die Lehrerin wartet geduldig.

Beispiele für Wert 2:

- Philipp kann nicht arbeiten, weil er andere/mehr Unterstützung bräuchte.
- Lehrerin merkt nicht, dass Philipp den Text nicht verstanden hat
- Die Lehrerin erklärt/veranschaulicht die Aufgabe nicht.
- Die Lehrerin stellt mehrfach die gleiche Frage, ohne wirklich zu helfen.

Beispiele für Wert 3:

4. Die Lehrerin sollte Philipp anders unterstützen und... (z. B. einzelne Arbeitsschritte vorgeben).
5. Die Lehrerin sollte nicht nur mündliche Hilfestellung geben.
6. Die Lehrerin sollte für Philipp wichtige Informationen im Text hervorheben.

Beispiel für Wert 4:

- Lehrerin bietet Unterstützung an, aber Philipp weigert sich.
- Philipp weiß die Geduld der Lehrerin nicht zu schätzen.
- Philipp kann sich nicht konzentrieren.

Kategorie 7:

Klassenführung, Regeln und Routinen: Es gibt keine Regelung für unterschiedliche Arbeitstempi, schon fertige Gruppe ruft rein, weil sie keine Anschlussaufgabe hat.

Beispiele für Wert 1:

- Ein Kind ruft rein.
- Eine andere Gruppe ist schon fertig.
- Philipp wird sauer.

Beispiele für Wert 2:

- Die Gruppen haben ein unterschiedliches Arbeitstempo und es entsteht dadurch Unruhe.
- Es gibt keine Regelung/Weiterbeschäftigung für schnellere Gruppen.
- Es gibt keine Regel, die das Reinrufen unterbindet.

Beispiele für Wert 3:

- Es wäre besser, wenn auch die schnelleren Gruppen noch beschäftigt wären.
- Es sollte eine Regel geben, was diejenigen machen, die schneller fertig sind.
- Die Lehrerin könnte Aufgaben/Rollen/Anschlussaufgaben in den Gruppen verteilen.
- Es könnte so geregelt werden, dass die Kinder, die schon fertig sind, Philipp helfen.

Beispiel für Wert 4:

- Philipp reagiert auf den Zwischenruf, weil er eine geringe Frustrationstoleranz hat.
- Die anderen Kinder wollen Philipp absichtlich ärgern.
- Die Klasse ist sehr leistungsstark in Mathe und arbeitet deshalb schnell.

Kategorie 8:

Feedback/Lernunterstützung: Die Lehrkraft gibt kein lernbezogenes, sondern ein persönliches Feedback. Sie unterstellt ein Nicht-Wollen und impliziert, dass es in Philipps Verantwortung liegt, ob er Fortschritte macht.

Beispiele für Wert 1:

- Die Lehrkraft sagt: „...“
- Die Lehrkraft gibt Philipp Feedback.
- Die Lehrkraft schimpft mit Philipp.

Beispiele für Wert 2:

- Die Lehrerin unterstellt Philipp Weigerung, obwohl Philipp sich nicht von vornherein weigert.
- Die Lehrerin sieht den Grund für die Probleme bei Philipp statt in der Aufgabe oder der Situation.
- Die Lehrerin gibt pauschal negatives/persönliches statt konstruktiven Feedbacks.
- Die Wortwahl des Feedbacks ist unangemessen.

Beispiele für Wert 3:

- Die Lehrerin könnte Verständnis äußern.
- Die Lehrerin könnte den Grund für die Probleme erfragen.
- Die Lehrerin könnte sachliches Feedback geben und einen konstruktiven Vorschlag machen.

Beispiel für Wert 4:

- Die Lehrerin gibt nicht auf und motiviert Philipp weiter.
- Das Feedback beweist, dass Philipp regelmäßig die Arbeit verweigert.
- Die Lehrerin ist zurecht von Philipp genervt.

Beispiele für nicht codierbare Aussagen:

Zuschreibungen von Eigenschaften oder Einstellungen, Interpretationen oder Unterstellungen, die dem Text nicht entnehmbar sind, sowie Kommentare, die sich auf andere Aspekte als den Unterricht beziehen.

- Philipp sieht sich selbst als schlechten Schüler.
- Lesen hat keinen Wert für Philip.
- Philipp spricht nicht so gerne.
- Philipp beschäftigt sich lieber mit anderen Dingen.

- Philipp will so gut sein wie die anderen.
- Die Lehrerin ist sympathisch.
- Das ist eine typische Textaufgabe.
- Solche Situationen kommen immer wieder vor.
- Die Klasse ist heterogen.

Tabelle B2

Teststatistiken des Mann-Whitney-U-Tests, Vergleich der Mittelwerte zu Verständlichkeit, Schwierigkeit und Interessantheit in Abhängig vom bearbeiteten Aufgabenformat

	Verständlichkeit	Schwierigkeit	Interessantheit
Mann-Whitney-U-Test	5663.000	5489.500	5546.000
Wilcoxon-W	11658.000	11160.500	11217.000
Z	-.012	-.823	-.785
Asymp. Sig. (2-seitig)	.991	.410	.432

Anmerkung. Gruppenvariable: Zugehörigkeit zur Testgruppe, die die offene oder geschlossene Aufgabe bearbeitet hat.

Tabelle B3*Zuordnung der genannten Kategorien zu Gelingensbedingungen*

Kategorie	Fokus
Vorleseverständnis, Sprachverständnis	Individuelle Voraussetzungen
Sehen	Individuelle Voraussetzungen
Hören	Individuelle Voraussetzungen
Sprechen und Sprache	Individuelle Voraussetzungen
Andere physische und kognitive Voraussetzungen	Individuelle Voraussetzungen
Vorläuferfähigkeiten	Individuelle Voraussetzungen
Leseverständnis	Individuelle Voraussetzungen
Phonem-Graphem-Korrespondenz	Individuelle Voraussetzungen
Silbenlesen, Synthese	Individuelle Voraussetzungen
Lesestrategien	Individuelle Voraussetzungen
Allgemeine Problemanalyse Lesen	Individuelle Voraussetzungen
Lesegeschwindigkeit und -flüssigkeit	Individuelle Voraussetzungen
Phase des Schriftspracherwerbs	Individuelle Voraussetzungen
Interessen und Lesemotivation	Individuelle Voraussetzungen
Mehrsprachigkeit	Individuelle Voraussetzungen
Häusliches Umfeld	Individuelle Voraussetzungen
Lesesituation, sozial	Unterrichtsbedingungen
Lesesituation, räumlich	Unterrichtsbedingungen
Förderung und Nachteilsausgleiche	Unterrichtsbedingungen
Soziale Situation und Integration	Klassenatmosphäre
Verhalten allgemein	Individuelle Voraussetzungen
Verhalten bei Leseaufgaben	Individuelle Voraussetzungen
Arbeitsorganisation	Individuelle Voraussetzungen
Affektiv-motivationale Aspekte	Individuelle Voraussetzungen
Aufgaben und Material	Unterrichtsbedingungen
Schulleistung Deutsch im Vergleich zur Klasse	Individuelle Voraussetzungen
Schulleistung in anderen Fächern	Individuelle Voraussetzungen
Verhalten der Lehrkraft	Lehrkraft
Selbstreflexion der Lehrkraft	Lehrkraft
Sicherung von Grundbedürfnissen	Individuelle Voraussetzungen

Tabelle B4*Inter-Rater-Reliabilität bei drei Raterinnen für die zweite offene Aufgabe*

Kategorie/Barriere	Fleiss Kappa
Raum	.439
Inhaltliche Klarheit	.515
Zielklarheit	.381
Allgegenwärtigkeit	.465
Übergänge	.333
Veranschaulichung	.266
Regeln und Routinen	.569
Feedback	.504

Tabelle B5*Inter-Rater-Reliabilität nach Ausschluss einer Raterin für die zweite offene Aufgabe*

Kategorie/Barriere	Cohens Kappa
Raum	.558
Inhaltliche Klarheit	.601
Zielklarheit	.573
Allgegenwärtigkeit	.644
Übergänge	.569
Veranschaulichung	.391
Regeln und Routinen	.733
Feedback	.564

Tabelle B6*Häufigkeiten für die Nennungen der Kategorien der zweiten offenen Aufgabe, getrennt nach Studiengängen*

Kategorie	nicht genannt		beschreibend genannt		beurteilt		verbessert		Falsch interpretiert	
	%		%		%		%		%	
	Gruppe 1	Gruppe 2	Gruppe 1	Gruppe 2	Gruppe 1	Gruppe 2	Gruppe 1	Gruppe 2	Gruppe 1	Gruppe 2
Raum	50.0	42.9	11.9	5.7	28.6	47.1	9.5	2.9	0.0	1.4
Inhaltliche Klarheit	45.2	25.7	11.9	7.1	21.4	47.1	19.0	18.6	2.4	1.4
Zielklarheit	66.7	75.7	14.3	4.3	11.9	14.3	7.1	5.7	0.0	0.0
Allgegenwärtigkeit	69.0	41.1	16.7	35.7	2.4	14.3	2.4	0.0	9.5	8.6
Übergänge	81.0	65.7	4.8	7.1	7.1	15.7	0.0	2.9	7.1	8.6
Veranschaulichung	73.8	70.0	4.8	1.4	21.4	22.9	0.0	4.3	0.0	1.4
Regeln und Routinen	50.0	42.9	42.9	50.0	4.8	0.0	2.4	5.7	0.0	0.0
Feedback		75.7		5.7		47.1		2.9		1.4

Anmerkung. Gruppe 1 = Studierende Allgemeinpädagogik ($n = 42$), Gruppe 2 = Studierende Sonderpädagogik ($n = 70$). Angabe in gültigen Prozent

Tabelle B7*Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der ersten Datenerhebung*

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2.869	35.862	35.862	2.869	35.862	35.862
2	1.482	18.528	54.391	1.482	18.528	54.391
3	1.097	13.707	68.098	1.097	13.707	68.098
4	.748	9.345	77.443			
5	.584	7.305	84.748			
6	.522	6.529	91.278			
7	.391	4.891	96.168			
8	.307	3.832	100.000			

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse.**Tabelle B8***Komponentenkorrelationsmatrix zur Faktorenanalyse der ersten Datenerhebung*

Komponente	1	2
1	1.000	.203
2	.203	1.000

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung.

Tabelle B9

Teststatistiken des Kruskal-Wallis-Tests bei unabhängigen Stichproben zum Einfluss der Studiendauer auf die gewichteten Mittelwerte für Faktor 1 und Faktor 2, getrennt nach Studienganggruppen

Studiengang		Faktor1 ge- wichtet	Faktor2 gewichtet
Allgemeinpädagogik	Kruskal-Wallis-H	5.341	1.476
	df	2	2
	Asymp. Sig.	.069	.478
Sonderpädagogik	Kruskal-Wallis-H	.099	1.631
	df	2	2
	Asymp. Sig.	.952	.442

Anmerkung. Gruppenvariable: Fachsemester

Tabelle B10

Paarweise Vergleiche der Fachsemester zum Einfluss der Studiendauer auf die gewichteten Mittelwerte für Faktor 1 und Faktor 2, getrennt nach Studienganggruppen

Studiengang	Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardtest- statistik	Sig.	Anp. Sig. ^a
Allgemeinpädagogik	1. BA – 5. BA	-7.558	4.727	-1.599	.110	.330
	1. BA – 3. MA	-9.156	4.308	-2.125	.034	.101
	5. BA – 3. MA	-1.597	5.151	-.310	.757	1.000
Sonderpädagogik	1. BA – 5. BA	-.885	6.712	-.132	.895	1.000
	1. BA – 3. MA	-1.678	5.337	-.314	.753	1.000
	5. BA – 3. MA	-.793	6.595	-.120	.904	1.000

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst. BA = Bachelorsemester, MA = Mastersemester

Tabelle B11

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studiengangsgruppen für die gewichteten Mittelwerte für Faktor 1, getrennt nach Studiendauer

1. Bachelorsemester	Gesamtzahl	45
	Mann-Whitney-U-Test	375.000
	Wilcoxon-W	726.000
	Teststatistik	375.000
	Standardfehler	43.495
	Standardisierte Teststatistik	2.943
	Asymptotische Sig. (zweiseitiger Test)	.003
5. Bachelorsemester	Gesamtzahl	22
	Mann-Whitney-U-Test	69.500
	Wilcoxon-W	160.500
	Teststatistik	69.500
	Standardfehler	14.958
	Standardisierte Teststatistik	.735
	Asymptotische Sig. (zweiseitiger Test)	.462
	Exakte Sig. (zweiseitiger Test)	.471
3. Mastersemester	Gesamtzahl	41
	Mann-Whitney-U-Test	168.500
	Wilcoxon-W	603.500
	Teststatistik	168.500
	Standardfehler	34.872
	Standardisierte Teststatistik	-.158
	Asymptotische Sig. (zweiseitiger Test)	.875
	Exakte Sig. (zweiseitiger Test)	.877

Tabelle B12

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studienganggruppen für die gewichteten Mittelwerte für Faktor 2, getrennt nach Studiendauer

1. Bachelorsemester	Gesamtzahl	45
	Mann-Whitney-U-Test	334.500
	Wilcoxon-W	685.500
	Teststatistik	334.500
	Standardfehler	43.213
	Standardisierte Teststatistik	2.025
	Asymptotische Sig. (zweiseitiger Test)	.043
5. Bachelorsemester	Gesamtzahl	22
	Mann-Whitney-U-Test	56.000
	Wilcoxon-W	147.000
	Teststatistik	56.000
	Standardfehler	14.852
	Standardisierte Teststatistik	-.168
	Asymptotische Sig. (zweiseitiger Test)	.866
	Exakte Sig. (zweiseitiger Test)	.896
3. Mastersemester	Gesamtzahl	41
	Mann-Whitney-U-Test	181.000
	Wilcoxon-W	616.000
	Teststatistik	181.000
	Standardfehler	34.679
	Standardisierte Teststatistik	.202
	Asymptotische Sig. (zweiseitiger Test)	.840
	Exakte Sig. (zweiseitiger Test)	.854

