

**Entwicklung eines
interaktionalen
Simulationsmodells von
Unterrichtshandeln im Rahmen
der Implementierung der VR-
Lernumgebung Clasivir**

DISSERTATION ZUR ERLANGUNG DES GRADES
EINES DOKTORS DER NATURWISSENSCHAFTEN

VORGELEGT BEIM FACHBEREICH PSYCHOLOGIE
DER JOHANN WOLFGANG GOETHE-UNIVERSITÄT
IN FRANKFURT AM MAIN

VON

Sebastian Breitenbach
aus Bad Soden-Salmünster

1. Gutachter: Prof. Dr. Holger Horz
2. Gutachter: Prof. Dr. Rolf van Dick

Frankfurt, 2023

Vom Fachbereich Psychologie der

Johann Wolfgang Goethe-Universität als Dissertation angenommen.

Dekanin: Prof. Dr. Sonja Rohrmann

1. Gutachter: Prof. Dr. Holger Horz
2. Gutachter: Prof. Dr. Rolf van Dick

Datum der Disputation: 21.02.2024

Inhaltsverzeichnis

<i>Vorwort</i>	vi
<i>Anmerkungen zur sprachlichen Gestaltung der Arbeit</i>	vii
<i>Anmerkungen zur Zitation von Webinhalten</i>	viii
1 Einleitung und Fragestellung	1
2 Terminologie	3
2.1 Etymologie und historischer Abriß	3
2.2 Definition des Begriffes Simulation	7
2.3 Abgrenzung zu <i>(serious) games</i>	12
2.4 Vorteile und Gründe für die Nutzung von Simulationen	16
2.4.1 Empirische Studien	22
2.5 Bisherige Nutzung angrenzender Termini	25
2.5.1 <i>virtual classroom</i>	25
2.5.2 <i>virtual learning environment (VLE)</i>	26
2.5.3 <i>virtual worlds</i>	27
2.5.4 <i>educational simulations</i>	28
2.5.5 <i>classroom simulation / virtual classroom simulation</i>	29
2.5.6 <i>computer simulation</i>	29
2.5.7 <i>presence</i> und <i>immersion</i>	30
2.5.8 <i>microteaching</i>	32
2.5.9 <i>critical incidents</i>	33
2.5.10 <i>(intelligent) agents</i>	35
2.5.11 <i>pedagogical agents</i> und <i>intelligent tutoring systems (ITS)</i>	36
2.5.12 Simulationsfidelität	38
2.5.13 Simulationsmodelle und Konzeptionsmodelle	41
2.5.14 <i>Entitäten, Entitätstypen und Entity-Relationship-Modelle (ERM)</i>	43
2.5.15 <i>cognitive load</i>	43
3 Taxonomie der Unterrichtssimulationen	45
3.1 Nach Granularität	45
3.1.1 <i>task simulation (TS)</i>	45
3.1.2 <i>school lesson simulations (SLS)</i>	46
3.1.3 <i>school teacher simulation (STS)</i>	46
3.2 Nach Immersionsstufe	47
3.2.1 <i>(Hyper-)Text-based simulation</i>	47
3.2.2 <i>2D (animated) graphical simulations</i>	48
3.2.3 <i>3D graphical simulations</i>	48

3.2.4	<i>interactive video-based simulations (IVBS)</i>	50
3.2.5	<i>Virtual Reality (VR) / Augmented Reality (AR) simulations</i>	51
3.2.6	Analoge Simulationen	53
3.3	Nach Interaktionstyp	54
3.3.1	Typ 1 – Deterministischer Entscheidungsbaum	56
3.3.2	Typ 2 – Entscheidungsbaum mit einfacher Randomisierung (Stochastische Simulation)	57
3.3.3	Typ 3 – Entscheidungsbaum mit komplexer algorithmischer Steuerung	58
3.3.4	Typ 4 – Komplexe algorithmische Steuerung ohne Entscheidungsbaum	60
3.3.5	Typ 5 – Komplexe algorithmische Steuerung mit eingeschränktem Entscheidungsbaum	64
3.3.6	Typ 6 – <i>human-in-the-loop</i>	64
3.3.7	Typ 7 – <i>digital/virtual puppetry</i>	65
3.3.8	Typ 8 – Reale Handlungsdurchführung	68
3.4	Andere Möglichkeiten der Taxonomie	68
4	Simulatoren in der Lehrkräftebildung	71
4.1	Aha! Classroom Sim	71
4.2	Behavior Breakthroughs	73
4.3	Breaking Bad Behaviors	74
4.4	ClassSim	78
4.5	Cook School District Simulation	80
4.6	Der Simulierte Klassenraum (SKR) / Simulated Classroom	85
4.7	Diversity Simulation	87
4.8	First Class	88
4.9	Inner-City Simulation Laboratory / Teaching Problems Laboratory	92
4.10	IVT (Interactive Virtual Training) / IVT-T (Interactive Virtual Training for Teachers))	94
4.11	Kognito Interactive Simulations	96
4.12	Lesson Sketch / SimTeach / ThEMaT I-III	98
4.13	Mr. Land's Sixth Grade	99
4.14	SimAula 1.0	102
4.15	simClass / simClass II	103
4.16	simSchool 1.0	105
4.16.1	Einführung und Entwicklungsgeschichte	105
4.16.2	Grafischer Aufbau und UI	107
4.16.3	Interaktionsmodellierung	107

4.16.4	Theoretische Grundlage / Konzeptionsmodell	109
4.16.5	Algorithmische Umsetzung – Simulationsmodell	112
4.16.6	Aufgaben (Tasks)	115
4.16.7	Kommentare	118
4.16.8	Anwendung, Zielgruppe und wahrgenommene Fidelität	119
4.17	simSchool 2.0	121
4.18	Simfta / SimInClass	124
4.19	STAR Simulator / TeachME / TeachLivE (TLE) / SimLab / Mursion	127
4.20	The Tigerlake Project	130
4.21	SimTeacher	131
4.22	Transition to Teaching Classroom Management	132
4.23	VCS – Virtual Classroom Simulator	134
4.24	Verbal Abuse Simulation	135
4.25	Virtual School Class (VSC)	136
4.26	Virtual Simulated ClassRoom (V.C.R.)	137
4.27	VR-Klassenraum	137
4.28	VR-Tool von Bailenson et al. (ohne Bezeichnung)	139
4.29	VR-Tool von Yaacov J. Katz (ohne Bezeichnung)	140
4.30	VR-Tool von Stavroulia et al. (ohne Bezeichnung)	140
4.31	Andere Simulationen	142
5	Entwicklung des interaktionalen Simulationsmodells	143
5.1	Planung der Simulation Clasivir	143
5.1.1	DOIT-Modell: Didaktik	144
5.1.2	DOIT-Modell: Organisation	149
5.1.3	DOIT-Modell: Individuum	151
5.1.4	DOIT-Modell: Technik	152
5.2	Funktionsweise von Clasivir	153
5.3	Generelle Überlegungen zum Simulationsmodell	155
5.4	Implementation eines Prototyps	159
5.5	Persönlichkeitsbezogene Variablen: Namen der SuS / Diversity	160
5.6	Persönlichkeitsbezogene Variablen: Emotionsmodell	163
5.7	Persönlichkeitsbezogene Variablen: Kognitive Fähigkeit	164
5.8	Situative Variablen: Konzentration	165
5.9	Situative Variablen: Emotionaler Status	171
5.10	Situative Variablen: Störungsverhalten	172
5.11	Situative Variablen: Aufgabenmodellierung	175
5.11.1	Adaptive Aufgabenpassung	175
5.11.2	Verarbeitungsalgorithmus	179
5.11.3	<i>functional off-task</i>	180

5.11.4	Vearbeitungszeitalgorithmus	181
5.11.5	Lösungsalgorithmus	181
5.11.6	Hereinrufen-Algorithmus	183
5.11.7	Meldealgorithmus	184
5.11.8	Meldungsabbruchalgorithmus	185
5.12	Situative Variablen: Lehrkraftaktionen	185
5.12.1	Aufgaben/Fragen	186
5.12.2	Informationen geben	186
5.12.3	<i>classroom management</i>	187
5.12.4	SuS aufrufen	189
5.12.5	Feedback auf Antwort geben	190
6	Methodische Anlage der Untersuchung	190
6.1	Forschungsdesign	191
6.1.1	Hypothesen	192
6.2	Quantitative Validierung des Modells	194
6.2.1	Stichprobe	195
6.2.2	Material	197
6.2.3	Durchführung	198
6.3	Mockup-Videos	200
6.3.1	Kodierung	203
6.3.2	Zur Konstruktion der Mockup-Videos	204
6.4	Methodenkritik	206
6.4.1	Betrachtung des Kodierungsprozesses	207
6.4.2	Betrachtung der Clasivir-Funktionsweise	212
7	Ergebnisse	213
7.1	Quantitative Ergebnisse der Modellvalidierung	213
7.1.1	Datenaufbereitung	213
7.1.2	Deskriptivstatistische Ergebnisse	218
7.1.3	Inferenzstatistischer Schluss	228
7.2	Qualitative Ergebnisse der Modellvalidierung	230
8	Zusammenfassung und Diskussion	235
8.1	Einordnung der Ergebnisse	235
8.2	Grenzen der Untersuchung	237
8.3	Ausblick	238
	Anhang	242
	Screenshots und Abbildungen der Simulatoren	242
	Konzeptionsmodell-Simulationsmodell-Übersetzungsgraphen	261

Tabellen der Clasivir-Variablen	266
Fragebogen	284
Literatur	289
Abbildungsverzeichnis	321
Tabellenverzeichnis	325

Vorwort

Ich möchte an dieser Stelle die Gelegenheit nutzen mich bei den Menschen zu bedanken, die mir geholfen haben diese Arbeit zu verfassen.

Mein besonderer Dank gilt Prof. Dr. Holger Horz, der es mir ermöglicht hat meine Arbeit an der Akademie für Bildungsforschung und Lehrkräftebildung zu verfassen und der mich ausgiebig im Forschungsprozess unterstützt hat. Vor allem möchte ich hervorheben, dass sich die Vernetzung mit anderen im Feld von *digital literacy* arbeitenden Personen als überaus hilfreich herausgestellt hat.

Herrn Prof. Dr. Rolf van Dick danke ich für die Bereitschaft als Zweitgutachter für diese Arbeit zu Verfügung zu stehen.

Zudem gilt mein Dank Dr. Johannes Appel, mit dem ich zahlreiche Gespräche zur inhaltlichen Natur meiner Dissertation führen konnte, die weit über das Maß seiner Rolle als Projektkoordinator der Projekte Level/The Next Level hinausreichten.

Ebenfalls möchte ich mich bei meinen zahlreichen Kolleginnen und Kollegen bedanken, die mir in unterschiedlichster Weise geholfen haben diese Arbeit zu einem Erfolg zu machen. Insbesondere möchte ich danken: Prof. Dr. Matthias Herrle, Dr. Sebastian Stehle, Dr. Jette Horstmeyer, Benjamin Tolweth, Prof. Dr. Miriam Hansen und Dr. Julia Mendzheritskaya. Einen besonderen Dank möchte ich auch Katharina Basar dafür aussprechen, die mich mit kollegialem Rat in verschiedensten projekt- und dissertationsbezogenen Fragen unterstützt hat.

Größter Dank jedoch gilt meiner Mutter Barbara Schelling-Breitenbach und meiner Lebensgefährtin Eva-Johanna Morgen für das Korrekturlesen dieser Arbeit und das große Verständnis, das sie für die langen Arbeitsstunden an diesem Text aufgebracht haben. Meine Schwester, Janina Stein, hat mir als Lehrkraft wichtige Hinweise zur Konstruktion meines Fragebogens geben können. Letztlich sei noch meiner Familie insgesamt ein Dank dafür ausgesprochen, dass sie mich immer unterstützt haben.

Anmerkungen zur sprachlichen Gestaltung der Arbeit

Diese Arbeit verwendet bei Personenbezeichnungen das generische Femininum, falls alle Geschlechter gemeint sind. Dies ist ein Kompromiss zwischen der Lesbarkeit des Textes und dem Bemühen um eine gendergerechte sprachliche Ausgestaltung. Menschen diversen, männlichen oder weiteren Geschlechts sind in der Verwendung des generischen Femininums immer mitgemeint. Auch Komposita werden in der weiblichen Form geführt (z.B. Schülerinnenkamera). Falls möglich, werden geschlechtsneutrale Bezeichnungen verwendet (z.B. Studierende, Lehrkräfte). Die sich bereits eingebürgerte Abkürzung ‚SuS‘ für Schülerinnen und Schüler wird ebenfalls verwendet.

Der überwiegende Teil der verwendeten Fachtermini wurde aus dem internationalen, häufig anglo-amerikanischen, Forschungskontext entnommen. Im Bemühen auf den internationalen Forschungsdiskurs zu verweisen wurde darauf verzichtet diese Termini einzudeutschen. Da z.B. ‚Klassenführung‘ oder ‚Klassenmanagement‘ nicht exakt dasselbe meint wie *classroom management*, wird nur der englischsprachige Begriff verwendet. Zur besseren Übersicht sind diese Termini durchgehend kursiv gesetzt. Englischsprachige Eigennamen, z.B. von Simulatoren, sind hingegen nicht kursiv. Da geplant ist, mit dieser Arbeit an den internationalen Forschungsdiskurs anzudocken, sind auch neu entwickelte Konstrukte, z.B. in der Taxonomie, in Englisch gehalten. Um eine bessere Lesbarkeit zu gewährleisten, wird der englische Begriff *pre-service teacher* durchgehend mit Lehramtsstudierende übersetzt.

Zahlenwerte sind an allen Stellen im anglo-amerikanischen Stil mit einem Punkt als Dezimaltrenner und einem Komma als Tausendertrennzeichen gehalten. Im statistischen Teil sind führende Nullen wie üblich entfernt.

Anmerkungen zur Zitation von Webinhalten

Da es für die Beschäftigung mit digitalen Unterrichtssimulationen unerlässlich ist, auch Webinhalte zu rezipieren, existieren derartige Quellen im Literaturverzeichnis dieser Arbeit. Viele Quellen sind jedoch nicht mehr im Original verfügbar, da sie zwischenzeitlich vom Netz genommen wurden. Man findet sie allerdings noch über die ‚Internet Waybackmaschine‘ der non-profit Organisation ‚The Internet Archive‘. In diesem Falle ist die URL der Waybackmaschine angegeben, über die man die Seite tatsächlich noch erreichen kann. Die restliche Zitationsweise der Webseiten folgt den Vorgaben der APA.

1 Einleitung und Fragestellung

Ziel dieser Arbeit ist die Entwicklung eines interaktionalen Simulationsmodells. Dies ist als eine Vorarbeit zu verstehen, um im nächsten Schritt einen VR-Simulator, genannt Clasivir (Classroom Simulator in Virtual Reality) zu entwickeln, der in der Lehrkräftebildung eingesetzt werden kann. Das interaktionale Simulationsmodell beschreibt dabei, wie sich die künstliche Intelligenz der SuS zu den Handlungen der Simulationspartizipantinnen, welche die Rolle einer Lehrkraft übernehmen, verhält. Um das Simulationsmodell zu entwickeln war es notwendig bisherige Termini und Simulationen zu untersuchen und eine eigene Taxonomie der Simulationen zu entwickeln.

[Kapitel 2](#) (S. 3) widmet sich der verwendeten Terminologie im Rahmen der pädagogischen Simulationsforschung. Dies ist kein reiner Selbstzweck, sondern im höchsten Maße erforderlich, um verschiedene Konstrukte wie Simulation, Simulator und (*serious*) *games* zu definieren sowie voneinander abzugrenzen. Dies scheint insbesondere deswegen geboten, da viele Begriffe äußerst unscharf und z.T. widersprüchlich verwendet werden. [Unterkapitel 2.4](#) (S. 16) fasst die Kernargumente zur Nutzung von Simulationen speziell in der Lehrkräftebildung zusammen und geht auf die Ergebnisse der vereinzelt vorliegenden empirischen Studien und Meta-Analysen ein.

In [Kapitel 3](#) (S. 45) wird eine Taxonomie von Unterrichtssimulationen entworfen. Dies war unter anderem deswegen geboten, da eine solche Taxonomie nicht vorlag und eine Analyse bereits vorhandener Simulationen eines wissenschaftlich-terminologischen Fundus bedurfte, der bislang noch nicht existierte. Die Analyse der bereits bestehenden Simulatoren in [Kapitel 4](#) (S. 71) dient als literaturanalytischer Teil zum einen dazu, einen Überblick über den Stand der Forschung zu gewinnen und damit das Fundament zur induktiven Taxonomieerstellung zu legen, zum anderen der Simulationsmodellanalyse, was wiederum die Basis für das folgende Kapitel legt.

Auf [Kapitel 5](#) (S. 143) liegt das Hauptaugenmerk dieser Arbeit. Dort wird beschrieben, wie durch Analyse von empirischen Arbeiten aus den Bereichen (pädagogische) Psychologie, Psychometrie, Pädagogik und Simulationsforschung ein interaktionales Simulationsmodell generiert werden kann. Dieses Simulationsmodell steuert, wie sich SuS in der Simulation in Abhängigkeit zu den von den Simulationspartizipantinnen gewählten Handlungsalternativen verhalten. Das Modell soll anschließend validiert werden.

Es ist wichtig zu betonen, dass diese Arbeit in der pädagogischen Psychologie angesiedelt ist und sich nicht vorrangig mit der Entwicklung einer Software beschäftigt, auch wenn die Entwicklung eines Prototyps des Simulationsmodells programmiertechnische und 3D-visuelle Aspekte umfasst.

In [Kapitel 6](#) (S. 190) wird das Studiendesign und die methodische Anla-

ge der Untersuchung behandelt. Die zur Simulationsfidelität erstellten drei Mockup-Videos werden in ihrer Konstruktion sowie ihren Chancen und Limitationen beschrieben. Das erste Mockup-Video zeigt in animierter Form die Handlungen und Verhaltensweisen von SuS einer realen Unterrichtsstunde, das zweite Mockup-Video wird durch die Vorhersagen des Clasivir-Simulationsmodells gesteuert. Das dritte Mockup-Video ist hingegen randomisiert, d.h. die Verhaltensweisen der SuS sind vollständig zufällig. Diese Mockup-Videos werden quantitativ im Rahmen eines Online-Fragebogens für Lehramtsstudierende analysiert und die Ergebnisse inferenzstatistisch ausgewertet. Der [Fragebogen](#) selbst kann im Anhang gefunden werden, die verwendeten Instrumente werden in [Abschnitt 6.2.2](#) (S. 197) beschrieben. Neben der Hauptthese $H_{A,0}$, dass es statistisch keinen Unterschied in dem wahrgenommenen Fidelitätsgrad des SuS-Verhaltens zwischen dem Clasivir-Simulationsmodell und der Übertragung der realen Unterrichtsstunde gibt, werden auch verschiedene andere Hypothesen überprüft, siehe [Abschnitt 6.1.1](#) (S. 192). Zusätzlich zur quantitativen Analyse wurde auch ein explorativer qualitativer Zugang gewählt. Dazu wurden die Freitextantworten der Partizipantinnen nach der qualitativen Inhaltsanalyse von Mayring (Mayring & Fenzl, 2019) ausgewertet. In [Kapitel 7](#) (S. 190) werden die deskriptiv- und inferenzstatistischen Ergebnisse der quantitativen Untersuchung präsentiert und die qualitative Analyse durchgeführt. In [Kapitel 8](#) (S. 213) werden die Ergebnisse diskutiert und in den Forschungskontext eingeordnet.