Anhang C: Zweiter Entwicklungsschritt

C1: Bepunktung der Items in Vignette *Luise*

	Trifft ganz und gar nicht zu	Trifft kaum zu	Trifft weniger zu	Trifft eher zu	Trifft weitgehend zu	Trifft voll und ganz zu
Die Arbeitssituation ist für Luise angemessen (Raum)	6	5	4	3	2	1
Die Aufgabenstellung ist für Luise angemessen (Aufgabenstellung, Inhaltliche Klarheit)	6	5	4	3	2	1
Die Lehrerin ist freundlich zu Luise (Ablenker)						
Die Lehrerin stellt nicht sicher, dass Luise die Aufgabe versteht (Lernunterstützung, Verständnissicherung)	1	2	3	4	5	6
Die Lehrerin bietet Luise passende Lernunterstützung an (Lernunterstützung)	6	5	4	3	2	1
Luise bemüht sich, die Aufgabe zu bearbeiten (Ablenker)						
Die Lehrerin versäumt es, Luise das Vorgehen bei der Bearbeitung zu verdeutlichen (Zielklarheit)	1	2	3	4	5	6
Luise lässt sich beim Bearbeiten der Aufgabe ablenken (Ablenker)						
Die Lehrerin hat die Gesamtsituation im Griff und reagiert passend auf Störungen (Allgegenwärtigkeit)	6	5	4	3	2	1
Die Lehrerin bietet Luise keine Anleitung zur selbständigen Bearbeitung der Aufgabe an (Reibungslosigkeit und Schwung, Übergänge)	1	2	3	4	5	6

Die besondere Schwierigkeit der Aufgabe liegt darin, dass die Kinder die Frage selbst formulieren sollen (Ablenker)

Der Ablauf der Gruppenarbeit erfolgt nach festen Regeln und Routinen (Regeln und Routinen) 6 5 4 3 2 1

Die Lehrerin gibt Luise unangemessene Rückmeldungen (Feedback) 1 2 3 4 5 6

Es handelt sich bei der Sachaufgabe um ein gängiges Aufgabenformat (Ablenker)

Tabelle C2

Häufigkeiten für die Beantwortungen der Items zu Barrieren in Vignette Luise

Item	n	Trifft voll und ganz zu	Trifft weitgehend zu	Trifft eher zu	Trifft weniger zu	Trifft kaum zu	Trifft ganz und gar nicht zu
Raum	141	0	0	3	27	57	54
Inhaltliche Klarheit	140	1	1	10	30	58	40
Veranschaulichung (n)	142	32	35	26	31	5	13
Lernunterstützung	142	1	3	11	39	59	29
Zielklarheit (n)	139	23	58	28	16	12	2
Allgegenwärtigkeit	141	0	3	16	41	51	30
Übergänge (n)	141	31	42	33	23	9	3
Regeln und Routinen	132	4	11	18	39	42	18
Feedback (n)	138	15	31	44	29	12	7

Anmerkung. Angaben in absoluten Zahlen. (n) = negativ formulierte Items, die richtige Antwort war jeweils *trifft voll und ganz zu*

Tabelle C3*Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der zweiten Datenerhebung*

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen ^a
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	
1	2.008	22.311	22.311	2.008	22.311	22.311	1.909
2	1.298	14.426	36.737	1.298	14.426	36.737	1.489
3	1.150	12.782	49.519				
4	1.046	11.624	61.143				
5	.877	9.739	70.883				
6	.717	7.963	78.845				
7	.695	7.717	86.563				
8	.646	7.183	93.746				
9	.563	6.254	100.000				

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse.**Tabelle C4***Komponentenkorrelationsmatrix zur Faktorenanalyse der zweiten Datenerhebung*

Komponente	1	2
1	1.000	.161
2	.161	1.000

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung.**Tabelle C5***Teststatistiken des Kruskal-Wallis-Tests bei unabhängigen Stichproben zum Einfluss der Studiendauer auf die gewichteten Mittelwerte für Faktor 1*

Gesamtzahl	142
Teststatistik	8.090 ^a
Freiheitsgrad	2
Asymptotische Sig. (zwei-seitiger Test)	.018

Anmerkung. a. Die Teststatistik wird für Bindungen angepasst.

Tabelle C6

Paarweise Vergleiche der Fachsemester zum Einfluss der Studiendauer auf die gewichteten Mittelwerte für Faktor 1

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig.a
2. BA - 4. MA	-13.414	8.501	-1.578	.115	.344
2. BA - 6. BA	-22.928	8.191	-2.799	.005	.015
4. MA - 6. BA	9.514	9.159	1.039	.299	.897

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst. BA = Bachelorsemester, MA = Mastersemester.

Anhang D: Dritter Entwicklungsschritt

D1: Lösung Vignette 1

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
<p>Unterrichtsbeginn Deutsch 4. Klasse. Die Kinder sitzen an Partnerischen.</p> <p>Der Lehrerin fällt auf, dass Murat noch Kaugummi kaut. Lehrerin: „Murat, spuck bitte dein Kaugummi aus!“ Murat steht auf, spuckt sein Kaugummi in den Müll, setzt sich wieder hin. Lehrerin: „Hanna, spuck dein Kaugummi aus!“ Hanna steht auf, spuckt ihr Kaugummi aus. In der Zwischenzeit fängt die Klasse an zu kichern – Murat hat sich ein neues Kaugummi in den Mund gesteckt. Lehrerin: „Murat, spuck dein Kaugummi aus, muss ich das jedem einzeln sagen?“</p>	Regeln und Routinen Barriere	<p>Hinweis auf Partnertische (für Sitzordnung relevant).</p> <p>Es sind keine Regeln bzgl. des Kaugummikauens erkennbar, so dass Zeit verloren geht und Unruhe entsteht. Die Lehrerin fordert die Kinder zudem nacheinander auf, statt kurz und bündig darauf zu reagieren, was zusätzlich Zeit kostet.</p>
<p>Wir wollen uns heute noch mal mit den Wortarten beschäftigen. Ich schreib euch drei Sätze an die Tafel und unterstreiche einige Wörter. Ihr müsst die nicht abschreiben. Anschließend sucht ihr euch einen Satz aus und überlegt erst mal allein für euch, mit welcher Frage ihr nach den unterstrichenen Wörtern fragen könnt (erklärt ein Beispiel). Jetzt schreib ich erst mal an.“</p>	Klarheit gelingend	Die Lehrerin benennt das Thema und macht deutlich, was die Kinder zu tun haben. Sie verdeutlicht durch ein Beispiel.
<p>Die Lehrerin schreibt an die Tafel. Es wird immer unruhiger. Tom beginnt, Papierkügelchen zu werfen, Luise zeigt Franz, die hinter ihr sitzt, ihre neuen Sticker. Jonas ruft schließlich laut: „Oh Mann, Leute!“ Lehrerin: „Ruhe jetzt! Jonas!“</p>	Allgegenwärtigkeit Barriere	Die Lehrerin hat die Klasse nicht im Blick, während sie anschreibt. Sie reagiert zu spät auf die sich ausbreitende Unruhe und ermahnt das falsche Kind.
<p>Sucht einen Satz aus und notiert die Frage zu den unterstrichenen Wortarten. Ihr habt 5 Minuten Zeit.“</p>	Übergänge Barriere	Erster Hinweis zum Übergang, strikte Zeitvorgabe.

Die Stillarbeit beginnt. Emma stöhnt: „Oh Mann, Paul redet dauernd vor sich hin, das macht der immer! Das nervt!“ Lehrerin: „Versucht jetzt bitte mal, 5 Minuten still zu arbeiten.“	Sitzordnung/Raum Barriere	Der Sitzplatz ist für Emma ungünstig, weil sie sich gestört fühlt. Die Lehrkraft berücksichtigt das nicht.
Nach 5 Minuten beendet die Lehrerin die Stillarbeit und beginnt, Ergebnisse zu sammeln, ein paar Kinder schreiben noch. Lehrerin: „Legt mal die Stifte hin und macht mit, sonst werden wir nicht fertig, ihr sollt das nachher auch noch geordnet aufschreiben.“	Übergänge Barriere	Die Lehrerin beginnt nach genau 5 Minuten mit dem Sammeln von Ergebnissen, obwohl einige Kinder noch schreiben. Ihre Ankündigung macht deutlich, dass das Zeit-Management nicht gelingt und zu wenig Zeit für die geplanten Schritte zur Verfügung steht.
Paul: „Das ist doof, ich kann das irgendwie nicht.“ Lehrerin: „Du passt ja auch nie auf und quasselst immer!“	Feedback Barriere	Die Lehrerin gibt kein konstruktives Feedback und unterstützt nicht, sondern gibt ein negatives, persönliches Feedback.

	Trifft sehr unwahrscheinlich zu	Trifft eher unwahrscheinlich zu	Trifft eher wahrscheinlich zu	Trifft sehr wahrscheinlich zu
Die technische Ausrüstung in der Klasse ist ausreichend (Ablenker)				
Die Lehrerin hat die Sitzordnung im Blick (Raum)	4	3	2	1
Die Lehrerin reagiert zum falschen Zeitpunkt und unpassend auf Störungen (Allgegenwärtigkeit)	1	2	3	4
Die Klassengemeinschaft ist stabil (Ablenker)				
Für einen reibungslosen Ablauf fehlen Regeln und Routinen (Regeln und Routinen)	1	2	3	4
Der Unterricht ist sinnvoll strukturiert (Übergänge)	4	3	2	1
Die Lehrerin gibt Rückmeldungen zum Lernprozess (Feedback)	4	3	2	1
In der Unterrichtsgestaltung fehlt Methodenvielfalt (Ablenker)				
Den Schülerinnen und Schülern sind einzelne Teilschritte unklar (Klarheit, gelingend)	4	3	2	1

D2: Lösung Vignette 2

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
<p>Sachunterricht Grundschule, Thema Magnetismus. Die Kinder sitzen an Gruppentischen. Kati kritzelt an ihrem Platz auf einem Zettel rum, während der Lehrer an der Tafel den Ablauf für die Gruppenarbeit erklärt.</p> <p>Lehrer: „Kati, jetzt dreh dich doch bitte mal zur Tafel und schau hin. (erklärt weiter)</p>	<p>Raum/Sitzplatz</p> <p>Barriere</p>	<p>Aus der Bemerkung des Lehrers („dreh dich zur Tafel“) wird deutlich, dass Kati nicht frontal zur Tafel sitzt, sondern sich umdrehen muss, um etwas zu sehen.</p>
<p>Also, ihr stellt erst mal Hypothesen auf und dann überprüft ihr die im Experiment und dokumentiert die Ergebnisse und macht euch vielleicht auch schon Notizen dazu, wenn euch was Bestimmtes auffällt, zum Beispiel, ähm, also einfach was euch auffällt.</p>	<p>Klarheit</p> <p>Barriere</p>	<p>Die Instruktion umfasst mehrere Fremdwörter und mehrere Anweisungen. Sie erfolgt nur mündlich und es fehlen Beispiele und Veranschaulichungen.</p>
<p>Die Aufteilung für die Gruppen kennt ihr, wer schon fertig ist, bleibt am Gruppentisch und kann wie immer Zeichnungen zu den Notizen machen. Los geht's!“</p>	<p>Regeln und Routinen</p> <p>gelingend</p>	<p>Es gibt Regeln für die Gruppeneinteilung und eine Routine, wie sich schnellere Gruppen beschäftigen können.</p>
<p>Die SuS beginnen eifrig, sich in den Gruppen zu beschäftigen. Der Lehrer wendet sich einer Gruppe zu, in der es heftige Diskussionen gibt, und bemüht sich intensiv um Klärung. Mit der Zeit wird es immer unruhiger in der Klasse. Lehrer: „Jetzt seid doch bitte mal noch einen Moment leise, damit die letzte Gruppe auch die Chance hat, das fertig zu machen! Seid ihr überhaupt schon alle fertig?“</p>	<p>Allgegenwärtigkeit</p> <p>Barriere</p>	<p>Der Lehrer hat den Rest der Klasse nicht im Blick, während er sich um eine Gruppe kümmert, und reagiert nicht auf die entstehende Unruhe.</p>
<p>Die Gruppenarbeit wird nach einiger Zeit durch einen Gong beendet und die SuS kommen im Kreis zusammen, um ausführlich die Ergebnisse zu besprechen. Eine Gruppe berichtet, dass Geldmünzen magnetisch</p>	<p>Übergänge</p> <p>gelingend</p> <p>(auch Hinweis zu gelingenden Regeln und Routinen)</p>	<p>Der Lehrer markiert einen Übergang, der zügig gelingt. Es bleibt genug Zeit für eine ausführliche Ergebnissammlung.</p>

sind (1 Cent), eine andere Gruppe hat das Gegenteil festgestellt (50 Cent). Die SuS sollen ihre Vermutungen zu den Gründen äußern.

Önder: „Vielleicht hat die andere Gruppe den Magneten falsch gehalten?“ Lehrer: „Was meinst du, Linn?“ Linn: „Das ist doch beides Metall. Vielleicht haben die den Magneten auch zu weit weg gehalten!“ Lehrer: „Der nächste, Louis.“ Louis: „Aber 50 Cent sind viel schwerer!“ Lehrer: „Kay, was meinst du?“ (...)

Feedback
Barriere

Der Lehrer geht nicht auf die Äußerungen der Kinder ein und gibt kein Feedback.