Diese Arbeit verfügt über einen recht umfangreichen Anhang. Alle im Rahmen von [Kapitel 4](#) (S. 71) analysierten Simulatoren sind dort mit Screenshots vertreten. Des Weiteren sind dort Graphen zur Übersetzung der Konzeptionsmodelle in Simulationsmodelle zu finden. Zuletzt wird eine Liste aller in Clasivir verwendeten Variablen präsentiert. Auf letztere sei besonderes Augenmerk gelegt, da hier auch Wahrscheinlichkeiten und Zusammenwirken der verschiedenen Variablen beschrieben werden.

2 Terminologie

2.1 Etymologie und historischer Abriss

Der Begriff Simulation leitet sich etymologisch von dem lateinischen Begriff *simulationem* ab, der Akkusativ-Form von *simulatio* (Nominativ). Er bezeichnet eine Imitation, eine Fingierung/Verstellung, eine Vortäuschung, einen Schein, eine Heuchelei (veraltet) oder auch ‚geheime Gedanken hegen‘ (Pfeifer, 1999). Das Verb *simulare* (*simulo*) wird mit ‚ähnlich machen, nachahmen, zum Schein äußern oder vorgeben‘ übersetzt („simulare“, 2018) und weist damit eine deutliche Nähe zur heutigen Sprachverwendung von ‚Simulation‘ auf (Vgl. auch Badiée & Kaufman, 2015, S.2). Eine Verwandtschaft besteht zwischen dem lateinischen Begriff *similis*, welcher ‚ähnlich‘ bedeutet. Während die Verwendung des deutschen Begriffes ‚simulieren‘ in der Umgangssprache des 18. Jahrhunderts auch ‚nachdenken, grübeln‘ und ‚aussinnen‘ umfasste, so ist diese Verwendung heute als veraltet anzusehen (Pfeifer, 1999). Noch im heutigen Sprachgebrauch hingegen wird das Verb ‚simulieren‘ im Sinne von ‚vortäuschen oder heucheln‘ verwendet, z.B. ‚eine Krankheit simulieren‘. Der lateinische Begriff *simulator*, der ‚Nachahmer‘ oder ‚Heuchler‘ bedeutet, wurde im Verlauf des 19. Jahrhunderts zum deutschen ‚Simulant‘. Die Verwendung des Begriffes von Simulator als ‚technische Vorgänge oder Naturprozesse wirklichkeitsgetreu (im Modell) nachzubilden‘ wurde so laut Pfeifer (1999) erst Mitte des 20. Jahrhunderts verwendet. Zumindest für das englischsprachige Äquivalent ‚simulation‘ kann eine frühere Benutzung nachgewiesen werden. So schildert La Marcus A. Thompson in seinem am 10.05.1896 eingereichten Patent einen von ihm entworfenen *Pleasure-Railway*, der die Vergnügungen einer Zugfahrt simulieren soll, ohne dass sich der Raum, der den Zug darstellen soll, bewegt. So schreibt Thompson über die dazu nötigen Dekorationen:

Preferably these decorations are made of papier-mâché in connection with rocks and plants and are formed so as to simulate natural scenery and elaborated by painting the walls or the canvas stretched on the same [...] (Thompson, 1896)

In diesem Zusammenhang ist es von Bedeutung zu erwähnen, dass die Praxis Simulationen im Lehrkräftebildungskontext durchzuführen bereits durch eine Schrift von Edwards (1865, zit. nach Cruickshank, 1988) zu belegen ist, freilich ohne dass sie als Simulation bezeichnet wurde:

Another essential requisite in a normal school is that it gives its pupils [hier sind zukünftige Lehrkräfte gemeint, Anm. d. Verf.]

an opportunity of some kind of practice in teaching under the supervision and subject of the criticism of experienced and skillful instructors. This is accomplished in several ways: by exercises in conducting the regular classes of the Normal Schools; by classes of normal pupils assuming for the time the character of children and receiving instruction and answering as they think children would; and by a separate school of children in which the novice is entrusted with the charge of the class (Edwards, 1865, S.280).

Was hier beschrieben wird, ähnelt dem heutigen Verständnis von *micro-teaching-peerteachings*.

Es ist jedoch richtig, dass der Begriff Simulation im modernen Verständnis erst in den 1930–1950er Jahren eine breitere Verwendung erfuhr, als die ersten Typen von Flug- und Radarsimulationen konstruiert und in größerem Umfang eingesetzt wurden. Einer der frühesten analogen Flugsimulatoren, der den Instrumentenflug simulierte, war der *Link Trainer*, den Ed Link 1929 entwickelte und den der U.S. Army Air Corps 1934 erstmals zu Schulungszwecken nutzte (Ward et al., 2006, S.252; Zazula et al., 2013, S.51; Page, 2000, S.3).

Abhängig von der Wahl der Simulationsdefinition war der erste digitale, computerisierte (Flug-)simulator entweder der ASCA oder der UDOFT. Der ASCA (Airplane Stability and Control Analyzer) wurde 1943, damals noch als analoges System, in Auftrag gegeben und 1950 am Servomechanics Laboratory des MIT konstruiert (Rolfe & Staples, 1986, S.33; Fischler, 2006, S.36). Im Gegensatz zum ASCA war der Universal Digital Operational Flight Trainer (UDOFT) ein *situational simulator* (Zur Definition siehe Abschnitt 3.4, S. 68), den die U.S. Navy im größeren Umfang zur Schulung einsetzte. Er wurde in den 1950er Jahren an der University of Pennsylvania konstruiert und bis in die 1960er Jahre weiter optimiert (Zazula et al., 2013, S.52). Zur weiteren Entwicklung der Flugsimulatoren sei an Rolfe und Staples (1986) verwiesen, welche die Geschichte der Flugsimulatoren ausgiebig beleuchten. Die erwähnten Simulatoren sind hier aufgeführt, da sie die ersten digitalen Simulatoren darstellen.

Die ersten explizit als solche bezeichneten Simulationen in der Lehrkräftebildung wurden Anfang der 1960er Jahre entwickelt (Brown, 1999, S.313). Die früheste Simulation ist wohl die von der Teacher Research Division des Oregon State System of Higher Education unter Leitung von Kersh (1963, 1963) entwickelte Simulation Mr. Land's Sixth Grade (Cruickshank, 1988, S.134; Brown, 1999, S.313; G. R. Girod et al., 2007, S.209). Zum Teil wurde die Simulation auch nur als Classroom Simulator (CS) bezeichnet (Kersh, 1962; Cruickshank, 1972, S.18). In dieser Simulation projizieren drei 16-mm

Projektoren lebensgroße Videoclips von zweiundzwanzig SuS einer fiktiven Grundschulklasse per Rückprojektion auf drei Wände, so dass der Eindruck einer realen Klasse entsteht. Zusätzlich regt Kersh (1962) an, den Raum mit zusätzlichen Einrichtungsgegenständen wie Tafel oder weiterer Möblierung zu dekorieren, um die Immersion der Partizipantinnen zu maximieren. Die Parallelen zur aktuellen Diskussion um VR (*virtual reality*) sind verblüffend: In dem Simulator kann man einen Urgroßvater sämtlicher situierter, hochimmersiver/high-fidelity Unterrichtssimulatoren erkennen. Die Interaktionsmodellierung war freilich recht basal. Es existierten sechzig einführende Ankervideos zu Klassenraumproblemen, welche zwei oder drei alternative Enden besaßen. Welches Ende ausgewählt wurde, bestimmten die von den Partizipantinnen durchgeführten Handlungen. Diese Handlungen waren aus einer Liste von maximal zehn möglichen Verhaltensweisen auszuwählen. Die Liste enthielt ebenfalls eine vordefinierte Zuordnung von Handlungsalternativen zu Szenarienausgängen. Die möglichen Verhaltensweisen und die Zuordnung zu Szenarienausgängen wurden durch ein Expertengremium von drei Lehrkräften normativ definiert (Cruickshank, 1988). Jede der Simulationssequenzen sollte bis zu zehn Mal wiederholt werden, bis die Partizipantinnen eine akzeptable Reaktion zeigen. Die paradigmatische Nähe zum operanten Konditionierung wird bei Cruickshank (1988, S.136f.) thematisiert.

Ausgehend von diesen Erfahrungen entwickelte Cruickshank mit Kollegen 1967 eine eigene Simulation, genannt *Teaching Problems Laboratory*, welches später in das *Inner-City Simulation Laboratory* mündete (Cruickshank, 1988). Lehramtsstudierende wurden mit 31 Problemen einer fiktiven fünften Klasse konfrontiert, zu denen in verschiedenen medialen Formen (16mm-Projektionen, Rollenspiel, schriftliche Dokumente) eine Szenarieneinführung gegeben wurde. Anschließend wurden schriftliche Reflexionsaufgaben mit Fokus auf Alternativhandlungskonstruktionen verteilt und in Kleingruppen diskutiert. Teilnehmerinnen wurden dazu ermuntert mögliche Alternativhandlungen im Rollenspiel zu erproben. Andere Simulationen dieser Zeit, wie die low-tech Simulation ‚Hey Teach‘ aus dem Jahr 1979, benutzten ausschließlich gedruckte Materialien (Brown, 1999, S.313). Ziel dieser speziellen Simulation war die Schulung administrativer Kompetenzen. Es besteht gewisse Ähnlichkeit zwischen den ersten beiden beschriebenen Simulationen und den analogen Flugsimulatoren der 1940er Jahre, welche Projektionen verwendeten (Vgl. Page, 2000, S.4). Auch wenn beide Simulationen aufgrund ihres Typus nicht vollständig miteinander vergleichbar sind, so lässt sich ein Technologievorsprung von etwa 15-20 Jahren zwischen Flugsimulationen und Simulationen in der Lehrkräftebildung bis in die heutige Zeit hin konstatieren. Solche Simulationen werden auch heute noch verwendet.

Von den 1960ern bis zum Ende der 1970er Jahre dominierte dieser Typus

von analogen, medienbasierten (Cruickshank, 1988, S.138) Simulationen die Lehrkräftebildung. Die zweite Phase des Simulationseinsatzes in der Lehrkräftebildung ordnen sowohl Brown (1999) als auch Cruickshank (1988) den Jahren 1980–1988 zu, mit einem proklamierten Paradigmenwechsel hin zur digitalen Simulation. Der vorgeschlagene Beginn dieser zweiten Phase ist diskussionswürdig, aber plausibel. Die dafür gewählte Begründung kann man hingegen kritisch beurteilen. Brown (1999, S.313) führt an, dass in den 1990 Jahren ein verstärktes Interesse für künstliche Intelligenz (KI) und *virtual reality* (VR) diese Forschungsepoche kennzeichnen. Während es richtig ist, dass die angeführten Forscher (VanLehn et al., 1994; VanLehn, 1996; Ur & VanLehn, 1995) wichtige praktische theoretisch-konzeptionelle Grundlagen zur Agentenmodellierung gelegt haben, so fand diese Arbeit doch einen verschwindend geringen Niederschlag in der Entwicklung und Forschungsliteratur von Simulationen in der Lehrkräftebildung. (Vgl. auch Turan, 2015). Ähnliches gilt für den Bereich der virtuellen Realität – diese wurde in den 1990er Jahren begeistert aufgenommen, aber mit Erwartungen überfrachtet. Die Technologie litt in der damaligen Umsetzung noch unter großen Kinderkrankheiten, so dass sie keine große Verbreitung und Anwendung fand (Kwon, 2019, S.101f.). In der Studie von Katz (1999, S.154) wurde erstmals die VR-Technologie der HMDs (*head-mounted displays*) im Lehrkräftebildungskontext verwendet. Ziel dieser Simulation war die Empathieschulung von Vorschullehrkräften. Das dafür benutzte Tool war, genauso wie die Applikationen von Stavroulia et al. (2018) und Bailenson et al. (2008), eine *virtual world/virtual environment* (siehe Abschnitt 2.5.3, S. 27 und Abschnitt 2.5.2, S. 26) und damit kein Simulator im engeren Sinne, da sich die Möglichkeiten der Interaktion auf das Sehen und Bewegen in der virtuellen Welt beschränkten. TeachLivE wird zwar als *virtual* bzw. *mixed reality* Simulator bezeichnet, ist jedoch rein bildschirmbasiert (Barmaki & Hughes, 2015, S.532ff. Stavroulia et al., 2019, S.196), so dass diese Selbstbezeichnung sehr davon abhängt, was man genau unter diesen Begriffen versteht. Die einzigen modernen Simulatoren, welche VR im Sinne der Verwendung von HMDs und modellierten Interaktionen verwenden, sind Breaking Bad Behaviors (Lugrin et al., 2016) und der VR-Klassenraum von Wiepke und Paulicke (2018).

Zur Einordnung des Beginns der zweiten Phase der Simulationsnutzung in der Lehrkräftebildung ist festzuhalten, dass die Mitte der 1980er Jahre tatsächlich einen paradigmatischen Wechsel hin zu digitalen Simulationen markiert. Allerdings waren diese Simulationen aufgrund der limitierten Rechenverarbeitungs- und Grafikkapazitäten der populären Heimcomputer wie Apple II, C64 und IBM PC in ihren Interaktionsmodellen und in ihrer Simulationsfidelität einfach strukturiert und sehr linear. Die vermutlich erste nicht nur rein textbasierte, graphische Simulation in der Lehrkräftebildung

dürfte die *Curry Teacher Simulation* gewesen sein, welche in ihrer ersten Iteration 1984 erschien und noch bis 1994 in einer für damalige Verhältnisse schon recht veralteten GUI beforscht wurde (Strang et al., 1989, S.147; Strang & Moore, 1994). Im Gegensatz zu den militärischen Simulationen gibt es in der Literatur keinen Hinweis, dass universitäre Großrechner, sogenannte Mainframes, für erziehungswissenschaftliche Simulatoren genutzt wurden. An dieser etwas randständigen Nutzung von digitalen Simulationen im Lehrkräftebildungskontext hat sich bis heute wenig geändert.

Der eigentliche Umbruch zur dritten Phase der Simulationsnutzung in der Lehrkräftebildung entstand Mitte der 2000er Jahre. In sehr kurzer Zeit entwickelten viele Forschungszentren unabhängig voneinander eigene Konzepte, die in ihrer Wirkung, Verbreitung, Interaktionsmodellierung und didaktischen Gestaltung sehr unterschiedlich waren. In der folgenden Aufzählung der wichtigsten digitalen Simulatoren bezeichnen die Schrägstriche immer andere Namen für denselben Basissimulator: SimSchool (2003), Lesson Sketch/Sim Teach/ThEMat (ab 2004), ClassSim (2004), SimTeacher (2006), STAR Simulator/TeachMe/Teach Live/Sim Lab Mursion (2007). Unter diesem Gesichtspunkt sind auch die Äußerungen von Fischler (2006, S.2) und Skrødal (2010, S.27) zu interpretieren, dass es zum damaligen Zeitpunkt noch keine guten *educational simulators* im engeren Sinne des Lehrkräftebildungskontextes gab. Noch 2016 konstatieren Kaufman und Ireland, dass die von ihnen analysierten Simulationen in der Lehrkräftebildung hauptsächlich Forschungsprototypen in experimentellen Settings waren.

Eine nähere Beschreibung der erwähnten Simulatoren wird im [Abschnitt 4](#) (S. 71) gegeben.

2.2 Definition des Begriffes Simulation

Eine wichtige Erkenntnis aus diesem historischen Abriss ist unter anderem, dass Simulationen keineswegs Computertechnologie benötigen (Skrødal, 2010, S.37; Dieker, Rodriguez et al., 2014, S.53). In bestimmten Bereichen wie der Medizin, Aviation und dem Militär wurden solche analogen Simulationen seit langer Zeit verwendet und werden in dieser Form zum Teil noch heute genutzt (Alinier, 2010; Dieker, Rodriguez et al., 2014, S.23; Kaufman & Ireland, 2016, S.260). In der umgangssprachlichen Verwendung des Begriffes Simulation versteht man heute allerdings oft rein digitale Simulationen, häufig mit der impliziten Annahme einer grafischen Oberfläche und mit terminologischer Nähe zu kommerziellen Simulationsspielen wie ‚The Sims‘, ‚SimCity‘, dem ‚Landwirtschafts-/Train-Simulator‘ oder ähnlichen Titeln. An dieser Stelle wird besonders deutlich, dass es notwendig ist, den Begriff der Simulation zu definieren und von anderen Begriffen wie Rollenspiel, *microteaching*/-

peerteaching, educational game, serious game und *virtual world* abzugrenzen (Vgl. auch Feinstein et al., 2002). Das Problem der inkongruenten Begriffsverwendung thematisiert auch die Metastudie von Turan (2015, S.36), weswegen es zentral ist einen Versuch zu unternehmen die Forschungstermini zu definieren.

Wie kann man den Begriff Simulation allgemeingültig definieren? Als Diskussionsgrundlage seien fünf Definitionen vorgestellt:

- (a) [A simulation is] the experimentation with a model in order to obtain information about the dynamic behavior of a system (Helsgaun, 2000, zit. nach. Skrødal, 2010, S.27).
- (b) A simulation is a simplified but accurate, valid, and dynamic model of reality implemented as a system (Sauvé et al., 2007).
- (c) Zwei Faktoren müssen vorhanden sein: „(1) there is a model of behavior based on a real or theoretical system and (2) the user can experiment by observing the consequences of actions“ (Fischler, 2006, S.37).
- (d) A simulation is defined as the abstraction or simplification of a situation or process (Brown, 1999, S.309). Eine ähnliche Definition wird in Deale und Pastore 2014 gegeben.
- (e) Statt eine kurze zusammenfassende Definition zu geben, schlägt Fischler eine Liste von einundzwanzig Charakteristika vor, die hier kurz vorgestellt werden sollen, für deren genaue Definition an Fischler (2006, S.38ff.) verwiesen sei.
 - (1) Learn by doing
 - (2) Learn from Mistakes
 - (3) Surprises
 - (4) Risk Free
 - (5) Time Compression
 - (6) High Fidelity
 - (7) Interaction Effect
 - (8) Meaningful Learning
 - (9) Applied Learning
 - (10) Learner-Centered
 - (11) Permeate Reality

- (12) Post-Analysis
- (13) Authentic Assessment
- (14) Dynamic Database
- (15) Progressive Complexity
- (16) Diminishing Assistance
- (17) Cognitive Mentoring
- (18) Expert Advice
- (19) Resource Library
- (20) Active Participation
- (21) Visualizing Impossibilities

In den Definitionen (a)-(c) findet sich explizit der zentrale Begriff ‚Model‘, in der Definition (d) wird implizit darauf verwiesen, indem von Abstraktion oder Simplifizierung eines Prozesses oder einer Situation gesprochen wird. Neben der Notwendigkeit ein System zu abstrahieren und es als Modell mit lernrelevanten Faktoren zu implementieren ist der zweite zentrale Bestandteil aller Definitionen die Möglichkeit mit diesem Modell zu experimentieren, d.h. durch Manipulation der Eingangsvariablen ein tieferes Verständnis über das System zu gewinnen (Definition (a), (c), Vgl. auch Deale und Pastore 2014, S.198).