	Trifft sehr un- wahr-schein- lich zu	Trifft eher un- wahr-schein- lich zu	Trifft eher wahr-schein- lich zu	Trifft sehr wahr-schein- lich zu
In der Klasse gibt es Aufmerksamkeitsprobleme (Ablenker)				
Die Rückmeldungen des Lehrers beziehen sich zu wenig auf den Lernfortschritt (Feedback)	1	2	3	4
Einzelne Unterrichtsphasen sind gut aufeinander abgestimmt (Übergänge, gelingend)	1	2	3	4
Der Lehrer sorgt unzureichend dafür, dass alle gut an ihrem Platz arbeiten können (Raum)	1	2	3	4
Die Aufgabenstellung ist sinnvoll erklärt (Ablenker)				
Für die anschauliche Unterrichtsgestaltung fehlen technische Hilfsmittel (Ablenker)				
Der Lehrer versäumt, Unterrichtsprozesse aufmerksam zu steuern (Allgegenwärtigkeit)	1	2	3	4
Der Lehrer strukturiert Inhalte	4	3	2	1
Der Unterrichtsfluss gelingt durch eingeübte Rituale	1	2	3	4

D3: Lösung Vignette 3

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
<p>Kunstunterricht 5. Klasse Stadtteilschule, die Klasse ist zu Unterrichtsbeginn noch sehr unruhig. Lehrerin: „Wir machen jetzt zum Locker werden mal 5 Minuten Bodypercussion!“ Sie fängt an, die SuS zögern noch. Lehrerin: „Los, alle mitmachen! Einfach machen, was ich mache! (...) Und jetzt im Loop!“ Ein paar SuS steigen wieder aus. Lehrerin: „Na gut, hab ich vielleicht nicht so gut vorgemacht.</p>	<p>Regel und Routinen</p> <p>Barriere</p> <p>(auch Hinweis auf mangelnde Klarheit)</p>	<p>Das Ritual ist nicht eingeübt, nicht alle Kinder machen mit. Mit dem Wort „Loop“ scheint sich die Lehrerin für die Kinder nicht klar auszudrücken. Der Stundeneinstieg gelingt dadurch weniger gut.</p>
<p>Aber so ist es ja jetzt toll ruhig hier, so dass wir anfangen können.</p>	<p>Feedback</p> <p>gelingend</p>	<p>Die Lehrerin gibt positives Feedback.</p>
<p>Wir haben letzte Woche mit dem Thema „Zentangle“ begonnen und Konturen gezeichnet, heute geht es um die Muster. Ich zeig euch ein paar Beispiele am Smartboard, da seht ihr das am besten.</p>	<p>Klarheit</p> <p>Barriere</p>	<p>Die Lehrerin sagt nur, dass es um Muster geht, ohne eine klare Instruktion zu geben oder konkrete Arbeitsschritte zu benennen.</p>
<p>Antoni, setz dich heut ruhig hier vorne hin.</p>	<p>Raum</p> <p>gelingend</p>	<p>Die Lehrerin hat die Sitzordnung im Blick.</p>
<p>(...) Fangt schon mal an, ich geh noch mal ins Lehrerzimmer, da hab ich noch Schritt-für-Schritt-Anleitungen für ein paar Muster.“</p> <p>Die Lehrerin verlässt den Klassenraum, die meisten Kinder arbeiten, aber es ist wieder unruhig. Nach einer Weile kommt die Lehrerin zurück.</p>	<p>Übergänge</p> <p>Barriere</p> <p>(auch Hinweis auf mangelnde Allgegenwärtigkeit)</p>	<p>Die Lehrerin hat nicht alle Materialien zur Hand und verlässt deshalb den Klassenraum, für einige Kinder gelingt so nicht der Übergang in die selbstständige Arbeit. Die Lehrerin bekommt das nicht mit.</p>
<p>Antoni hat noch nicht mit Zeichnen begonnen. Die Lehrerin beginnt, ihm noch mal den nächsten Arbeitsschritt zu erklären. Sie unterbricht: „Moment, Antoni. (Zur Klasse:) Ihr habt euch eben so toll beruhigt, bitte versucht mal, euch weiter auf eure</p>	<p>Allgegenwärtigkeit</p> <p>Barriere</p> <p>(auch Hinweise auf mangelnde Klarheit und gelingendes Feedback)</p>	<p>Die Lehrerin bekommt erst jetzt mit, dass Antoni nicht arbeitet und nochmals eine Erklärung braucht. Sie beschäftigt sich mit Antoni und hat die restliche Klasse nicht im Blick. Sie reagiert nicht zu einem passenden Moment auf die Unruhe, son-</p>

Arbeit zu konzentrieren. Zent-angle ist ja meditatives Zeichnen, da solltet ihr euch also drauf konzentrieren und das geht am besten, wenn ihr dabei nicht redet.“ (erklärt Antoni weiter)

dern unterbricht dafür die Erklärung. Inhaltlich ist ihr Feedback aber konstruktiv und bezieht sich auf den Lerngegenstand.

Die Kinder zeichnen eine Weile. Alina und Carolin sind schnell fertig und fragen: „Dürfen wir jetzt Stempelbilder machen?“ Lehrerin: „Oh, ihr wart ja schnell! Ja, dann könnt ihr stempeln oder was ihr wollt.“ Alina und Carolin laufen los und wollen Stempel holen. Lehrerin: „Alina, Carolin, bitte räumt die Zeichenstifte weg.“ Alina und Carolin gehen zurück und räumen die Stifte weg und wollen dann stempeln. Lehrerin: „Bitte auch die Bilder in die Mappen tun!“ Alina und Carolin laufen wieder zurück und räumen die Bilder weg und laufen dann wieder zum Regal mit den Stempeln. Antoni: „Oh Mann, habt ihr's bald?“

Regeln und Routinen
Barriere
(auch Hinweis auf mangelnde Allgegenwärtigkeit)

Die Lehrerin hat nicht im Blick wie weit die Kinder sind. Es gibt keine eindeutige Regelung dafür, was die Kinder tun, die schon fertig sind. Es ist keine eingeübte Routine erkennbar, dass im Anschluss an die Arbeit Materialien und Bilder weggeräumt werden.

	Trifft sehr unwahr- scheinlich zu	Trifft eher unwahr- scheinlich zu	Trifft eher wahr-schein- lich zu	Trifft sehr wahr-schein- lich zu
Die Lehrerin macht deutlich, was die Schülerinnen und Schüler tun sollen (Klarheit)	4	3	2	1
Die Binnendifferenzierung wird durch einen Mangel an Materialien erschwert (Ablenker)				
Die Lehrerin gibt unpassende Rückmeldungen (Feedback gelingend)	4	3	2	1
Die Schülerinnen und Schüler sind geübt im Wechsel von Unterrichtsaktivitäten (Regeln und Routinen)	4	3	2	1
Die Lehrerin bereitet das Thema zu uninteressant auf (Ablenker)				
Der Unterrichtsverlauf ist stockend (Übergänge)	1	2	3	4
Der Umgangston der Klasse ist freundschaftlich (Ablenker)				
Die Lehrerin lässt räumliche Gegebenheiten unbeachtet (Raum gelingend)	4	3	2	1
Die Lehrerin hat das gesamte Unterrichtsgeschehen im Griff (Allgegenwärtigkeit)	4	3	2	1

D4: Lösung Vignette 4

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
5. Stunde Mathe. Der Lehrer kommt rein und begrüßt die Kinder. Lehrer: „Heute ist es so schön ruhig hier! Ach, der Paul ist nicht da!“	Feedback Barriere	Der Lehrer kommentiert einen abwesenden Schüler negativ und schafft so keine positive Präsenz.
Gelächter, der Lehrer sorgt schnell für Ruhe und kündigt eine Aufgabe an.	Allgegenwärtigkeit gelingend	Der Lehrer unterbindet zügig die Unruhe und verliert so keine Zeit.
Die Kinder sitzen an Gruppentischen und sollen sich gegenseitig erklären, welche „schlaueren Rechenwege“ sie nutzen, um Aufgaben mit Zehnerübergang zu rechnen. Lehrer: „Ihr habt jetzt genau 15 Minuten Zeit. Anna, du machst das mit mir, kannst da sitzen bleiben. (Geht zu Anna)	Raum Barriere (auch Hinweis auf mangelhaften Übergang)	Anna sitzt am Gruppentisch, soll an der Gruppenarbeit aber nicht teilnehmen, das ist als Arbeitssituation ungünstig. Der Lehrer macht eine genaue Zeitvorgabe, ohne dass deutlich wird, wie die 15 Minuten strukturiert werden sollen oder wer die Zeit im Blick hat.
Erklär mal wie du vorgehst. Anna: „Ich weiß ja eben nicht, wie ich das machen soll.“ Lehrer: „Dann versuch mal so: 30 plus 10 ist 40, dann denkst du an die Zahlengeschwister: 8 plus 2 ist 10, macht zusammen 50, und dann bleiben von der 7 noch 5 übrig, macht zusammen 55. (schreibt auf ein Blatt): 38+17= 30+10=40 8+2=10 40+10=50 50+5=55 Anna: „Das ist total unfair! Dann muss ich ja 5 Aufgaben rechnen!“	Klarheit Barriere (auch Hinweis auf unkonstruktives Feedback)	Der Lehrer geht nicht auf Annas Verständnisprobleme ein, sondern gibt eine Lösung vor. Diese erklärt er unzureichend, so dass Anna nicht folgen kann
Am Tisch entsteht Unruhe, die Kinder kichern. Der Lehrer macht den lachenden Kindern ein Zeichen, still weiterzuarbeiten und erklärt gleichzeitig weiter. Anschließend geht er zu den anderen Gruppen.	Allgegenwärtigkeit gelingend	Der Lehrer hat die ganze Klasse im Blick und reagiert niedrigschwellig auf die Störung, ohne den Unterrichtsfluss zu unterbrechen.

Nach 15 Minuten stellt sich der Lehrer vor die Klasse und wartet. Die Kinder reden noch. Lehrer: „Die 15 Minuten sind um, ich warte! (...) Ich möchte jetzt, dass jede Gruppe mal einen Rechenweg vorstellt.	Übergänge Barriere (auch Hinweis auf fehlende Regeln und Routinen)	Der Lehrer markiert nicht das Ende der Gruppenarbeit und schafft keinen Übergang zur Ergebnissammlung.
Wer fängt an?“ Die Kinder am vordersten Tisch schweigen zunächst und schubsen sich dann gegenseitig an. Lehrer: „So, wer trägt vor?“ Jan: „Okay, dann stell ich meinen Weg vor.“ Lucas murmelt leise: „Der will immer bestimmen.“ (...)	Regeln und Routinen Barriere	Es sind keine Regeln oder Routinen erkennbar, wie und von wem Ergebnisse von Gruppenarbeiten vorgestellt werden.

	Trifft sehr unwahrscheinlich zu	Trifft eher unwahrscheinlich zu	Trifft eher wahrscheinlich zu	Trifft sehr wahrscheinlich zu
Die räumlichen Arbeitsbedingungen sind für alle Schülerinnen und Schüler gut (Raum)	4	3	2	1
Es kommt zu unnötigen Verzögerungen im Unterrichtsverlauf (Übergänge)	1	2	3	4
Den Schülerinnen und Schülern ist das Ziel der Stunde unklar (Klarheit)	1	2	3	4
Den Schülerinnen und Schülern mangelt es an Lernmotivation (Ablenker)				
Der Lehrer kann Unterrichtsgeschehen durch passendes Feedback steuern (Feedback)	4	3	2	1
Der Lehrer handelt ergebnisorientiert (Ablenker)				
Den Schülerinnen und Schülern fehlen Routinen, was sie wann und wie zu tun haben (Regeln und Routinen)	1	2	3	4
Der Lehrer kann verschiedene Unterrichtsprozesse gleichzeitig steuern (Allgegenwärtigkeit gelingend)	1	2	3	4
Die Schülerinnen und Schüler haben genügend Materialien zur Verfügung (Ablenker)				

Tabelle D5

Häufigkeiten für die Beantwortungen der Items zu Barrieren und der Items zu gelingenden Bedingungen in der dritten Datenerhebung

Items und Vignetten	N	Trifft sehr un- Trifft eher un- Trifft eher wahr-schein- wahr-schein- wahrscheinlich Trifft sehr wahr- lich zu lich zu zu scheinlich zu			
V1 Raum Barriere	219	40	118	51	10
V1 Allgegenwärtigkeit Barriere	220	15	39	86	80
V1 Regeln und Routinen Barriere	221	7	27	110	77
V1 Übergänge Barriere	218	40	122	54	2
V1 Feedback Barriere	220	82	103	28	7
V1 Klarheit gelingend	222	8	32	120	62
V2 Feedback Barriere	218	14	52	116	36
V2 Übergänge gelingend	222	6	81	107	28
V2 Raum Barriere	221	28	90	84	19
V2 Allgegenwärtigkeit Barriere	222	11	65	105	41
V2 Klarheit Barriere	219	23	98	85	13
V2 Regeln und Routinen gelingend	222	29	73	78	42
V3 Klarheit Barriere	221	24	58	86	53
V3 Feedback gelingend	221	33	87	79	22
V3 Regeln und Routinen Barriere	221	29	80	86	26
V3 Übergänge Barriere	217	19	77	80	41
V3 Raum gelingend	218	25	73	90	30
V3 Allgegenwärtigkeit Barriere	219	44	73	83	19
V4 Raum Barriere	217	13	93	99	12
V4 Struktur Barriere	220	20	99	87	14
V4 Klarheit Barriere	220	27	83	90	20
V4 Feedback Barriere	218	38	106	63	11
V4 Regeln und Routinen Barriere	219	10	90	104	15
V4 Allgegenwärtigkeit	218	25	78	89	26

Anmerkung. Angaben in absoluten Zahlen. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4

Tabelle D6*Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1*

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weglassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weglassen
Raum Barriere	14.26	3.279	.144	.115
Allgegenwärtigkeit Barriere	14.09	3.016	.125	.124
Regeln und Routinen Barriere	13.99	3.200	.175	.087
Übergänge Barriere	14.22	3.050	.301	-.010
Feedback Barriere	13.93	3.381	.104	.150
Klarheit gelingend	15.17	4.533	-.271	.431

Tabelle D7*Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2*

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weglassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weglassen
Feedback Barriere	13.15	2.644	.332	-.279
Übergänge gelingend	13.22	4.428	-.265	.295
Raum Barriere	13.52	2.964	.175	-.096
Allgegenwärtigkeit Barriere	13.14	2.850	.248	-.173
Klarheit Barriere	13.34	3.343	.082	.016
Regeln und Routinen gelingend	13.35	4.190	-.243	.358

Tabelle D8*Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3*

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
Klarheit Barriere	12.85	2.757	.316	-.502
Feedback gelingend	12.50	5.951	-.487	.405
Regeln und Routinen Barriere	12.58	3.070	.264	-.376
Übergänge Barriere	12.43	3.148	.226	-.328
Raum gelingend	12.67	5.508	-.396	.340
Allgegenwärtigkeit Barriere	12.43	2.923	.294	-.437

Tabelle D9*Trennschärfen der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4*

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Cronbachs Alpha, wenn Item weggelassen
Raum Barriere	12.78	3.867	.212	.251
Struktur Barriere	12.84	3.311	.399	.107
Klarheit Barriere	12.80	2.997	.431	.048
Feedback Barriere	12.49	3.360	.323	.155
Regeln und Routinen Barriere	12.71	3.336	.441	.091
Allgegenwärtigkeit gelingend	12.75	6.463	-.520	.698

Tabelle D10

Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der dritten Datenerhebung (Studierende der Universität Hamburg)

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen ^a
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	
1	3.940	21.887	21.887	3.940	21.887	21.887	3.259
2	2.443	13.570	35.457	2.443	13.570	35.457	3.037
3	2.199	12.216	47.673	2.199	12.216	47.673	2.928
4	1.152	6.399	54.072				
5	.997	5.536	59.608				
6	.923	5.126	64.735				
7	.806	4.478	69.213				
8	.769	4.273	73.486				
9	.702	3.902	77.388				
10	.607	3.375	80.763				
11	.587	3.259	84.022				
12	.519	2.885	86.907				
13	.468	2.600	89.507				
14	.440	2.445	91.953				
15	.407	2.263	94.216				
16	.370	2.058	96.274				
17	.344	1.912	98.186				
18	.326	1.814	100.000				

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. a. Wenn Komponenten korreliert sind, können die Summen der quadrierten Ladungen nicht addiert werden, um eine Gesamtvarianz zu erhalten.

Tabelle D11

Komponentenkorrelationsmatrix zur Faktorenanalyse der dritten Datenerhebung

Komponente	1	2	3
1	1.000	.242	.159
2	.242	1.000	.188
3	.159	.188	1.000

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung.