Das Simulationen zugrunde liegende Lernverständnis ist ein situiertes (Vgl. G. Blomberg et al., 2014, für eine ähnliche Diskussion für den Bereich Video; Galarneau & Zibit, 2007, für den Bereich Spiele) – und zwar unabhängig von der Einordnung der einzelnen Simulationen in ein übergeordnetes Forschungsparadigma wie Behaviorismus, Kognitivismus oder Konstruktivismus. Es wird angenommen, dass sich Lernen mit Simulationen durch die inhärente Handlungsnähe und aktive Rolle der Partizipantinnen von Lernprozessen klassischer linearer Medien wie Bücher und Filme fundamental unterscheidet (Aldrich, 2009, S.xxii). Letzteren wird zwar eine Bedeutung für die Entwicklung deklarativen Wissens zugeschrieben, jedoch sei handlungsnahes oder prozedurales Wissen durch Simulationen oder Hospitationen besser zu erlernen (Chini et al., 2016, S.2; Aldrich, 2009, S.xxxi). Dieses handlungsnahes Wissen sei der Königsweg zu den „big skills“ oder auch „21st century skills“, die die angeblich zentralen Fähigkeiten wie Führung, Projektmanagement und Innovation beinhalten (Aldrich, 2009, S.xxviii; Kallonis & Sampson, 2010, S.42). Vor dieser, mit großem Pathos vorgetragenen Glorifizierung des Mediums sei eindringlich gewarnt. Nicht nur deswegen, weil die emotionalisierende Berichterstattung in den Medien den Forschungsdiskurs

verwässert und ungut an die Diskussion um *edutainment*, inklusive Ernüchterung, erinnert, sondern auch, weil dadurch wieder über Medien und nicht über die eigentlich wichtigen didaktischen Settings diskutiert wird (Akilli, 2007, S.2). Spricht Aldrich etwa davon, dass der – ohne Zweifel einflussreiche – Spieleentwickler Will Wright ein „Beethoven“ oder „Shakespeare“ des digitalen Mediums sei (Aldrich, 2009, S.xxxi) und er Bücher hingegen als „Weißbrot für den Geist“ bezeichnet (Aldrich, 2009, S.46), so leistet Aldrich einen Beitrag dazu überzogene Erwartungen an das Medium per se zu stellen – auch wenn er seine Aussagen später etwas relativiert.

Trotz dieses immer wieder anzutreffenden Pathos herrscht ein Forschungskonsens darüber, dass Simulationen nur ein Werkzeug sind, die traditionelle Lernkontexte und Hospitationen nicht substituieren, sondern bereichern sollen (Aldrich, 2009, S. xxv; Bradley & Kendall, 2014, S.3, S.9; Dieker et al., 2017, S.78; Fischler, 2006, S.2; Brown, 1999, S.316; Cheong & Kim, 2009, S.72f.). Dies ist insbesondere unter der Perspektive zu betrachten, dass reale Felderfahrungen in Schulklassen primär auf die Planung der Stunden fokussieren (Makroebene) und weniger das interaktionale Geschehen in den Blick nähmen (Meso- und Mikroebene) (Bradley & Kendall, 2014, S.3).

Simulationen und Simulatoren werden in der Literatur oft unspezifisch und synonym zueinander verwendet. Es macht aber durchaus Sinn diese Begriffe semantisch voneinander zu trennen; während eine Simulation auf den prozessual-didaktischen Aspekt des Umgangs abzielt, bezeichnet ein Simulator ein technisches Artefakt. Dieses Artefakt kann eine systemisch logische Verkettung von Programmcode, 2D oder 3D Objekten und Eingabemedien wie Maus/Tastatur/Motion-Controller sein, jedoch auch eine systemisch logische Verkettung analoger Komponenten wie schriftliche Fallschilderungen oder Charakterbeschreibungen für ein Rollenspiel. Eine Simulation nimmt demzufolge nicht nur das technische Artefakt Simulator in den Blick, sondern auch die didaktische Vor- und Nachbereitung und die Art und Weise wie der Simulator genutzt wird. Da ein Simulator ohne didaktisches Rahmenmodell in der Diskussion um Lernen mit Simulationen nur einen eingeschränkten Wert besitzt, wird in dieser Arbeit primär der Begriff Simulation genutzt und der Begriff ‚Simulator‘ nur dann verwendet, wenn er auf die technische Umsetzung rekurriert.

Kritisch am Begriff Simulation ist zu sehen, dass er aufgrund der Nähe zu mathematisch-naturwissenschaftlichen Simulationen wie der Wettersimulation eine große Assoziation mit einer hohen Prädiktorengenauigkeit aufweist, die im Bildungskontext so nicht gegeben ist (Aldrich, 2009, S.xxxii). Ein mathematisch-naturwissenschaftlicher Simulator kann zwar mit passender didaktischer Rahmung auch zu einer Lerngelegenheit werden, allerdings wurde in dieser Simulatorform keine indigene instruktionale Komponente angelegt

(Aldrich, 2009, S.xxxii). Die alternative Bezeichnung, die Aldrich vorschlägt und auch selbst benutzt, nämlich schlicht *Sims*, erweitert den Begriffsraum, ist aber sprachlich aufgrund der Nähe zum Begriff ‚Simulation‘ und zu einem bekannten Spielefranchise unglücklich gewählt und im Rahmen dieser Arbeit zu offen. Aldrich versteht unter *Sims* sowohl alle Arten von Spielen (Brettspiele, Rollenspiele, Computerspiele) als auch immersive Lernsimulationen, denen er *serious games* und *educational simulators* unterordnet. Unter *educational simulators* versteht Aldrich in etwa das, was in dieser Arbeit unter Simulationen subsumiert wird (Aldrich, 2009, S.9). Wie bereits in der Unterscheidung von Simulation und Simulator angedeutet, ist das erfolgreiche Lernen innerhalb einer Simulation von deren didaktischer Rahmung abhängig (Hixon & So, 2009, S.300). Allein die verbrachte Zeit im Simulator ist kaum entscheidend für den Lernerfolg (Ward et al., 2006, S.252, S.258).

Die Feedbackkomponenten der Simulation seien insbesondere in situationalen Simulationen wie einer Klassenraumsimulation von zentraler Bedeutung (Badiee & Kaufman, 2015, S.2; Brown, 1999, S.314; Alessi & Trollip, 2001, S.245, S.260; Fischler, 2006, S.66; Gibson, 2009a, S.693; empirische Belege Dieker et al., 2017, S.73). Feedbackkomponenten können Hilfestellungen, Korrekturen, Coaching und Auswahl von Aufgaben oder einer Aufgabenkette sein (Alessi & Trollip, 2001, S.245). Laut Dieker, Rodriguez et al. (2014) muss es zur optimalen Nutzung einer Simulation einen zyklischen Prozess aus Durchspielen von Handlungsalternativen, Feedback, Reflexion und Modifizierung der Handlungsweisen geben (Dieker, Rodriguez et al., 2014, S.23; Badiee & Kaufman, 2015, S.3; Kaufman & Ireland, 2016, S.261; Anwendung findet dies auch bei Chini et al., 2016, S.6). Dieses Konzept teilt die Simulationsnutzung in drei Teile auf: (a) Eine Komponente zur Vorbereitung auf die Nutzung des Simulators mit Fokus auf Lernzieldefinition und einer präemptiven Handlungsreflexion, genannt BAR – Before Action Review, (b) die eigentliche Nutzung des Simulators (A – Action) und (c) die Nachbereitung im Sinne eines Feedbacks zu den im Simulator durchgeführten Handlungen (ARR – After Action Review). Es gibt einige Hinweise darauf, dass gerade die Feedbackgespräche nach Simulatornutzung (ARR) für die Vermittlung handlungsnahen Wissens zentral sind (Fischler, 2006, S.141; Straub et al., 2014). Werden keine Feedbackgespräche durchgeführt, so lassen sich nur geringe positive bis keine Effekte auf handlungsnahen Kompetenzen beobachten. Eine interessante neue Komponente ist die videographische Aufzeichnung von Mimik und Gesten während der Simulationsnutzung, die anschließend angesehen und über die reflektiert werden kann (Delamarre et al., 2017, S.401).