Tabelle D12

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studienganggruppen im 1. und 2. Bachelorsemester anhand gewichteter Mittelwerte je Vignette (Studierende der Universität Hamburg)

	V2 gesamt ge- wichtet	V3 gesamt ge- wichtet	V4 gesamt ge- wichtet	V2 Barrieren gewichtet	V2 gelingend gewichtet	V3 Barrieren gewichtet	V3 gelingend gewichtet	V4 Barrieren gewichtet	V4 gelingend gewichtet
Mann-Whitney-U-Test	812.000	815.000	895.000	879.000	791.500	849.000	875.000	858.000	735.000
Wilcoxon-W	1163.000	3230.000	3310.000	3364.000	1142.500	3264.000	3290.000	3273.000	1060.000
Z	-.808	-.685	-.017	-.256	-.990	-.401	-.185	-.326	-1.060
Asymp. Sig. (2-seitig)	.419	.494	.987	.798	.322	.689	.853	.745	.289

Anmerkung. Gruppierungsvariable Studiengang. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4

Tabelle D13

Teststatistiken des Kruskal-Wallis-Test bei unabhängigen Stichproben zum Einfluss der Studiendauer auf die gewichteten Mittelwerte je Vignette (Studierende der Universität Hamburg im Studiengang Sonderpädagogik)

	V2 gesamt ge- wichtet	V3 gesamt ge- wichtet	V4 gesamt ge- wichtet	V2 Barrieren ge- wichtet	V2 gelingend gewichtet	V3 Barrieren ge- wichtet	V3 gelingend gewichtet	V4 Barrieren ge- wichtet	V4 gelingend gewichtet
Kruskal-Wallis-H	3.054	1.854	.591	.516	1.846	.648	1.519	.805	2.612
df	2	2	2	2	2	2	2	2	2
Asymp. Sig.	.217	.396	.744	.772	.397	.723	.468	.669	.271

Anmerkung. Gruppenvariable: Fachsemester. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4

Tabelle D14

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte je Vignette

	V2 gesamt ge- wichtet	V3 gesamt ge- wichtet	V4 gesamt ge- wichtet	V2 Barrieren gewichtet	V2 gelingend gewichtet	V3 Barrieren gewichtet	V3 gelingend gewichtet	V4 Barrieren gewichtet	V4 gelingend gewichtet
Mann-Whitney-U-Test	1551.500	1781.000	1486.500	1108.000	1080.500	1539.500	1030.500	1376.000	1390.000
Wilcoxon-W	22666.500	1952.000	22192.500	22223.000	1251.500	22245.500	1201.500	22082.000	1561.000
Z	-1.118	-.177	-1.310	-2.811	-2.922	-1.106	-3.085	-1.735	-1.699
Asymp. Sig. (2-seitig)	.263	.860	.190	.005	.003	.269	.002	.083	.089

Tabelle D15

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 2

Gesamtzahl	223
Mann-Whitney-U-Test	2582.000
Wilcoxon-W	2753.000
Teststatistik	2582.000
Standardfehler	262.178
Standardisierte Teststatistik	2.811
Asymptotische Sig. (zweiseitiger Test)	.005

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D16

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 2

Gesamtzahl	222
Mann-Whitney-U-Test	1080.500
Wilcoxon-W	1251.500
Teststatistik	1080.500
Standardfehler	258.551
Standardisierte Teststatistik	-2.922
Asymptotische Sig. (zweiseitiger Test)	.003

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D17

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 3

Gesamtzahl	221
Mann-Whitney-U-Test	1030.500
Wilcoxon-W	1201.500
Teststatistik	1030.500
Standardfehler	258.156
Standardisierte Teststatistik	-3.085
Asymptotische Sig. (zweiseitiger Test)	.002

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D18

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 4

Gesamtzahl	221
Mann-Whitney-U-Test	2278.000
Wilcoxon-W	2449.000
Teststatistik	2278.000
Standardfehler	259.887
Standardisierte Teststatistik	1.735
Asymptotische Sig. (zweiseitiger Test)	.083

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D19

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden der Universität Hamburg und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 4

Gesamtzahl	218
Mann-Whitney-U-Test	1390.000
Wilcoxon-W	1561.000
Teststatistik	1390.000
Standardfehler	241.273
Standardisierte Teststatistik	-1.699
Asymptotische Sig. (zweiseitiger Test)	.089

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D20

Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der dritten Datenerhebung (Studierende anderer Universitäten)

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen ^a
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	
1	4.301	23.896	23.896	4.301	23.896	23.896	3.658
2	2.391	13.284	37.180	2.391	13.284	37.180	2.928
3	2.012	11.178	48.358	2.012	11.178	48.358	3.096
4	1.048	5.823	54.181				
5	1.015	5.640	59.821				
6	.913	5.073	64.894				
7	.799	4.437	69.331				
8	.730	4.056	73.387				
9	.675	3.750	77.137				
10	.661	3.671	80.808				
11	.595	3.303	84.111				
12	.575	3.193	87.304				
13	.502	2.791	90.094				
14	.437	2.430	92.524				
15	.380	2.108	94.633				
16	.365	2.028	96.661				
17	.323	1.797	98.458				
18	.278	1.542	100.000				

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. a. Wenn Komponenten korreliert sind, können die Summen der quadrierten Ladungen nicht addiert werden, um eine Gesamtvarianz zu erhalten.

Tabelle D21

Komponentenkorrelationsmatrix zur Faktorenanalyse der dritten Datenerhebung (Studierende anderer Universitäten)

Komponente	1	2	3
1	1.000	.175	.333
2	.175	1.000	.174
3	.333	.174	1.000

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung.

Tabelle D22

Teststatistiken des Kruskal-Wallis-Tests bei unabhängigen Stichproben zum Einfluss der Studiendauer auf die gewichteten Mittelwerte je Vignette (Studierende der Allgemeinpädagogik, andere Universitäten)

	V2 gesamt ge- wichtet	V3 gesamt ge- wichtet	V4 gesamt ge- wichtet	V2 Barrieren gewichtet	V2 gelingend gewichtet	V3 Barrieren gewichtet	V3 gelingend gewichtet	V4 Barrieren gewichtet	V4 gelingend gewichtet
Kruskal-Wallis-H	2.877	.183	.157	2.653	.035	1.252	3.768	.073	.293
df	2	2	2	2	2	2	2	2	2
Asymp. Sig.	.237	.913	.924	.265	.983	.535	.152	.964	.864

Anmerkung. Gruppenvariable: Fachsemester. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4

Tabelle D23

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studienganggruppen im 4. bis 6. Bachelorsemester anhand gewichteter Mittelwerte je Vignette (Studierende anderer Universitäten)

	V2 gesamt ge- wichtet	V3 gesamt ge- wichtet	V4 gesamt ge- wichtet	V2 Barrieren gewichtet	V2 gelingend gewichtet	V3 Barrieren gewichtet	V3 gelingend gewichtet	V4 Barrieren gewichtet	V4 gelingend gewichtet
Mann-Whitney-U-Test	762.500	823.500	863.500	685.500	865.000	836.000	983.000	855.000	965.000
Wilcoxon-W	1258.500	1319.500	1359.500	1181.500	3010.000	1332.000	3128.000	1351.000	3110.000
Z	-1.920	-1.442	-1.129	-2.525	-1.127	-1.345	-.194	-1.195	-.363
Asymp. Sig. (2-seitig)	.055	.149	.259	.012	.260	.179	.846	.232	.717

Anmerkung. Gruppenvariable: Studiengänge. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4

Tabelle D24

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studiengangsgruppen im 4. bis 6. Bachelorsemester anhand gewichteter Mittelwerte für Vignette 2 (Studierende anderer Universitäten)

Gesamtzahl	96
Mann-Whitney-U-Test	762.500
Wilcoxon-W	1258.500
Teststatistik	762.500
Standardfehler	127.619
Standardisierte Teststatistik	-1.920
Asymptotische Sig. (zweiseitiger Test)	.055

Anmerkung. Gruppenvariable: Studiengänge

Tabelle D25

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Studiengangsgruppen im 4. bis 6. Bachelorsemester anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 2 (Studierende anderer Universitäten)

Gesamtzahl	96
Mann-Whitney-U-Test	685.500
Wilcoxon-W	1181.500
Teststatistik	685.500
Standardfehler	127.531
Standardisierte Teststatistik	-2.525
Asymptotische Sig. (zweiseitiger Test)	.012

Anmerkung. Gruppenvariable: Studiengänge

Tabelle D26

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte je Vignette

	V2 gesamt ge- wichtet	V3 gesamt ge- wichtet	V4 gesamt ge- wichtet	V2 Barrieren gewichtet	V2 gelingend gewichtet	V3 Barrieren gewichtet	V3 gelingend gewichtet	V4 Barrieren gewichtet	V4 gelingend gewichtet
Mann-Whitney-U-Test	1223.500	1122.000	1022.500	862.000	792.000	1207.500	773.000	950.000	1029.000
Wilcoxon-W	11808.500	1293.000	11607.500	11447.000	963.000	11647.500	944.000	11535.000	1200.000
Z	-.432	-.927	-1.496	-2.347	-2.741	-.472	-2.815	-1.880	-1.569
Asymp. Sig. (2-seitig)	.666	.354	.135	.019	.006	.637	.005	.060	.117

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D27

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 2

Gesamtzahl	163
Mann-Whitney-U-Test	1748.000
Wilcoxon-W	1919.000
Teststatistik	1748.000
Standardfehler	188.719
Standardisierte Teststatistik	2.347
Asymptotische Sig. (zweiseitiger Test)	.019

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D28

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 2

Gesamtzahl	163
Mann-Whitney-U-Test	792.000
Wilcoxon-W	963.000
Teststatistik	792.000
Standardfehler	187.129
Standardisierte Teststatistik	-2.741
Asymptotische Sig. (zweiseitiger Test)	.006

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D29

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu gelingenden Situationen in Vignette 3

Gesamtzahl	162
Mann-Whitney-U-Test	773.000
Wilcoxon-W	944.000
Teststatistik	773.000
Standardfehler	185.769
Standardisierte Teststatistik	-2.815
Asymptotische Sig. (zweiseitiger Test)	.005

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Tabelle D30

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich von Studierenden anderer Universitäten und Lehrkräften anhand gewichteter Mittelwerte der Items zu Barrieren in Vignette 4

Gesamtzahl	163
Mann-Whitney-U-Test	1660.000
Wilcoxon-W	1831.000
Teststatistik	1660.000
Standardfehler	188.807
Standardisierte Teststatistik	1.880
Asymptotische Sig. (zweiseitiger Test)	.060

Anmerkung. Gruppenvariable: Expertise (Studierende/Lehrkräfte)

Anhang E: Vierter Entwicklungsschritt

E1: Lösung und Werte Vignette 1

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
<p>Unterrichtsbeginn Deutsch, 4. Klasse Grundschule, Thema Satzglieder. Die Kinder sitzen an Partnertischen. Der Lehrerin fällt auf, dass Murat noch Kaugummi kaut. Lehrerin: „Murat, spuck bitte dein Kaugummi aus!“ Murat steht auf, spuckt sein Kaugummi in den Müll, setzt sich wieder hin. Lehrerin: „Hanna, spuck dein Kaugummi aus!“ Hanna steht auf, spuckt ihr Kaugummi aus. In der Zwischenzeit fängt die Klasse an zu kichern – Murat packt ein neues Kaugummi aus. Lehrerin: „Murat, heb dir das bitte für die Pause auf. So, nun lasst uns endlich anfangen.“</p>	Regeln und Routinen Barriere	<p>Hinweis auf Partnertische (für Sitzordnung relevant).</p> <p>Es sind keine Regeln bzgl. des Kaugummikauens erkennbar, so dass Zeit verloren geht und Unruhe entsteht. Die Lehrerin fordert die Kinder zudem nacheinander auf, statt kurz und bündig darauf zu reagieren, was zusätzlich Zeit kostet.</p>
<p>Wir haben gestern mit den W-Fragen Satzglieder bestimmt. Das wollen wir jetzt noch mal so wie gestern wiederholen, damit Ihr die Wer-oder-Was-, die Wem-oder-Was- und die Wen-oder-Was-Fragen noch sicherer könnt. Ich schreibe drei Sätze an die Tafel und unterstreiche einige Wörter. Ihr schreibt die nicht ab. Anschließend sucht ihr euch einen Satz aus und überlegt allein für euch, mit welcher Frage ihr nach den unterstrichenen Wörtern fragen könnt (erklärt ein Beispiel). Also ich schreibe an und Ihr sucht dann einen Satz aus und findet eine Frage.“</p>	Klarheit gelingend	<p>Die Lehrerin benennt das Thema und macht deutlich, was sie tut und was die Kinder zu tun haben. Sie verdeutlicht durch ein Beispiel. Am Ende wiederholt sie noch mal.</p>
<p>Die Lehrerin schreibt an die Tafel. Als sich Luise zu Franzl umdreht, um ihr ihre neuen Sticker zu zeigen, gibt ihr die Lehrerin kurz ein Zeichen, aufmerksam zu bleiben. Anschließend fordert sie die Kinder auf: „Sucht einen Satz aus und notiert die Frage zu den unterstrichenen Wortarten. Ihr habt 5 Minuten Zeit.“</p>	Allgegenwärtigkeit Gelingend Erster Hinweis zum Übergang (Barriere)	<p>Die Lehrerin hat die Klasse im Blick und reagiert prompt und kurz und bündig auf die Störung.</p> <p>Erster Hinweis zum Übergang, strikte Zeitvorgabe.</p>
<p>Emma stöhnt: „Oh Mann, Paul macht sich immer so breit, ich hab gar keinen</p>	Raum/Sitzplatz Barriere	<p>Der Sitzplatz ist für Emma ungünstig, weil sie sich gestört</p>

Platz!“ Lehrerin: „Versucht jetzt bitte mal, still zu arbeiten.“	Übergänge Barriere	fühlt. Die Lehrkraft berücksichtigt das nicht.
Nach 5 Minuten beendet die Lehrerin die Stillarbeit und beginnt, Ergebnisse zu sammeln. Lehrerin: „Jetzt legen alle die Stifte hin und machen mit, sonst werden wir nicht fertig, ihr sollt das nachher ja auch noch geordnet aufschreiben.“		Die Lehrerin beginnt nach genau 5 Minuten mit dem Sammeln von Ergebnissen, obwohl einige Kinder noch schreiben. Ihre Ankündigung macht deutlich, dass das Zeit-Management nicht gelingt und zu wenig Zeit für die geplanten Schritte zur Verfügung steht.
Paul: „Das ist doof, ich kann das irgendwie nicht.“ Lehrerin: „Du konntest doch letztes Mal schon gut die Werfrage stellen, versuch doch das mal zuerst.“	Feedback gelingend	Die Lehrerin gibt ermutigendes und konstruktives Feedback, das sich auf den Lerngegenstand bezieht.