2.3 Abgrenzung zu (*serious*) games

Nicht alle Simulationen sind Spiele, es existiert aber ein ganzes (Computer-) Spielgenre, das unter dem Label „*simulation games*“ firmiert. Auch werden Simulationen mit definierter Zielgruppe und didaktischer Rahmung verwirrenderweise in der Literatur als *games* bezeichnet. Letztlich enthalten Simulationen, die explizit für den Bildungskontext entwickelt wurden, oft Spielelemente oder -mechanismen. Daher erscheint es geboten, einen Versuch zu unternehmen die Begriffe (*computer*) *games* und *serious games* zu definieren und sie semantisch von dem Begriff der Simulation zu trennen. Der Begriff *game* ist im Kontext einer Abgrenzung zur Simulation viel zu offen und divers (Aldrich, 2009, S.xxxii). Als Spiele werden nicht nur Brettspiele, Rollenspiele und Computerspiele bezeichnet, sondern der Begriff bezeichnet viele Aktivitäten, die zum Zwecke des Zeitvertreibes durchgeführt werden. Da es ein recht kompliziertes Unterfangen wäre, den Begriff *games* in all seinen Facetten adäquat zu beleuchten, soll sich an dieser Stelle darauf konzentriert werden, aus welcher Definition eine Abgrenzung zum Begriff Simulation gewonnen werden kann. Brown (1999, S.309) hingegen definiert den Begriff *game* sogar so weit, dass er Simulationen miteinschließt. Brown bezeichnet *games* als „situations in which participants follow a set of rules to achieve a specified goal“. Auch für Cruickshank (1972, S.18f.) gibt es keine klare Distinktion zwischen Spielen und Simulationen, er mutmaßt, dass diese auch in der Literatur deswegen nicht gemacht wird, weil sich Simulationen und Spiele mehr im Grad der Modellfidelität als in ihrer grundsätzlichen Art unterscheiden. Spiele sind für Cruickshank demzufolge auf übersimplifizierenden Modellen oder Regeln basierend. Fischler (2006, S.29ff.) arbeitet sich in seiner Dissertationsschrift am gleichen Thema ab und präsentiert zwei Distinktionsmerkmale. Zum einen sei Ziel einer *educational simulation* in erster Linie eine didaktisch durchdachte Lerngelegenheit bereitzustellen, während das Ziel einer *video-game simulation* primär die Unterhaltung ist. Dieser Abgrenzung ist zuzustimmen (Zum selben Ergebnis kommt auch Akilli, 2007, S.3f.); außerdem legt die Abgrenzung den Finger in die Wunde des Problems, dass, sobald off-the-shelf Spiele in einem formalen Lernkontext verwendet werden, ihnen zum Teil didaktische Konzepte übergestülpt werden, welche die Spieledesignerinnen nicht antizipiert haben (Fischler, 2006, S.29f.). Auch wenn es gegenteilige Beispiele gibt, so zeigt sich häufig eine mangelhafte Passung zwischen didaktischem Ziel und Spielmechanik. Dem zweiten von Fischler (2006, S.29ff.) vorgeschlagenen Distinktionsmerkmal ist nicht so umfänglich zuzustimmen: Er behauptet, dass *video-game simulations* ein *fantasy setting* bedingen würden, während *educational simulations* ein realistisches Setting erfordern. Viele *video-game simulations* bedienen sich realistischer Settings

(z.B. Wirtschafts- und Transportkreisläufe, Stadtbausimulationen), während es auch einige *educational simulations* gibt, welche in einem *fantasy setting* operieren. In der Diskussion um die *virtual worlds* wird als Stärke der Simulation angepriesen, dass man Experimente durchführen kann, die in der Realität so nicht möglich wären. Es ist allerdings richtig, dass im speziellen Fall der *classroom simulations* ein *fantasy setting* eine Ausnahme darstellt. Das zweite wichtige Unterscheidungsmerkmal zwischen Spielen und Simulationen benennen Baidee und Kaufman treffend:

Simulations are distinguished from games in that they do not involve explicit competition; instead of trying to „win,“ simulation participants take on roles, try out actions, see the results, and try new actions without causing real-life harm (Badiee & Kaufman, 2015, S.1).

Der zentrale Aspekt dieser Unterscheidung, nämlich die Kompetitivität von Spielen, wird ebenfalls bei Alessi und Trollip (2001, S.277f.), Cruickshank (1972, S.18), bei Kaufman und Ireland (2016, S.261) und bei Akilli (2007, S.4) erwähnt. Während es in Spielen also primär darum geht in unterhaltsamer Weise andere Mitspielerinnen, zu denen auch der Computer zählen kann, zu besiegen oder eine möglichst hohe Punktzahl zu erreichen, geht es bei *educational simulations* vor allem um das Lernen von Inhalten, welche die Erstellerinnen der Simulation in der Konstruktion derselben berücksichtigt haben.

Der Begriff *serious game* lässt sich mit ‚ernsthafte Spiel‘ übersetzen und deutet ein Wortspiel mit vermeintlichen Oxymora an (Djaouti et al., 2011, S.25f.). Die erste Verwendung des Begriffes wird häufig auf Sawyer und Resjki (2002) zurückgeführt, welche den Begriff einzig im Titel verwenden. Tatsächlich nutze Abt (1970) den Begriff aber schon in einem ähnlichen Bedeutungsumfang (Djaouti et al., 2011, S.26). Für eine ausführliche Begriffs-exegese sei an Djaouti et al. (2011) verwiesen.

Im Gegensatz zum neutralen Begriff Simulation setzt der Terminus *serious game* einen assoziativen Schwerpunkt auf die für Bildungsinstitutionen kontroversesten Teile: Spiel und Spaß (Aldrich, 2009, S.xxiii). Lerntheoretisch legt die Verwendung dieses Begriffes eine Nähe zu *game-based learning* (GBL) oder *game-informed learning* nahe (Fischler, 2006, S.28ff.). Aldrich sieht *serious games* als Nachfolger des mittlerweile aus der Mode gekommenen Begriffes *edutainment*. (Aldrich, 2009, S.xxxiii). Ähnlich betrachten auch Höntzsch et al. (2011) den Begriff *serious games*, dem sie sowohl Simulationen, *edutainment* als auch Werbespiele zuordnen. Alessi und Trollip (2001, S.270) hingegen betrachten beide Begriffe eher als Marketinginstrumente und fassen sie unter dem Schlagwort *educational games* zusammen. Dennoch hat

sich im allgemeinen und medialen Sprachgebrauch eine Distinktion sowohl zwischen den Begriffen *serious games* und *edutainment* gebildet.

Der Neologismus *edutainment* setzt sich aus den Begriffen *education* und *entertainment* zusammen. In einer weiten Definition werden damit demzufolge alle medialen Produkte bezeichnet, die Bildung mit einer Form von Unterhaltung verbinden. Die Verwendung des Begriffes in diesem Sinne hat nach dem Ende der Euphoriewelle um *edutainment* Mitte der 2000er Jahre stark abgenommen. In der engeren Definition assoziiert man *edutainment* mit einer bestimmten Form von Lernsoftware, oft mit der Zielsetzung deklaratives Wissen zu erzeugen. Das Lehr-/Lernparadigma dieser Art von Software ist oft ein vor-theoretisches, häufig beeinflusst von den neusten technologischen Möglichkeiten Wissen aufzubereiten. So war Ende der 1990er Jahre das Buzzword der multimedialen Aufbereitung in aller Munde. Die didaktische Umsetzung in den erfolgreichen kommerziellen Produkten war oft von zweifelhafter Natur und fußte auf einem implizit-intransparenten behaviouristischen Lernparadigma. Die Ernüchterung, nachdem die hohen Erwartungen an *edutainment* nicht erfüllt werden konnten, folgte rasch.

Serious games hingegen sind stark mit dem gesteigerten Bewusstsein für realweltliche Probleme (Aldrich, 2009, S.26) Anfang der 2000er Jahre verknüpft. Denkt man an die prototypische Produzentin eines *serious game*, so assoziiert man damit NGOs, Regierungsorganisationen und idealistische Indie-Studios. Ziel dieser Spiele ist es in erster Linie den Spielerinnen eine Lerngelegenheit zu bieten, die im Unterschied zu gewöhnlichen Computerspielen bereits in ihrer instruktional-didaktischen Anlage von den Macherinnen direkt implementiert oder zumindest antizipiert wurde (Fischler, 2006, S.28; Höntzsch et al., 2011, S.3). Dies haben *serious games* mit Simulationen gemeinsam (Alessi & Trollip, 2001, S.215). Lerntheoretisch bemühen sich diese Spiele um etwas, dass Aldrich (2009, S.26) als „*stealth learning*“ bezeichnet. Ohne die didaktischen Ziele vollständig offenzulegen oder prominent anzusprechen soll Lernen quasi unbemerkt geschehen, ganz im Gegensatz zum *edutainment*, welches auf eine klar ausdefinierte *high-stakes* Feedbackkomponente setzt. Der Transfer von deklarativem Wissen spielt in *serious games* oftmals eine untergeordnete Rolle, die zu vermittelnden Wissensdimensionen konzentrieren sich auf konzeptuelles und metakognitives Wissen. *Serious games* haben den Anspruch dieses Wissen ohne umfangreiches normatives Feedback zu vermitteln. Diese Merkmale sollen zur Veranschaulichung an einem Beispiel illustriert werden. Aldrich nennt als Beispiel für das ‚prototypische serious game‘ SimCity (Aldrich, 2009, S.29). Diese Zuordnung dürfte einige Fragen aufwerfen, die sich nur dann etwas auflösen, wenn man versteht, dass Aldrich eine andere Definition von *serious games* ins Feld führt, die sich eher an Cruickshank (1972) orientiert. Für ihn ist SimCi-