	Trifft gar nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft völlig zu
Es herrscht insgesamt eine positive Klassenatmosphäre	1	2	3	4
Übergänge zwischen Unterrichtseinheiten gelingen reibungslos	4	3	2	1
Die Lehrkraft hat die ganze Klasse im Blick	1	2	3	4
Räumliche Gegebenheiten werden berücksichtigt	4	3	2	1
Die Lehrkraft gibt konstruktives Feedback	1	2	3	4
Unterrichtsinhalte und -ziele werden deutlich	1	2	3	4
Der Unterrichtsverlauf lässt etablierte Regeln und Routinen erkennen	4	3	2	1
Die beschriebene Situation ist authentisch	1	2	3	4
Ich kann mir die beschriebene Situation gut vorstellen	1	2	3	4

E2: Lösung und Werte Vignette 2

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
<p>Sachunterricht, 4. Klasse Grundschule, Thema Magnetismus. Die Kinder sitzen an Gruppentischen. Katri kritzelt an ihrem Platz auf einem Zettel rum, während der Lehrer an der Tafel den Ablauf für die Gruppenarbeit erklärt.</p> <p>Lehrer: „Kati, jetzt dreh dich doch bitte mal zur Tafel um, sonst kriegst du wieder nix mit. (erklärt weiter)</p>	<p>Raum/Sitzplatz</p> <p>Barriere</p>	<p>Aus der Bemerkung des Lehrers („dreh dich doch bitte mal zur Tafel um“) wird deutlich, dass Kati nicht frontal zur Tafel sitzt, sondern sich umdrehen muss, um etwas zu sehen.</p>
<p>Also, ihr stellt erst mal Hypothesen auf und dann überprüft ihr die im Experiment und dokumentiert die Ergebnisse und macht euch vielleicht auch schon Notizen dazu, wenn euch was Bestimmtes auffällt, zum Beispiel, ähm, also einfach was euch auffällt.</p>	<p>Klarheit</p> <p>Barriere</p>	<p>Die Instruktion umfasst mehrere Fremdwörter und mehrere Anweisungen. Sie erfolgt nur mündlich und es fehlen Beispiele und Veranschaulichungen.</p>
<p>Die Aufteilung für die Gruppen kennt ihr, wer schon fertig ist, bleibt am Gruppentisch und kann wie immer Zeichnungen zu den Notizen machen. Los geht's!“</p>	<p>Regeln und Routinen</p> <p>gelingend</p>	<p>Es gibt Regeln für die Gruppeneinteilung und eine Routine, wie sich schnellere Gruppen beschäftigen können.</p>
<p>Die SuS beginnen sofort eifrig, sich in den Gruppen zu beschäftigen. Der Lehrer wendet sich einer Gruppe zu, in der es Diskussionen gibt, und redet längere Zeit mit der Gruppe. Mit der Zeit wird es immer unruhiger in der Klasse. Lehrer: „Jetzt seid doch bitte mal noch einen Moment leise, damit die letzte Gruppe auch die Chance hat, das fertig zu machen! Seid ihr überhaupt schon alle fertig?“</p>	<p>Allgegenwärtigkeit</p> <p>Barriere</p>	<p>Der Lehrer hat den Rest der Klasse nicht im Blick, während er sich um eine Gruppe kümmert, und reagiert nicht auf die entstehende Unruhe.</p>
<p>Die Gruppenarbeit wird nach einiger Zeit durch einen Gong beendet und die SuS kommen zügig im Kreis zusammen, um ausführlich die Ergebnisse zu</p>	<p>Übergänge</p> <p>gelingend</p> <p>(auch Hinweis zu gelingenden Regeln und Routinen)</p>	<p>Der Lehrer markiert einen Übergang, der zügig gelingt. Es bleibt genug Zeit für eine ausführliche Ergebnissammlung.</p>

besprechen. Eine Gruppe berichtet, dass Geldmünzen magnetisch sind (1 Cent), eine andere Gruppe hat das Gegenteil festgestellt (50 Cent). Die SuS sollen ihre Vermutungen zu den Gründen äußern.

Önder: „Vielleicht hat die andere Gruppe den Magneten falsch gehalten?“ Lehrer: „Was meinst du, Linn?“ Linn: „Das ist doch beides Metall. Vielleicht haben die den Magneten auch zu weit weg gehalten!“ Lehrer: „Der nächste, Louis.“ Louis: „Aber 50 Cent sind viel schwerer!“ Lehrer: „Kay, was meinst du?“ Kay: „Die sind doch aus unterschiedlichen Metallen, vielleicht hat das damit zu tun?“ Lehrer: „Ja, das ist richtig.“ (...)

Feedback
Barriere

Der Lehrer geht zunächst nicht auf die Äußerungen der Kinder ein und gibt kein Feedback. Die Letzte Äußerung wertet er als richtig, was die vorhergehenden Äußerungen als falsch markiert, wobei darauf nicht konstruktiv eingegangen wird.

	Trifft gar nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft völlig zu
Es herrscht insgesamt eine positive Klassenatmosphäre	1	2	3	4
Übergänge zwischen Unterrichtseinheiten gelingen reibungslos	1	2	3	4
Die Lehrkraft hat die ganze Klasse im Blick	4	3	2	1
Räumliche Gegebenheiten werden berücksichtigt	4	3	2	1
Die Lehrkraft gibt konstruktives Feedback	4	3	2	1
Unterrichtsinhalte und -ziele werden deutlich	4	3	2	1
Der Unterrichtsverlauf lässt etablierte Regeln und Routinen erkennen	1	2	3	4
Die beschriebene Situation ist authentisch	1	2	3	4
Ich kann mir die beschriebene Situation gut vorstellen	1	2	3	4

E3: Lösung und Werte Vignette 3

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
Kunstunterricht, 5. Klasse Stadtteilschule, Thema meditatives Zeichnen. Lehrerin: „Wir haben ja letzte Woche mit dem Thema „Zentangle“, dem meditativen Zeichnen, begonnen und Konturen gezeichnet, heute geht es wie angekündigt um die Muster. Ich zeig euch ein paar Beispiele am Smartboard, da seht ihr das am besten.“	Klarheit gelingend	Die Lehrerin benennt das Thema, auf das sie vorher schon verwiesen hat („wie angekündigt“).
Jonas, setz dich heut ruhig kurz hier vorne hin.	Raum gelingend	Die Lehrerin hat die Sitzordnung im Blick und bietet einem Kind für die Aufgabe einen anderen Sitzplatz an.
(Zeigt einige Muster am Smartboard). Wer fertig ist, darf frei arbeiten und nächste Woche beginnen wir mit der Ausstellung der Bilder.	Klarheit gelingend (Fortsetzung)	Die Lehrerin veranschaulicht die Aufgabe am Smartboard und nennt das Ziel, wann die Bilder fertig sein sollen und was damit passiert.
Fangt schon mal an, ich geh noch mal ins Lehrerzimmer, da hab ich noch Schritt-für-Schritt-Anleitungen für ein paar Muster.“ Die Lehrerin verlässt den Klassenraum, die meisten Kinder arbeiten, aber es ist wieder unruhig. Nach einer Weile kommt die Lehrerin zurück.	Übergänge Barriere (auch Hinweis auf mangelnde Allgegenwärtigkeit)	Die Lehrerin hat nicht alle Materialien zur Hand und verlässt deshalb den Klassenraum, für einige Kinder gelingt so nicht der Übergang in die selbstständige Arbeit. Die Lehrerin bekommt das nicht mit.
Alina: „Ich bin schon ganz weit, Antoni hat noch nicht mal angefangen!“ Die Lehrerin beginnt daraufhin, Antoni noch mal den nächsten Arbeitsschritt zu erklären. Sie unterbricht: „Moment, Antoni.“	Allgegenwärtigkeit Barriere	Die Lehrerin bekommt erst jetzt mit, dass Antoni nicht arbeitet und nochmals eine Erklärung braucht. Sie beschäftigt sich mit Antoni und hat die restliche Klasse nicht im Blick. Sie reagiert auf die Unruhe mit einer Unterbrechung der Erklärung.
(Zur Klasse:) Ihr habt euch eben so toll beruhigt, jetzt ist es schon wieder zu laut hier, bitte versucht mal, euch weiter auf eure Arbeit zu konzentrieren. Zentangle ist ja meditatives	Feedback gelingend	Inhaltlich ist das Feedback der Lehrerin aber konstruktiv und bezieht sich auf den Lerngegenstand.

Zeichnen, da solltet ihr euch also drauf konzentrieren und das geht am besten, wenn ihr dabei nicht redet.“ (erklärt Antoni weiter)

Die Kinder zeichnen eine Weile. Alina und Carolin sind schnell fertig und fragen: „Dürfen wir jetzt Stempelbilder machen?“ Lehrerin: „Oh, ihr wart ja schnell! Ja, dann könnt ihr stempeln oder was ihr wollt.“ Alina und Carolin laufen los und wollen Stempel holen. Lehrerin: „Alina, Carolin, bitte räumt die Zeichenstifte weg.“ Alina und Carolin gehen zurück und räumen die Stifte weg und wollen dann stempeln. Lehrerin: „Bitte auch die Bilder in die Mappen tun!“ Alina und Carolin laufen wieder zurück und räumen die Bilder weg und laufen dann wieder zum Regal mit den Stempeln. Antoni: „Oh Mann, habt ihr’s bald?“

Regeln und Routinen
Barriere
(auch Hinweis auf mangelnde Allgegenwärtigkeit)

Die Lehrerin hat nicht im Blick wie weit die Kinder sind. Es gibt keine eindeutige Regelung für die Freiarbeit und es ist keine eingeübte Routine erkennbar, dass im Anschluss an die Arbeit Materialien und Bilder weggeräumt werden.

	Trifft gar nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft völlig zu
Es herrscht insgesamt eine positive Klassenatmosphäre	1	2	3	4
Übergänge zwischen Unterrichtseinheiten gelingen reibungslos	4	3	2	1
Die Lehrkraft hat die ganze Klasse im Blick	4	3	2	1
Räumliche Gegebenheiten werden berücksichtigt	1	2	3	4
Die Lehrkraft gibt konstruktives Feedback	1	2	3	4
Unterrichtsinhalte und -ziele werden deutlich	1	2	3	4
Der Unterrichtsverlauf lässt etablierte Regeln und Routinen erkennen	4	3	2	1
Die beschriebene Situation ist authentisch	1	2	3	4
Ich kann mir die beschriebene Situation gut vorstellen	1	2	3	4

E4: Lösung und Werte Vignette 4

Vignettenabschnitt	Zuordnung und Ausprägung	Kommentar
Mathematik, 2. Klasse Grundschule, Thema Addieren bis 100 mit Zehnerübergang. Der Lehrer kommt rein und begrüßt die Kinder. Es herrscht noch Gekicher. Der Lehrer sorgt schnell für Ruhe und kündigt eine Aufgabe an:	Allgegenwärtigkeit Gelingend (auch Hinweis auf gelingenden Übergang)	Der Lehrer sorgt unterbindet schnell die Unruhe.
„So, ihr sitzt ja schon wie besprochen in euren Arbeitsgruppen zusammen und sollt gleich auch in diesen Gruppen arbeiten. Lennet, wenn es eurer Gruppe wieder zu laut ist, dann könnt ihr in den Gruppenraum gehen, da liegt auch alles, was ihr braucht.“	Raum/Sitzordnung Gelingend (auch Hinweis auf gelingende Regeln und Routinen)	Die Kinder sitzen bereits an den richtigen Plätzen und der Lehrer hat das im Blick. Raumbedingungen berücksichtigt er und bietet alternative Arbeitsplätze an, an denen die Materialien zugänglich sind.
Ihr sollt heute mal überlegen, welche schlaun Rechenwege ihr bei schwierigen Aufgaben nutzen könnt, also zum Beispiel sowas wie $38+27$. Fangt mal an, ich geh dabei rum.“	Klarheit Barriere	Der Lehrer erklärt nicht, was er mit „schlau“ meint und erklärt auch ansonsten die Aufgabe nicht. Es fehlen Hinweis, wie das als Gruppenarbeit bearbeitet werden soll.
Die Kinder fangen zügig an und der Lehrer geht zur ersten Gruppe und gibt einige Anregungen, welches Material die Kinder nutzen könnten. Da eine andere Gruppe nicht zu arbeiten scheint, geht der Lehrer anschließend zu diesem Gruppentisch und stellt fest, dass jeder für sich rechnet.	Allgegenwärtigkeit gelingend (auch Hinweis auf gelingenden Übergang)	Der Lehrer hat die ganze Klasse im Blick und bemerkt, was vor sich geht.
Lehrer: „Ihr sollt Euch das aber auch gegenseitig vorstellen und erklären!“ Milla: „Wir wollten das jetzt erst aufschreiben und dann in der Gruppe besprechen. Oder sollen wir das nur mündlich machen?“ Lehrer: „Ach so, nee, so könnt ihr das auch machen.“	Klarheit Barriere (Fortsetzung)	Es zeigt sich, dass die Kinder nicht wissen, wie sie vorgehen sollen und dass die Arbeitsform unklar ist. Der Lehrer benennt weiterhin kein klares Ziel.
Nach einiger Zeit kündigt der Lehrer das Ende der Gruppenarbeit an. Die Kinder kommen tischweise zügig in den Stuhlkreis und setzen sich auf ihre Plätze.	Übergänge Gelingend Regeln und Routinen gelingend	Es gelingt ein reibungsloser Übergang. Daran, dass die Kinder tischweise in den Stuhlkreis kommen, sind etablierte Regeln und Routinen erkennbar.

Lehrer: „So, dann zeigt mal der Reihe nach, wie ihr gerechnet habt. Yasmin: „Ich hab das nicht aufgeschrieben.“ Lehrer: „Erzähl halt mal.“ Yasmin: „Wir haben 50 plus 40 gerechnet und erst die Null weggetan und dann wieder hingeschrieben.“ Lehrer: „Da hast du es dir aber leicht gemacht, das ist ja gar kein Zehnerübergang!“

Feedback
Barriere

Der Lehrer gibt kein konstruktives Feedback. Er geht nicht auf den Rechenweg ein, sondern formuliert einen Vorwurf.

	Trifft gar nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft völlig zu
Es herrscht insgesamt eine positive Klassenatmosphäre	1	2	3	4
Übergänge zwischen Unterrichtseinheiten gelingen reibungslos	1	2	3	4
Die Lehrkraft hat die ganze Klasse im Blick	1	2	3	4
Räumliche Gegebenheiten werden berücksichtigt	1	2	3	4
Die Lehrkraft gibt konstruktives Feedback	4	3	2	1
Unterrichtsinhalte und -ziele werden deutlich	4	3	2	1
Der Unterrichtsverlauf lässt etablierte Regeln und Routinen erkennen	1	2	3	4
Die beschriebene Situation ist authentisch	1	2	3	4
Ich kann mir die beschriebene Situation gut vorstellen	1	2	3	4

Tabelle E5

Paarweise Vergleiche des Friedman-Tests zur Einschätzung der Authentizität der Vignetten (Bachelorstudierende)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 3 Authentizität- Vignette 2 Authentizität	.041	.175	.236	.813	1.000
Vignette 3 Authentizität- Vignette 4 Authentizität	-.106	.175	-.603	.546	1.000
Vignette 3 Authentizität- Vignette 1 Authentizität	.459	.175	2.623	.009	.052
Vignette 2 Authentizität- Vignette 4 Authentizität	-.064	.175	-.367	.713	1.000
Vignette 2 Authentizität- Vignette 1 Authentizität	.417	.175	2.387	.017	.102
Vignette 4 Authentizität- Vignette 1 Authentizität	.353	.175	2.020	.043	.260

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E6

Häufigkeiten für die Beantwortungen der Items zu Barrieren und der Items zu gelingenden Bedingungen in der vierten Datenerhebung

Items	N	Trifft gar nicht zu	Trifft eher nicht zu	Trifft eher zu	Trifft völlig zu
V1 Übergänge (Barriere)	192	17	89	80	6
V1 Allgegenwärtigkeit (Gelingend)	193	13	34	113	33
V1 Raum (Barriere)	187	18	85	78	6
V1 Feedback (Gelingend)	192	23	74	79	161
V1 Klarheit (Gelingend)	193	13	52	87	41
V1 Regeln (Barriere)	193	26	77	75	15
V2 Übergänge (Gelingend)	193	8	83	92	10
V2 Allgegenwärtigkeit (Barriere)	191	42	93	50	6
V2 Raum (Barriere)	188	16	91	75	6
V2 Feedback (Barriere)	191	61	91	33	6
V2 Klarheit (Barriere)	190	20	85	76	9
V2 Regeln (Gelingend)	192	18	58	95	21
V3 Übergänge (Barriere)	192	30	110	48	4
V3 Allgegenwärtigkeit (Barriere)	191	39	89	54	9
V3 Raum (Gelingend)	189	18	92	73	6
V3 Feedback (Gelingend)	193	27	102	56	8
V3 Klarheit (Gelingend)	193	12	79	80	22
V3 Regeln (Barriere)	193	33	114	39	7
V4 Übergänge (Gelingend)	193	3	42	92	56
V4 Allgegenwärtigkeit (Gelingend)	193	9	65	88	31
V4 Raum (Gelingend)	189	3	19	112	55
V4 Feedback (Barriere)	190	29	104	48	9
V4 Klarheit (Barriere)	193	28	86	61	18
V4 Regeln (Gelingend)	190	3	52	103	32

Anmerkung. Angaben in absoluten Zahlen. V1 = Vignette 1, V2 = Vignette 2, V3 = Vignette 3, V4 = Vignette 4.

Tabelle E7*Reliabilitätsstatistik zur Gesamtskala der vierten Datenerhebung*

Cronbachs Alpha ^a	Anzahl der Items
-.104	24

Anmerkung. a. Der Wert ist negativ aufgrund einer negativen mittleren Kovarianz zwischen den Items.

Tabelle E8*Reliabilitätsstatistik zu Vignette 1 der vierten Datenerhebung*

Cronbachs Alpha ^a	Anzahl der Items
-.815	6

Anmerkung. a. Der Wert ist negativ aufgrund einer negativen mittleren Kovarianz zwischen den Items. Dies verstößt gegen die Annahmen über die Zuverlässigkeit des Modells. Sie sollten die Item-Kodierungen überprüfen.

Tabelle E9*Reliabilitätsstatistik zu Vignette 2 der vierten Datenerhebung*

Cronbachs Alpha ^a	Anzahl der Items
-.007	6

Anmerkung. a. Der Wert ist negativ aufgrund einer negativen mittleren Kovarianz zwischen den Items. Dies verstößt gegen die Annahmen über die Zuverlässigkeit des Modells. Sie sollten die Item-Kodierungen überprüfen.

Tabelle E10*Reliabilitätsstatistik zu Vignette 3 der vierten Datenerhebung*

Cronbachs Alpha ^a	Anzahl der Items
-.409	6

Anmerkung. a. Der Wert ist negativ aufgrund einer negativen mittleren Kovarianz zwischen den Items. Dies verstößt gegen die Annahmen über die Zuverlässigkeit des Modells. Sie sollten die Item-Kodierungen überprüfen.

Tabelle E11*Reliabilitätsstatistik zu Vignette 4 der vierten Datenerhebung*

Cronbachs Alpha	Anzahl der Items
.086	6

Tabelle E12*Erklärte Gesamtvarianz der Faktoren für die Faktorenanalyse der vierten Datenerhebung*

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte
	Gesamt	% der Varianz	Kumulierte	Gesamt	% der Varianz	Kumulierte %	Summe der quadrierten Ladungen ^a
			%				
1	4.285	17.853	17.853	4.285	17.853	17.853	3.460
2	2.879	11.997	29.851	2.879	11.997	29.851	3.164
3	2.489	10.373	40.223	2.489	10.373	40.223	2.857
4	2.068	8.616	48.839	2.068	8.616	48.839	3.103
5	1.390	5.791	54.630				
6	1.100	4.583	59.213				
7	1.008	4.198	63.411				
8	.881	3.669	67.080				
9	.817	3.405	70.486				
10	.753	3.138	73.623				
11	.685	2.854	76.477				
12	.655	2.729	79.206				
13	.616	2.567	81.772				
14	.584	2.433	84.205				
15	.506	2.107	86.312				
16	.464	1.935	88.246				
17	.457	1.903	90.150				
18	.439	1.831	91.981				
19	.401	1.672	93.652				
20	.382	1.591	95.243				
21	.318	1.326	96.569				
22	.299	1.244	97.813				
23	.275	1.148	98.961				
24	.249	1.039	100.000				

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. a. Wenn Komponenten korreliert sind, können die Summen der quadrierten Ladungen nicht addiert werden, um eine Gesamtvarianz zu erhalten.

Tabelle E13

Komponentenkorrelationsmatrix zur Faktorenanalyse der vierten Datenerhebung

Komponente	1	2	3	4
1	1.000	.179	.004	-.267
2	.179	1.000	-.100	-.142
3	.004	-.100	1.000	.070
4	-.267	-.142	.070	1.000

Anmerkung. Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Promax mit Kaiser-Normalisierung.

Tabelle E14

Teststatistiken des Kruskal-Wallis-Tests bei unabhängigen Stichproben zum Einfluss der Expertise auf die gewichteten Mittelwerte je Vignetten und auf die Items zu Barrieren und zu gelingenden Situationen

	Vignette 1 ge-wichtet	Vignette 2 ge-wichtet	Vignette 3 ge-wichtet	Vignette 4 ge-wichtet	Alle Barrieren gewichtet	Alle Gelingen-den gewichtet
Kruskal-Wallis-H	3.910	5.983	.911	6.299	17.766	6.094
df	2	2	2	2	2	2
Asymp. Sig.	.142	.050	.634	.043	<.001	.048

Anmerkung. Gruppenvariable: Expertise

Tabelle E15

Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für

Vignette 2

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
4. Bachelor-Lehrer	-17.721	14.210	-1.247	.212	.637
4. Bachelor-Master	-20.345	8.711	-2.335	.020	.059
Lehrer-Master	2.624	14.852	.177	.860	1.000

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst

Tabelle E16

Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für Vignette 4

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
4. Bachelor-Lehrer	-2.212	14.210	-.156	.876	1.000
4. Bachelor-Master	-21.543	8.711	-2.473	.013	.040
Lehrer-Master	19.331	14.852	1.302	.193	.579

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E17

Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für Items zu Barrieren

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
4. Bachelor-Master	-28.304	8.712	-3.249	.001	.003
4. Bachelor-Lehrer	-47.804	14.212	-3.364	<.001	.002
Master-Lehrer	-19.500	14.853	-1.313	.189	.568

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E18

Paarweise Vergleiche der Expertisegruppen anhand gewichteter Mittelwerte für Items zu gelingenden Situationen

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Lehrer-4. Bachelor	34.009	14.212	2.393	.017	.050
Lehrer-KP2	34.323	14.853	2.311	.021	.063
4. Bachelor-KP2	-.315	8.712	-.036	.971	1.000

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E19

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen der Klassenatmosphäre (Bachelorstudierende)

Gesamtzahl	109
Teststatistik	46.502
Freiheitsgrad	3
Asymptotische Sig. (zwei-seitiger Test)	<.001

Tabelle E20

Paarweise Vergleiche der Einschätzungen der Klassenatmosphäre (Bachelorstudierende)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 1 Klassenatmosphäre-Vignette 3 Klassenatmosphäre	-.050	.175	-.289	.773	1.000
Vignette 1 Klassenatmosphäre-Vignette 2 Klassenatmosphäre	-.197	.175	-1.128	.259	1.000
Vignette 1 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-.853	.175	-4.879	<.001	.000
Vignette 3 Klassenatmosphäre-Vignette 2 Klassenatmosphäre	.147	.175	.839	.401	1.000
Vignette 3 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-.803	.175	-4.590	<.001	.000
Vignette 2 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-.656	.175	-3.751	<.001	.001

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E21

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen der Klassenatmosphäre (Masterstudierende)

Gesamtzahl	63
Teststatistik	31.081
Freiheitsgrad	3
Asymptotische Sig. (zwei-seitiger Test)	<.001

Tabelle E22

Paarweise Vergleiche der Einschätzung der Klassenatmosphäre (Masterstudierende)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 1 Klassenatmosphäre-Vignette 3 Klassenatmosphäre	-.214	.230	-.932	.352	1.000
Vignette 1 Klassenatmosphäre-Vignette 2 Klassenatmosphäre	-.230	.230	-1.001	.317	1.000
Vignette 1 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-.952	.230	-4.140	<.001	.000
Vignette 3 Klassenatmosphäre-Vignette 2 Klassenatmosphäre	.016	.230	.069	.945	1.000
Vignette 3 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-.738	.230	-3.209	.001	.008
Vignette 2 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-.722	.230	-3.140	.002	.010

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanz (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E23

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen der Klassenatmosphäre (Lehrkräfte)

Gesamtzahl	18
Teststatistik	16.297
Freiheitsgrad	3
Asymptotische Sig. (zwei-seitiger Test)	

Tabelle E24

Paarweise Vergleiche der Einschätzung der Klassenatmosphäre (Lehrkräfte)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 1 Klassenatmosphäre-Vignette 3 Klassenatmosphäre	-.167	.430	-.387	.699	1.000
Vignette 1 Klassenatmosphäre-Vignette 2 Klassenatmosphäre	-.361	.430	-.839	.401	1.000
Vignette 1 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-1.250	.430	-2.905	.004	.022
Vignette 3 Klassenatmosphäre-Vignette 2 Klassenatmosphäre	.194	.430	.452	.651	1.000
Vignette 3 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-1.083	.430	-2.517	.012	.071
Vignette 2 Klassenatmosphäre-Vignette 4 Klassenatmosphäre	-.889	.430	-2.066	.039	.233

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanz (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E25

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu Barrieren je Vignette (Bachelorstudierende)

Gesamtzahl	109
Teststatistik	14.463
Freiheitsgrad	3
Asymptotische Sig. (zweiseitiger Test)	.002

Tabelle E26

Paarweise Vergleiche der Einschätzungen von Items zu Barrieren (Bachelorstudierende)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 1 Mittelwert Barrieren-Vignette 4 Mittelwert Barrieren	-.138	.175	-.787	.431	1.000
Vignette 1 Mittelwert Barrieren-Vignette 2 Mittelwert Barrieren	-.394	.175	-2.256	.024	.144
Vignette 1 Mittelwert Barrieren-Vignette 3 Mittelwert Barrieren	-.587	.175	-3.358	<.001	.005
Vignette 4 Mittelwert Barrieren-Vignette 2 Mittelwert Barrieren	.257	.175	1.469	.142	.851
Vignette 4 Mittelwert Barrieren-Vignette 3 Mittelwert Barrieren	.450	.175	2.571	.010	.061
Vignette 2 Mittelwert Barrieren-Vignette 3 Mittelwert Barrieren	-.193	.175	-1.102	.271	1.000

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E27

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu Barrieren je Vignette (Masterstudierende)

Gesamtzahl	66
Teststatistik	22.126
Freiheitsgrad	3
Asymptotische Sig. (zweiseitiger Test)	<.001

Tabelle E28

Paarweise Vergleiche der Einschätzungen von Items zu Barrieren (Masterstudierende)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 1 Mittelwert Barrieren-Vignette 2 Mittelwert Barrieren	-.674	.225	-3.000	.003	.016
Vignette 1 Mittelwert Barrieren-Vignette 4 Mittelwert Barrieren	-.841	.225	-3.742	<.001	.001
Vignette 1 Mittelwert Barrieren-Vignette 3 Mittelwert Barrieren	-.909	.225	-4.045	<.001	.000
Vignette 2 Mittelwert Barrieren-Vignette 4 Mittelwert Barrieren	-.167	.225	-.742	.458	1.000
Vignette 2 Mittelwert Barrieren-Vignette 3 Mittelwert Barrieren	-.235	.225	-1.045	.296	1.000
Vignette 4 Mittelwert Barrieren-Vignette 3 Mittelwert Barrieren	.068	.225	.303	.762	1.000

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E29

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu Barrieren je Vignette (Lehrkräfte)

Gesamtzahl	18
Teststatistik	1.253
Freiheitsgrad	3
Asymptotische Sig. (zweiseitiger Test)	.740

Tabelle E30

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu gelingenden Situationen je Vignette (Bachelorstudierende)

Gesamtzahl	109
Teststatistik	44.468
Freiheitsgrad	3
Asymptotische Sig. (zweiseitiger Test)	<.001

Tabelle E31

Paarweise Vergleiche der Einschätzungen von Items zu gelingenden Situationen (Bachelorstudierende)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 3 Mittelwert gelingende Situationen-Vignette 2 Mittelwert gelingende Situationen	.188	.175	1.075	.282	1.000
Vignette 3 Mittelwert gelingende Situationen-Vignette 1 Mittelwert gelingende Situationen	.693	.175	3.961	<.001	.000
Vignette 3 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-1.009	.175	-5.771	<.001	.000
Vignette 2 Mittelwert gelingende Situationen-Vignette 1 Mittelwert gelingende Situationen	.505	.175	2.885	.004	.023
Vignette 2 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-.821	.175	-4.695	<.001	.000
Vignette 1 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-.317	.175	-1.810	.070	.422