ty nämlich deswegen keine Simulation, sondern ein *serious game*, da „die zugrundeliegende Mechanik des Spieles zu ungenau zur Modellierung einer realen Stadt ist“ (Aldrich, 2009, S.29). Diese Definition ist keineswegs unbrauchbar, aber sie deckt nur einen Aspekt der *serious games* ab. Da die hier vorgeschlagenen Merkmale den Begriff anders zu fassen versuchen, soll in dieser Arbeit auch ein anderes Beispiel verwendet werden, nämlich das *serious game* ‚Papers, Please‘. In ‚Papers, Please‘ übernimmt die Spielerin die Kontrolle eines Grenzsoldaten im fiktiven, totalitären Staat Arstotzka. Das Spiel wurde von Lucas Pope entwickelt, 2013 veröffentlicht – unter anderem auf der populären kommerziellen Spieleplattform Steam – und lehnt sich grafisch an die niedrig aufgelöste EGA/VGA Grafik Ende der 1980er/Anfang der 1990er Jahre an, übernimmt jedoch eine düstere Farbpalette. Die Aufgabe der Spielerin ist es die Dokumente der Einreisenden zu kontrollieren und diesen bei gefälschten Dokumenten den Zutritt zu verwehren. Dabei wird die Spielerin mehrfach vor moralische Dilemmata gestellt: So kann sie beispielsweise eine Reisende mit falschem Pass mit ihrer Familie vereinen, muss jedoch mit Repressionen des Staates rechnen. Trifft man zu oft inkorrekte Entscheidungen, so gerät man in den Blick der Geheimpolizei; das eigene Gehalt wird gekürzt und man kann nicht mehr genug Nahrung und Medikamente für die eigene Familie kaufen. Der wichtigste Designaspekt ist, dass das Spiel nie direkte normative Bewertungen der Spielerinnenhandlungen durchführt, so dass die Spielerin lernt von den Handlungskonsequenzen im Spiel eigene Schlussfolgerungen zu ziehen. Deswegen ist das Spiel ‚Papers, Please‘ in gewissem Maße auch ein Prototyp des (erfolgreichen) *serious game*: Es behandelt ein eher gesellschaftlich-komplexes Problem (Verhalten und Moral in diktatorischen Staaten, Umgang mit Überwachung), liefert jedoch keine einfache Lösung, sondern lässt die Spielerin selbst überlegen. ‚Papers, Please‘ bietet zudem eine an kommerzielle Spiele angelehnte Spielmechanik (Kontrollieren der Dokumente), die in sich selbst herausfordernd ist. Dass das Spiel kommerziellen Erfolg hatte, ist für *serious games* eher die Ausnahme, bestätigt aber die Annahme, dass hier besonders geschickt Spielmechanik, Lernen und Unterhaltung miteinander verwoben wurden. Warum ist ‚Papers, Please‘ nun keine Simulation? Weder die kommerzielle Nutzung, noch das fiktive Szenario sind ausschlaggebend, sondern das Ziel der Simulation. ‚Papers, Please‘ hat nicht zum Ziel die Zoll- oder Grenzkontrolle beispielsweise zur Schulungszwecken abzubilden, sondern nutzt den Arbeitsalltag der Grenzbeamtin lediglich als Mittel zum Zweck, um Denkprozesse zu den übergeordneten gesellschaftlichen Problemen anzuregen (Vgl. Aldrich, 2009, S.31). Letztendlich muss aber zugestanden werden, dass die Abgrenzung von Simulationen und *serious games* nicht ganz trennscharf ist (Vgl. Delamarre et al., 2017) und Behelfskonstruktionen wie die Fidelität des Simulationsmodells dabei helfen

können eine Software dem ein oder anderen Begriff zuzuordnen.

2.4 Vorteile und Gründe für die Nutzung von Simulationen

Simulationen wurden seit der ersten fachbezogenen Nutzung von Seiten der Bildungsforschung großes Potential zugeschrieben (Kersh, 1962; Cruickshank, 1969), obwohl nie dauerhaft im Curriculum der Lehramtsausbildung implementiert (Clayton & Gizelis, 2005). Die Potentiale waren primär aus theoretischen Überlegungen abgeleitet und werden im Zuge der gesamtgesellschaftlichen Digitalisierung heute wieder verstärkt im Forschungsdiskurs rezipiert (Billingsley & Scheuermann, 2014, S.268; Gibson, 2012a, S.3). Zunächst sollen hier die sich aus theoretischen Überlegungen ergebenden Vorteile gesammelt und diskutiert werden, anschließend soll auf erste empirische Studien eingegangen werden, welche sich mit der Wirksamkeit von Simulationslernen mit Fokus auf den Lehrkräftebildungskontext oder anverwandten Fachfeldern beschäftigen. Die Clusterung zu verschiedenen Argumentationssträngen wurde vom Autor entwickelt.

(a) Ethisches Argument

(Badiee & Kaufman, 2015, S.2; Bradley & Kendall, 2014, S.4; Im Bereich Chirurgie: Ward et al., 2006, S.250; Brown, 1999, S.307; Chini et al., 2016, S.1; Christensen et al., 2011, S.215; Foley & McAllister, 2005, S.162; Cruickshank, 1969, S.24; Höntzsch et al., 2011, S.7; Kaufman & Ireland, 2016, S.261f. Hixon & So, 2009; Ferry et al., 2004, S.296; Diefenbach et al., 2019, S.196)

Das ethische Argument ist eines der am häufigsten genannten Argumente und zielt darauf ab, dass in Simulationen keine realen SuS suboptimalen Unterrichtsversuchen ausgesetzt werden. Genauso wie der Umgang mit Patientinnen im medizinischen Kontext sei der Umgang mit SuS ein Bereich, der höchste Sensibilität erfordere und besonderen gesetzlichen Schutzanforderungen genügen muss (In ihrer Eigenschaft als SuS z.B.: HSchG §3, Abs. 9, §82, §82a; In ihrer Eigenschaft als Jugendliche JuSchG).

Der Simulator bietet eine sichere Umgebung, in der man verschiedene Unterrichtsstile ausprobieren kann, sogar welche, von denen man annehmen könnte, dass sie weniger gut funktionieren. Da virtuelle SuS (auch Agenten genannt) nicht müde, verwirrt oder frustriert werden, kann man als Simulationspartizipantin ungestört und ohne ethische Barrieren, die in realen *microteachings* oder Hospitationen vorhanden sind, Handlungsalternativen durchspielen.

Gerade der Begriff Sicherheit wird von zahlreichen Autorinnen im Zusammenhang mit diesem Argument genannt (Alessi & Trollip, 2001, S.227; Fischler, 2006, S.36; de Jong et al., 2012, S.36; Sawchuk, 2011; Deale & Pastore, 2014, S.199, 204; Hixon & So, 2009; Dieker et al., 2007, S.3, 5; Dieker et al., 2008, S.4, 6; Dieker, Rodriguez et al., 2014, S.25; Dieker, Kennedy et al., 2014, S.51; Gibson, 2012a, S.2).

Ein tendenziell kaum erwähnter aber nichtsdestotrotz wichtiger Teilaspekt des ethischen Arguments ist es, eine mögliche empathische Funktion von Simulationen in den Blick zu nehmen. In Simulationen ist es möglich verschiedene Rollen einzunehmen und die Welt aus einer anderen Perspektive zu betrachten (Bradley & Kendall, 2014, S.4; de Jong et al., 2012, S.39). Dies ist nicht auf das Hineinversetzen in SuS begrenzt. Auch ermöglichen es Simulationen, dass beliebig viele Beobachter als nicht-sichtbare Geisteravatare die Unterrichtsversuche von anderen beobachten und sogar Feedback geben können, ohne in die soziale Realität der Situation durch den Invasivitätseffekt der puren Anwesenheit einzugreifen. Tatsächlich wurden diese Möglichkeiten selten erprobt; als erwähnenswerte Ausnahme seien hier die Studien von Katz (1999) und Stavroulia et al. (2018) genannt.

(b) **Administrativ-Ökonomisches Argument**

Nennung der administrativen Komponente : (Billingsley & Scheuermann, 2014, S.256f. Badiee & Kaufman, 2015, S.2; Deale & Pastore, 2014, S.204; Ferry & Kervin, 2007, S.190; Dieker et al., 2007, S.2; Ferry et al., 2005, S.85; Kersh, 1962; Cruickshank, 1969, S.25; Höntzsch et al., 2011, S.7; Kaufman & Ireland, 2016, S.261; Kervin et al., 2004, S.4; Gundel et al., 2019, S.248)

Nennung der ökonomischen Komponente: (Badiee & Kaufman, 2015, S.3; Bradley & Kendall, 2014, S.4; Im Bereich Flugsimulation Ward et al., 2006, S.253; Bulu, 2012, S.154; Alessi & Trollip, 2001, S.6, 228; Gibson, 2012a, S.2; Cruickshank, 1969, S.26; Höntzsch et al., 2011, S.7)

Die Nutzung von Simulationen kann administrative Entlastungen mit sich bringen und Lerngelegenheiten ermöglichen, die sonst nicht durchzuführen sind. Hospitationen oder *microteachings* mit realen SuS benötigen beispielsweise Lehrpersonal, Absprache mit Schulen, Studienseminaren, Universitäten, pädagogischen Hochschulen und Kultusministerien. Schulen können nur bis zu gewissen Grenzen Kapazitäten für die Durchführung solcher Projekte schaffen; sie sind gesetzlich dazu verpflichtet, dass der reguläre Unterricht nicht beeinträchtigt werden darf. Letztlich ist es nur möglich zu bestimmten Zeiten und für die Partizipantinnen einmalige oder zeitlich begrenzte Felderfahrungen zu

machen. Billingsley und Scheuermann (2014, S.256) weisen in diesem Kontext auch darauf hin, dass ein oder zwei Praxiserfahrungen für Sonderpädagoginnen nicht ausreichend sind, um sich optimal auf die anschließende Ausbildungsphase vorzubereiten. Alessi und Trollip (2001, S.228) führen dieses Argument generalisiert ins Feld: Es würde zu wenig gelernt, Theorie und Praxis miteinander zu verbinden. Simulationen seien hier eine Alternative, die mit wenig Aufwand betrieben werden könne, wenn sie erst einmal entwickelt sei. Dieses Argument ist jedoch nicht vollumfänglich auf alle Simulationen anzuwenden: Je größer der Anteil von erforderlicher menschlicher didaktisch-instruktionaler Interaktion, besonders in Simulationen mit der Interaktionsmodellierung des Types 6, 7 und 8 (siehe [Abschnitt 3.3](#), S. S. 54), umso mehr administrativer und finanzieller Aufwand wird generiert. Simulationen, die auf diese Art von Interaktionsmodellierung setzen, zeigen ebenso wie analoge Rollenspiele oder *microteachings* keine oder zumindest weniger starke positive Skalierungseffekte (*economies of scale*) (Aldrich, 2009, S.10). Eine flächendeckende Anwendung über große Zahlen an Studierenden ist innerhalb dieser Interaktionsmodellierungen kaum möglich.