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E32

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu gelingenden Situationen je Vignette (Masterstudierende)

Gesamtzahl	66
Teststatistik	48.116
Freiheitsgrad	3
Asymptotische Sig. (zweiseitiger Test)	<.001

Tabelle E33

Paarweise Vergleiche der Einschätzungen von Items zu gelingenden Situationen (Masterstudierende)

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 3 Mittelwert gelingende Situationen-Vignette 2 Mittelwert gelingende Situationen)	.758	.225	3.371	<.001	.004
Vignette 3 Mittelwert gelingende Situationen-Vignette 1 Mittelwert gelingende Situationen	.970	.225	4.315	<.001	.000
Vignette 3 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-1.485	.225	-6.607	<.001	.000
Vignette 2 Mittelwert gelingende Situationen-Vignette 1 Mittelwert gelingende Situationen	.212	.225	.944	.345	1.000
Vignette 2 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-.727	.225	-3.236	.001	.007
Vignette 1 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-.515	.225	-2.292	.022	.131

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanzen (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E34

Zusammenfassung der zweifaktoriellen Varianzanalyse für Ränge nach Friedman bei verbundenen Stichproben, Vergleich der Einschätzungen von Items zu gelingenden Situationen je Vignette (Lehrkräfte)

Gesamtzahl	18
Teststatistik	10.440
Freiheitsgrad	3
Asymptotische Sig. (zweiseitiger Test)	.015

Tabelle E35*Paarweise Vergleiche der Einschätzungen von Items zu gelingenden Situationen (Lehrkräfte)*

Sample 1-Sample 2	Teststatistik	Std.-Fehler	Standardteststatistik	Sig.	Anp. Sig. ^a
Vignette 1 Mittelwert gelingende Situationen-Vignette 3 Mittelwert gelingende Situationen	-.417	.430	-.968	.333	1.000
Vignette 1 Mittelwert gelingende Situationen-Vignette 2 Mittelwert gelingende Situationen	-.694	.430	-1.614	.107	.639
Vignette 1 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-1.333	.430	-3.098	.002	.012
Vignette 3 Mittelwert gelingende Situationen-Vignette 2 Mittelwert gelingende Situationen	.278	.430	.645	.519	1.000
Vignette 3 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-.917	.430	-2.130	.033	.199
Vignette 2 Mittelwert gelingende Situationen-Vignette 4 Mittelwert gelingende Situationen	-.639	.430	-1.485	.138	.826

Anmerkung. Jede Zeile prüft die Nullhypothese, dass die Verteilungen in Stichprobe 1 und Stichprobe 2 gleich sind. Asymptotische Signifikanz (zweiseitige Tests) werden angezeigt. Das Signifikanzniveau ist .050. a. Signifikanzwerte werden von der Bonferroni-Korrektur für mehrere Tests angepasst.

Tabelle E36

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1 (Bachelorstudierende)

Korrelationen^a

		Vignette 1 Klassenatmo- sphäre	Vignette 1 Mittelwert Barrieren	Vignette 1 Mittelwert gelingende Si- tuationen
Spearman-Rho Vignette 1 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)			
	N	109		
Vignette 1 Mittelwert Barrieren	Korrelationskoeffizient-.471**		--	
	Sig. (2-seitig)	<.001	.	
	N	109	109	
Vignette 1 Mittelwert gelingende Situationen	Korrelationskoeffizient.497**		-.574**	--
	Sig. (2-seitig)	<.001	<.001	.
	N	109	109	109

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = 4. Bachelorsemester

Tabelle E37

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2 (Bachelorstudierende)

Korrelationen^a

		Vignette 2 Klassenatmo- sphäre	Vignette 2 Mittelwert Barrieren	Vignette 2 Mittelwert gelingende Si- tuationen)
Spearman-Rho Vignette 2 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)	.		
	N	109		
Vignette 2 Mittelwert Barrieren	Korrelationskoeffizient--	-.352**	--	
	Sig. (2-seitig)	<.001	.	
	N	109	109	
Vignette 2 Mittelwert gelingende Situationen)	Korrelationskoeffizient--	.447**	-.576**	--
	Sig. (2-seitig)	<.001	<.001	.
	N	109	109	109

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = 4. Bachelorsemester

Tabelle E38

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3 (Bachelorstudierende)

Korrelationen^a

		Vignette 3 Klassenatmo- sphäre	Vignette 3 Mittelwert Barrieren	Vignette 3 Mittelwert gelingende Si- tuationen)
Spearman-Rho Vignette 3 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)	.		
	N	109		
Vignette 3 Mittelwert Barrieren	Korrelationskoeffizient--	-.399**	--	
	Sig. (2-seitig)	<.001	.	
	N	109	109	
Vignette 3 Mittelwert gelingende Situationen)	Korrelationskoeffizient--	.364**	-.561**	--
	Sig. (2-seitig)	<.001	<.001	.
	N	109	109	109

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = 4. Bachelorsemester

Tabelle E39

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4 (Bachelorstudierende)

Korrelationen^a

		Vignette 4 Klassenatmo- sphäre	Vignette 4 Mittelwert Barrieren	Vignette 4 Mittelwert gelingende Si- tuationen
Spearman-Rho Vignette 4 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)			
	N	109		
Vignette 4 Mittelwert Barrieren	Korrelationskoeffizient-.469**		--	
	Sig. (2-seitig)	<.001	.	
	N	109	109	
Vignette 4 Mittelwert gelingende Situationen	Korrelationskoeffizient.526**		-.550**	--
	Sig. (2-seitig)	<.001	<.001	.
	N	109	109	109

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = 4. Bachelorsemester

Tabelle E40

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1 (Masterstudierende)

Korrelationen^a

	Vignette 1 Klassenatmo- sphäre	Vignette 1 Mittelwert Barrieren	Vignette 1 Mittelwert gelingende Si- tuationen
Spearman-Rho Vignette 1 Klassenatmosphäre	Korrelationskoeffizient-- Sig. (2-seitig) N	.	
		66	
Vignette 1 Mittelwert Barrieren	Korrelationskoeffizient-.554** Sig. (2-seitig) N	-- . 66	
		66	
Vignette 1 Mittelwert gelingende Situationen	Korrelationskoeffizient.632** Sig. (2-seitig) N	-.586** .<.001 66	-- . 66
		66	66

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Masterstudierende

Tabelle E41

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2 (Masterstudierende)

Korrelationen^a

		Vignette 2 Klassenatmo- sphäre	Vignette 2 Mittelwert Barrieren	Vignette 2 Mittelwert gelingende Si- tuationen)
Spearman-Rho Vignette 2 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)			
	N	63		
Vignette 2 Mittelwert Barrieren	Korrelationskoeffizient--	.311*	--	
	Sig. (2-seitig)	.013	.	
	N	63	66	
Vignette 2 Mittelwert gelingende Situationen)	Korrelationskoeffizient--	.199	-.386**	--
	Sig. (2-seitig)	.118	.001	.
	N	63	66	66

Anmerkung. *. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig). **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Masterstudierende

Tabelle E42

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3 (Masterstudierende)

Korrelationen^a

		Vignette 3 Klassenatmo- sphäre	Vignette 3 Mittelwert Barrieren	Vignette 3 Mittelwert gelingende Si- tuationen)
Spearman-Rho Vignette 3 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)			
	N	65		
Vignette 3 Mittelwert Barrieren	Korrelationskoeffizient--	.172	--	
	Sig. (2-seitig)	.170	.	
	N	65	66	
Vignette 3 Mittelwert gelingende Situationen)	Korrelationskoeffizient--	.412**	-.437**	--
	Sig. (2-seitig)	<.001	<.001	.
	N	65	66	66

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Masterstudierende

Tabelle E43

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4 (Masterstudierende)

Korrelationen^a

		Vignette 4 Klassenatmo- sphäre	Vignette 4 Mittelwert Barrieren	Vignette 4 Mittelwert gelingende Si- tuationen
Spearman-Rho Vignette 4 Klassenatmosphäre	Korrelationskoeffizient-- Sig. (2-seitig)	.		
	N	66		
Vignette 4 Mittelwert Barrieren	Korrelationskoeffizient-- Sig. (2-seitig)	-.260*	-- .	
	N	66	66	
Vignette 4 Mittelwert gelingende Situationen	Korrelationskoeffizient-- Sig. (2-seitig)	.343** .005	-.464** <.001	-- .
	N	66	66	66

Anmerkung. *. Die Korrelation ist auf dem 0,05 Niveau signifikant (zweiseitig). **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Masterstudierende

Tabelle E44

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 1 (Lehrkräfte)

Korrelationen^a

	Vignette 1 Klassenatmo- sphäre	Vignette 1 Mittelwert Barrieren	Vignette 1 Mittelwert gelingende Si- tuationen
Spearman-Rho Vignette 1 Klassenatmosphäre	Korrelationskoeffizient-- Sig. (2-seitig) N	. 18	
Vignette 1 Mittelwert Barrieren	Korrelationskoeffizient-.079 Sig. (2-seitig) N	-- .756 18	-- .18
Vignette 1 Mittelwert gelingende Situationen	Korrelationskoeffizient.196 Sig. (2-seitig) N	-.685** .436 18	-- .18

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Lehrkräfte

Tabelle E45

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 2 (Lehrkräfte)

Korrelationen^a

		Vignette 2 Klassenatmo- sphäre	Vignette 2 Mittelwert Barrieren	Vignette 2 Mittelwert gelingende Si- tuationen)
Spearman-Rho Vignette 2 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)			
	N	18		
Vignette 2 Mittelwert Barrieren	Korrelationskoeffizient--	-.294	--	
	Sig. (2-seitig)	.236	.	
	N	18	18	
Vignette 2 Mittelwert gelingende Situationen)	Korrelationskoeffizient--	.418	-.674**	--
	Sig. (2-seitig)	.084	.002	.
	N	18	18	18

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Lehrkräfte

Tabelle E46

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 3 (Lehrkräfte)

Korrelationen^a

		Vignette 3 Klassenatmo- sphäre	Vignette 3 Mittelwert Barrieren	Vignette 3 Mittelwert gelingende Si- tuationen)
Spearman-Rho Vignette 3 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)	.		
	N	18		
Vignette 3 Mittelwert Barrieren	Korrelationskoeffizient--	.246	--	
	Sig. (2-seitig)	.324	.	
	N	18	18	
Vignette 3 Mittelwert gelingende Situationen)	Korrelationskoeffizient--	.813**	-.451	--
	Sig. (2-seitig)	<.001	.060	.
	N	18	18	18

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Lehrkräfte

Tabelle E47

Korrelationen nach Spearman-Rho zwischen den Beurteilungen der Klassenatmosphäre, der Items zu Barrieren und der Items zu gelingenden Situationen in Vignette 4 (Lehrkräfte)

Korrelationen^a

		Vignette 4 Klassenatmo- sphäre	Vignette 4 Mittelwert Barrieren	Vignette 4 Mittelwert gelingende Si- tuationen
Spearman-Rho Vignette 4 Klassenatmosphäre	Korrelationskoeffizient--	.		
	Sig. (2-seitig)			
	N	18		
Vignette 4 Mittelwert Barrieren	Korrelationskoeffizient--	.408	--	
	Sig. (2-seitig)	.093	.	
	N	18	18	
Vignette 4 Mittelwert gelingende Situationen	Korrelationskoeffizient--	.171	-.723**	--
	Sig. (2-seitig)	.498	<.001	.
	N	18	18	18

Anmerkung. **. Die Korrelation ist auf dem 0,01 Niveau signifikant (zweiseitig). a. Expertise = Lehrkräfte

Tabelle E48

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 1

	Vignette 1 Mittelwert Barrieren	Vignette 1 Mittelwert gelingende Situationen
Mann-Whitney-U-Test	304.000	317.000
Wilcoxon-W	1630.000	437.000
Z	-1.220	-1.018
Asymp. Sig. (2-seitig)	.222	.309

Anmerkung. Gruppenvariable: Vignette 1 markiert (dichotom)

Tabelle E49

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 2

	Vignette 2 Mittelwert Barrieren	Vignette 2 Mittelwert gelingende Situationen)
Mann-Whitney-U-Test	439.500	408.500
Wilcoxon-W	649.500	1489.500
Z	-.289	-.749
Asymp. Sig. (2-seitig)	.772	.454

Anmerkung. Gruppenvariable: Vignette 2 markiert (dichotom)

Tabelle E50

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 3

	Vignette 3 Mittelwert Barrieren	Vignette 3 Mittelwert gelingende Situationen)
Mann-Whitney-U-Test	308.000	329.500
Wilcoxon-W	461.000	482.500
Z	-1.646	-1.308
Asymp. Sig. (2-seitig)	.100	.191

Anmerkung. Gruppenvariable: Vignette 3 markiert (dichotom)

Tabelle E51

Teststatistiken des Mann-Whitney-U-Tests bei unabhängigen Stichproben, Vergleich der Masterstudierenden (Markierungsaufgabe bearbeitet/nicht bearbeitet) anhand der Mittelwerte für Items zu Barrieren und zu gelingenden Situationen in Vignette 4

	Vignette 4 Mittelwert Barrieren	Vignette 4 Mittelwert gelingende Situationen)
Mann-Whitney-U-Test	223.500	212.000
Wilcoxon-W	1601.500	317.000
Z	-2.301	-2.409
Asymp. Sig. (2-seitig)	.021	.016

Anmerkung. Gruppenvariable: Vignette 4 markiert (dichotom)

Zusammenfassung

Theoretischer Hintergrund: Die professionelle Wahrnehmung unterrichtsrelevanter Situationen gilt als Schlüsselkompetenz von Lehrkräften. Aus sonderpädagogischer Perspektive sind insbesondere solche Situationen relevant, in denen es zu Störungen des Lehr-Lern-Prozesses kommt, in denen also *Lernbarrieren* entstehen könnten. Potenzielle Lernbarrieren zu *erkennen*, wird damit zu einer wesentlichen Aufgabe von Lehrkräften.

Die Analyse von Lehr-Lern-Situationen muss über die Beobachtung individueller Bedingungen, die zu Lernbarrieren beitragen können, hinausgehen und auch alle anderen Bedingungen in den Blick nehmen (Langner & Jugel, 2019; Ricken, 2017; Vock & Gronastaj, 2017). Merkmale der Klassenführung und Lernunterstützung gelten dabei als besonders relevant (Holodynski et al., 2017). Die Beschreibung und Erfassung der Kompetenz Lernbarrieren zu erkennen, die in der fehlenden oder mangelhaften Umsetzung von Klassenführungs- und Lernunterstützungsmerkmalen liegen, fehlen aber bisher als Ansatz.

Vorgehen: In der vorliegenden Arbeit werden ausgehend von der Unterrichtsqualitätsforschung Merkmale der Klassenführung und Lernunterstützung herausgearbeitet. Deren fehlende oder fehlerhafte Umsetzungen werden als potenzielle Lernbarrieren verstanden und operationalisiert. Die Kompetenz, diese zu erkennen, wird auf der Basis des Kompetenzmodells von Blömeke et al. (2015a) als situationsspezifische Fähigkeit beschrieben und betont u. a. die bewusste Fokussierung und Wahrnehmung von Unterrichtsbedingungen und die Beurteilung hinsichtlich der förderlichen oder hinderlichen Auswirkung auf die Lehr-Lernsituation.

In vier Entwicklungsschritten wird ein Verfahren entwickelt, bei dem anhand der Beurteilung von Textvignetten die Kompetenz zum Erkennen von Lernbarrieren erfassbar werden soll. Textvignetten und Items werden theoriegeleitet konstruiert, erprobt und auf der Basis der gewonnenen Daten sowie mehrerer Expertendiskussionen variiert und weiterentwickelt. Die Stichproben der vier Entwicklungsschritte umfassen insgesamt 935 Lehrkräfte und Studierende verschiedener Lehramtsstudiengänge aus verschiedenen Fachsemestern. Daten zu den Beurteilungen der Vignetten werden deskriptiv und anhand non-parametrischer Verfahren ausgewertet.

Ergebnisse: Im ersten Entwicklungsschritt wird erprobt, wie spezifisch Impulse durch die Vignette und das Beurteilungsformat sein müssen, damit Beurteilungsprozesse hinsichtlich der Klassenführung und Lernunterstützung in angemessenem Schwierigkeitsgrad in Gang gesetzt werden. Dazu werden zwei unterschiedliche Vignettenentwürfe mit offenen und geschlossenen Beurteilungsimpulsen eingesetzt. In den Ergebnissen zeigt sich, dass weniger spezifische offene Impulse nicht dazu führen, dass Studierende Unterrichtsbedingungen fokussieren, sondern vor

allem Bedingungen des Kindes als Ursache für problematische Lehr-Lern-Situationen in Betracht ziehen. In der weiteren Entwicklung werden deshalb Vignetten eingesetzt, in denen konkrete Unterrichtssituationen beschrieben werden, die Lernbarrieren in der Klassenführung und Lernunterstützung enthalten, und die anhand von Items zur Klassenführung und Lernunterstützung zu beurteilen sind. Für diese Variante zeigt sich im ersten Entwicklungsschritt eine akzeptable Reliabilität und es deuten sich Gruppenunterschiede zugunsten der Sonderpädagogikstudierenden an, wobei Schwierigkeitsgrade noch nicht angemessen sind.

Im zweiten Entwicklungsschritt werden Vignette und Items inhaltlich, sprachlich und methodisch überarbeitet und erprobt. Die Ergebnisse liefern Hinweise darauf, dass vor allem Kontextinformationen zur Vignette einen starken Einfluss auf die Beurteilung haben und die beschriebene Unterrichtssituation eher holistisch als differenziert beurteilt wird. Gruppenunterschiede zugunsten höherer Fachsemester zeigen sich zwischen Studierenden des 2. und 6. Bachelorsemesters.

Der dritte Entwicklungsschritt hat zum Ziel, Schwierigkeitsgrade über die Darstellung der Merkmale in den Vignetten zu beeinflussen und differenziertere Beurteilungen anzuregen. Dies wird umgesetzt, indem auf der Basis realer Unterrichtsbeispiele vier neue Vignetten bei gleichbleibenden Konstruktionskriterien entwickelt und validiert werden. Neben Lernbarrieren werden auch gelingende Situationen in den Vignetten dargestellt. In den Ergebnissen zeigen sich zufriedenstellende mittlere Schwierigkeitsgrade und zufriedenstellende Reliabilitäten für drei der vier Vignetten. Die Vignetten selbst haben einen starken Einfluss auf die Beurteilungen und werden eher nach Gesamteindruck als differenziert beurteilt. Gruppenunterschiede zeigen sich nur im Vergleich der Lehrkräfte mit Studierenden, Lehrkräfte beurteilen Lernbarrieren zutreffender, gelingende Situationen aber weniger zutreffend als Studierende.

Der vierte Entwicklungsschritt hat zum Ziel, noch stärker eine Differenzierung zu betonen und Einflüsse auf die Vignettenbeurteilung zu überprüfen. Dazu werden die Vignetten nochmals überarbeitet und validiert, gelingende Situationen und Lernbarrieren werden in insgesamt ausgewogener Anzahl dargestellt, innerhalb der Vignetten aber in der Anzahl variiert. Die Gesamtwahrnehmung der Klassenatmosphäre wird zusätzlich erhoben. Außerdem wird in einer Teilstichprobe eine Variation der Instruktion erprobt. Die Ergebnisse zeigen, dass mittlere Schwierigkeitsgrade und Varianzen zufriedenstellend sind, Items zu Barrieren und gelingenden Situationen aber nicht differenziert beurteilt werden und getrennte Skalen bilden. Die Vignetten werden in allen Gruppen eher holistisch beurteilt – Lehrkräfte urteilen insgesamt eher negativ und erzielen so höhere Werte für Barrieren. Studierende der Allgemeinpädagogik urteilen eher positiv und erzielen deshalb höhere Werte für Gelingendes. Während die Urteile der Lehrkräfte

weitgehend unabhängig von der Einschätzung der Klassenatmosphäre und der Vignettenkonstruktion sind, zeigen sich vor allem bei Bachelorstudierenden starke Zusammenhänge der Beurteilungen mit der Einschätzung der Klassenatmosphäre. Ein relevanter Zusammenhang mit der Variation der Instruktion zeigt sich nicht.

Fazit: Das entwickelte textvignettenbasierte Verfahren eignet sich, um *unterschiedliche* Beurteilungen von Lernbarrieren und von gelingenden Situationen reliabel zu erfassen und Gruppenunterschiede abzubilden. Die *Expertise*, Gelingendes und Mislingendes *differenziert* zu beurteilen und das gleichbleibend gut in allen vier Vignetten, zeigt sich in keiner der Teilstichproben. Eine solche Expertise scheint sich ohne spezifische Lerngelegenheit nicht zu entwickeln. Beurteilungen sind in allen Gruppen eher holistisch und abhängig von Gesamteindrücken und unterschiedlichen schwerpunktmäßigen Fokussierungen auf misslingende oder gelingende Prozesse. Die Vignetten selbst haben unabhängig von den *gezielten* Konstruktionskriterien und -variationen einen starken Einfluss auf die Beurteilungen.

Summary

Theoretical background: The professional perception of teaching-relevant situations is considered a key competence of teachers. From a special education perspective, situations are particularly relevant in which disruptions to the teaching-learning process occur and in which learning barriers could arise. Recognizing potential barriers to learning therefore becomes an essential task for teachers.

The analysis of teaching-learning situations must go beyond the observation of individual conditions that can contribute to learning barriers and also take all other conditions into account (Langner & Jugel, 2019; Ricken, 2017; Vock & Gronastaj, 2017). Features of classroom management and learning support are considered particularly relevant (Holodynski et al., 2017). However, the description and recording of the competence to recognize learning barriers that lie in the absence or inadequate implementation of classroom management and learning support features is still missing as an approach.

Approach: In this work, characteristics of classroom management and learning support are identified based on teaching quality research. Their missing or incorrect implementations are understood and operationalized as potential learning barriers. The competence to recognize this is based on the competence model by Blömeke et al. (2015a) described as a situation-specific ability and emphasizes, among other things, the conscious focus and perception of teaching conditions and the assessment of the beneficial or hindering impact on the teaching-learning situation.

In four development steps, a process is being developed in which the competence to recognize learning barriers can be determined by assessing text vignettes. Text vignettes and items are constructed based on theory, tested and varied and further developed on the basis of the data obtained and several expert discussions. The samples from the four development steps were taken from a total of 935 teachers and students during various teacher training courses from different semesters. Data on the assessments of the vignettes are evaluated descriptively and using non-parametric methods.

Results: In the first development step, it is tested how specific impulses must be provided by the vignette and the assessment format so that assessment processes with regard to class management and learning support are initiated with an appropriate level of difficulty. For this purpose, two different vignette designs with open and closed assessment impulses are used. The results show that less specific open impulses do not lead to students focusing on teaching conditions, but rather considering the child's conditions as the cause of the problematic teach-

ing-learning situation. In further development, vignettes will therefore be used in which concrete teaching situations are described, which contain learning barriers in class management and learning support, and which can be assessed using items on class management and learning support. In the first development step, this variant shows acceptable reliability and there are group differences in favor of special education students, although levels of difficulty are not yet appropriate.

In the second development step, the vignette and items are revised and tested in terms of content, language and methodology. The results provide evidence that contextual information about the vignette has a strong influence on the assessment and that the teaching situation described is assessed more holistically than differentiated. Group differences in favor of higher semesters can be seen between students in the 2nd and 6th Bachelor semesters.

The third development step aims to influence levels of difficulty through the representation of the characteristics in the vignettes and to encourage more differentiated assessments. This is implemented by developing and validating four new vignettes based on real teaching examples while maintaining the same design criteria. In addition to learning barriers, successful situations are also presented in the vignettes. The results show satisfactory average levels of difficulty and satisfactory reliabilities for three of the four vignettes. The vignettes themselves have a strong influence on the assessments and are judged more on the basis of an overall impression than on a differentiated basis. Group differences only become apparent when comparing teachers with students; teachers judge learning barriers more accurately, but successful situations less accurately than students.

The fourth development step aims to emphasize differentiation even more and to examine influences on the vignette assessment. For this purpose, the vignettes are revised and validated again, successful situations and learning barriers are presented in a balanced number overall, but the number varies within the vignettes, and the overall perception of the class atmosphere is recorded. In addition, a variation of the instructions is tested in a subsample. The results show that medium levels of difficulty and variance are satisfactory, but items about barriers and successful situations are not assessed differentiated and form separate scales. The vignettes are assessed more holistically in all groups - teachers judge them more negatively overall and thus achieve higher scores for barriers. General education students are more likely to judge things positively and therefore achieve higher scores for success. While the teachers' judgments are largely independent of the assessment of the class atmosphere and the vignette construction,

there are strong connections between the assessments and the assessment of the class atmosphere, especially among bachelor's students. There is no relevant connection with the variation of the instruction.

Conclusion: The text vignette-based method developed is suitable for reliably recording different assessments of learning barriers and successful situations and for depicting group differences. The expertise in assessing successful situations and learning barriers in a differentiated manner and doing so consistently well in all four vignettes is not evident in any of the sub-samples. Such expertise does not appear to develop without specific learning opportunities. Assessments in all groups are more holistic and dependent on overall impressions and different focuses on failing or successful processes. The vignettes themselves have a strong influence on the assessments, regardless of the targeted design criteria.

Liste der aus dieser Dissertation hervorgegangenen Veröffentlichungen

- Wenck, S., & Ricken, G. (2017, 11-13 Mai). *Lernbarrieren erkennen - Analyse der Wahrnehmung von Lehramtsstudierenden mithilfe unterschiedlich konstruierter Textvignetten* [Tagungspräsentation]. Frühjahrstagung der Arbeitsgruppe Empirische Sonderpädagogische Forschung (AESF), Rostock, Deutschland.
- Wenck, S., & Ricken, G. (2017, 21-23 Mai). *Wahrnehmung von Barrieren in Textvignetten* [Kongressbeitrag]. New International Perspectives on Future Teachers' Professional Competencies, Hamburg, Deutschland.
- Wenck, S., & Ricken, G. (2018, 15-17 November). *Erkennen von Lernbarrieren (ErLe)* [Tagungspräsentation]. Herbsttagung der Arbeitsgruppe Empirische Sonderpädagogische Forschung (AESF), Wuppertal, Deutschland.
- Wenck, S., & Ricken, G. (2019, 23-25 September). *ErLe – Erkennen von Lernbarrieren in Textvignetten. Ich mache einen Sitzkreis und alles wird gut?!* [Tagungspräsentation]. Tagung der Deutschen Gesellschaft für Erziehungswissenschaft (DGFE), Wuppertal, Deutschland.
- Wenck, S., & Ricken, G. (2019). ErLe- Erkennen von Lernbarrieren in Textvignetten. In B. Krause, & R. Beyer (Hrsg.), *Wissenschaftliche Veranstaltungen* (Band 8) (S. 85-93). ZeE.
- Wenck, S., & Ricken, G. (2020). ErLe – Erkennen von Lernbarrieren in Textvignetten. In M. Grosche, J. Decristan, K. Urton, G. Bruns, B. Ehl, & N. Jansen (Hrsg.), *Sonderpädagogik und Bildungsforschung – Fremde Schwestern?* (S. 240–245). Klinkhardt.

Lebenslauf

(entfällt aus datenschutzrechtlichen Gründen)

Eidesstattliche Versicherung

Stephanie Wenck, Radekoppel 15c, 22397 Hamburg

Hiermit erkläre ich durch meine Unterschrift an Eides statt:

1. Die eingereichte Dissertation mit dem Titel „Erkennen von Lernbarrieren in Textvignetten – ErLe, Entwicklung und Validierung eines quantitativen Verfahrens zur Beurteilung situationsspezifischer Fähigkeiten“ habe ich selbstständig und ohne unzulässige fremde Hilfe verfasst. Hierbei habe ich weder Textstellen von Dritten oder aus eigenen Prüfungsarbeiten noch Grafiken oder sonstige Materialien ohne Kennzeichnung übernommen.
2. Es sind ausschließlich die von mir angegebenen Quellen und Hilfsmittel verwendet worden.
3. Sämtliche wörtliche und nicht wörtliche Zitate aus anderen Werken sind gemäß den wissenschaftlichen Zitierregeln kenntlich gemacht.
4. Die von mir vorgelegte Arbeit ist bisher noch nicht, auch nicht teilweise, veröffentlicht worden.
5. Die von mir vorgelegte Arbeit ist bisher noch in keiner Form als Bestandteil einer Prüfungs-/Qualifikationsleistung vorgelegt worden.
6. Die von mir eingereichte Dissertation habe ich unter Beachtung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis erstellt.
7. Die geltende Promotionsordnung ist mir bekannt.
8. Über die Bedeutung und die strafrechtlichen Folgen einer falschen eidesstattlichen Erklärung gemäß § 156 StGB bin ich mir bewusst.
9. Ich erkläre an Eides statt, dass meine Angaben der Wahrheit entsprechen und ich diese nach bestem Wissen und Gewissen gemacht habe.

Hamburg, 05.02.2024

Stephanie Wenck