Gleichzeitig steht der administrativ-ökonomische Arbeitsaufwand auch mit der gewählten Simulationsfidelität in einem positiven linearen Zusammenhang, d.h. je komplexer und lebensechter eine Simulation wird, umso mehr Wartungs- und Verwaltungsaufwand entsteht und umso eingeschränkter werden die Möglichkeiten den Simulator jederzeit zu nutzen. Außerdem steigen die Kosten für die Implementierung des Simulators. Ersichtlich wird das am Beispiel eines Flugsimulators. Wird als Hardware nur ein Mobiltelefon verwendet, ist die Simulationsfidelität nicht sehr hoch, jedoch ist heutzutage fast jeder in der Lage den Simulator unabhängig von Uhrzeit und Ort zu nutzen. Die Entwicklung ist vergleichbar kostengünstig. Steigert man die Fidelität, indem man die erweiterten graphischen Fähigkeiten und die Möglichkeiten von Peripherieanschlüssen wie Steuerknüppel eines PCs nutzt, so muss sichergestellt werden, dass die benötigte Prozessor- und Grafikleistung, sowie die erforderlichen Peripheriegeräte bereitgestellt werden. PC-Pools müssen eingerichtet oder Entleihkonzepte entwickelt und durchgeführt werden. Die Entwicklung von Texturen, Shadern, 3D-Modellen und verbesserten Interaktionsmodellierungen nimmt erheblichen Umfang ein und erfordert große finanzielle und programmiertechnische Anstrengungen, um eine akzeptable Qualität gewährleisten zu können. In Simulationen mit extrem hoher Fidelität werden voll ausgestattete Cockpits auf mechanisierten Bewegungssystemen installiert. Ein solches System

ist nur mittels festgelegter Übungszeiten, regelmäßiger Wartung und stationär zu betreiben, so dass einige der angesprochenen Vorteile wegfallen. Im Falle eines Flugsimulators ist es dennoch ersichtlich, dass auch ein solcher Simulator mit geringeren administrativ-ökonomischen Kosten verbunden ist wie ein realer Flug. Im Bereich der Klassenraum-simulation ist dies jedoch nicht zwangsläufig gegeben, so dass neben inhaltlichen und technischen Gründen auch die administrative Komponente bei der Implementierung solcher Simulationen miteinbezogen werden muss.

Inhaltlich verwandt mit dem administrativen Argument ist das Effizienz-Argument. Man geht davon aus, dass Simulationen gegenüber traditionellen Unterrichtsmethoden wie Vorlesung und Seminar, unter anderem aufgrund der direkten Feedbackkomponenten, effektiver seien und es daher sinnvoller sei, diese einzusetzen: So spare man administrativ bereitgestellte Lernzeit ein (Alessi & Trollip, 2001, S.230). Dieses Argument ist noch nicht hinreichend empirisch validiert und geht problematischerweise von einem Medienvergleich aus, der die didaktische Rahmung unberücksichtigt lässt. Näheres zu den empirischen Ergebnissen folgt später in diesem Kapitel.

(c) **Repeatability Argument**

(Badiee & Kaufman, 2015, S.2f. Billingsley & Scheuermann, 2014, S.257; Brown, 1999, S.308; Alessi & Trollip, 2001, S.228; Chini et al., 2016, S.2f. Andreasen & Haciomeroglu, 2009, S.1318; Dieker et al., 2008, S.6; Dieker, Rodriguez et al., 2014, S.29; Dieker, Kennedy et al., 2014, S.51; Ferry et al., 2004, S.295 [nur sehr kurz]; Gibson, 2012a, S.2 [nur sehr kurz]; Kaufman & Ireland, 2016, S.261)

Simulationen ermöglichen es ein Szenario mehrfach zu gleichen Anfangsbedingungen durchzuspielen und alternative Handlungsoptionen auszuprobieren. Diese Wiederholbarkeit kann genutzt werden, um viele verschiedene Unterrichtsstrategien (ethisch) zu erproben oder gleiche Unterrichtstrategien auf andere SuS anzuwenden. Zudem kann ein Szenario so lange wiederholt werden bis das instruktionale Ziel erreicht wurde. Zuletzt ist es möglich die Simulation an beliebigen Zeitpunkten zu stoppen und ggf. direktes Feedback zu geben. In realen Felderfahrungen ist dies nicht oder nur eingeschränkt mit hohem technischem Aufwand (z.B. Knopf im Ohr) möglich.

(d) **Controlled Exposure Argument**

(Badiee & Kaufman, 2015, S.2; Bradley & Kendall, 2014, S.3f. Billingsley & Scheuermann, 2014, S.257, 269; Chieu & Herbst, 2011, S.106;

Alessi & Trollip, 2001, S.227; Dieker et al., 2008, S.5; Dieker, Rodriguez et al., 2014, S.28; Höntzsch et al., 2011, S.7)

In einer Simulation können sowohl das abgebildete Szenario, als auch sämtliche modellierten Variablen gezielt gesteuert werden. So ist es möglich unwahrscheinliche, aber mögliche Szenarien zu simulieren, die angehenden Lehrkräften so während einer realen Felderfahrung in der Regel selten begegnen. Beispiele hierfür wären der Umgang mit körperlich-aggressivem Störungsverhalten oder mit bestimmten Lern- und Entwicklungsstörungen. So können Simulationen hilfreich sein, um angehende Lehrkräfte schon früh praktisch für Inklusionsthemen zu sensibilisieren. Viele Studierende, die reale Felderfahrungen wie Praktika oder Hospitationen machen, stehen unter einem hohen *cognitive load* (Sweller, 1988). Oft haben die Studierenden in Schulhospitationen Schwierigkeiten Lerngelegenheiten während des Unterrichts zu identifizieren; sie sind schon damit überfordert visuell auf relevante Cues zu fokussieren (Ferry et al., 2005, S.89; Fischler, 2006, S.38). Agieren Sie selbst aktiv, so verkompliziert Zeitdruck und affektiver Stress den Entscheidungsfindungsprozess der Lehramtsstudierenden (Brown, 1999, S.314; Deale & Pastore, 2014, S.203). Weiterhin gibt es oft keine direkte Möglichkeit die eigenen Handlungen zu reflektieren. In Hospitationen ist es keinesfalls garantiert, dass sich relevante Lerngelegenheiten bieten und falls dies doch passiert ist nicht garantiert, dass die gelernten Unterrichtsfähigkeiten wirklich *best-practice* darstellen (Billingsley & Scheuermann, 2014, S.257). Manche Forscherinnen behaupten, dass die Gefahr bestehe, dass Lehramtsstudierende nach Praxiserfahrungen einen engeren und dogmatischeren Unterrichtsstil, angelehnt an die didaktischen Vorstellungen der Instruktorin, pflegen würden (Brown, 1999, S.307). Diese Gefahr besteht allerdings bei allen – auch simulierten – Praxiserfahrungen, wenn die Feedbackkomponente nicht ausgeprägt oder die Erfahrung zu wenig theoretisch fundiert ist.

In einem Simulator können sämtliche realweltlich unveränderbaren Faktoren wie Klassengröße angepasst werden (Alessi & Trollip, 2001, S.227f.). Im besten Falle ist die Lerngelegenheit im Simulator adaptiv auf die jeweilige Nutzerin zugeschnitten und kontrolliert Rahmenvariablen so, dass es nicht zu einer Überforderung durch *cognitive load* kommen kann (Deale & Pastore, 2014, S.211). So kann es – abhängig von dem didaktischen Ziel der Simulationsnutzung – sinnvoll sein, die Klassengröße zunächst auf drei bis vier SuS zu limitieren, nur eine Schülerin mit Störungsverhalten zu simulieren oder die kognitive Leistungsfähigkeit einer Schülerin zu beeinflussen (Brown, 1999, S.314; Chini et al., 2016, S.3).

Meistert eine Simulationspartizipantin diese Herausforderung, können Simulationseigenschaften sukzessive komplexer werden und so der *cognitive load* erhöht werden bis zu dem Punkt, an dem optimales Lernen stattfindet. Einschränkend sei gesagt, dass die Kontrolle über Simulationsvariablen immer dann schwächer wird, je mehr menschliche Interaktion für die Simulation erforderlich ist. (Interaktionsmodellierung Typ 6, 7 und 8)

Zusammenhängend mit der Kontrolle aller unabhängigen Simulationsvariablen ist auch das diagnostische Argument. Eine Simulationsumgebung ist prinzipiell eine Experimentalsituation; es ist einfacher als in einer realweltlichen Situation diagnostische Kompetenzen zu messen (Dieker et al., 2007, S.3). Es ist möglich mit Simulationen Assessments durchzuführen (Ward et al., 2006, S.244ff.) und Simulationen als Performanz-Testinstrument zu benutzen (Alessi & Trollip, 2001, S.231). Als Beispiel für eine solche Simulationsnutzung seien die Studien von Südkamp et al. (2008), Südkamp und Möller (2009), Fiedler et al. (2002), Fiedler et al. (2007) und Fiedler (2012) und Hetmanek und van Gog (2017) genannt.

(e) **Effektivitätsargument**

Eine ganze Reihe von Studien formulieren theoretisch begründete Hypothesen zur erhöhten Wirksamkeit von Simulationen gegenüber traditionellen Methoden. Obgleich ein Großteil dieser Hypothesen noch nicht empirisch überprüft wurde ist es sinnvoll, sich mit den möglichen Potentialen von Simulationen zu beschäftigen. Eine zentrale theoretische Annahme ist, dass Unterrichten primär nicht deklarative, sondern prozedurale und metakognitive Wissensdimensionen benötigt, welche durch traditionelle Unterrichtsmethoden und Tests nur schlecht geschult werden können. Einzig eine situierte Durchführung der zu trainierenden Handlung (Ferry et al., 2005, S.85; Fischler, 2007, S.10f.) schließe aus, dass sich ‚träges Wissen‘ (Whitehead, 2012; Renkl, 1996; G. A. Blomberg, 2011, S.5) entwickle. ‚Träges Wissen‘ bezeichnet vorhandenes deklaratives Wissen, das nicht im Moment der Anwendung abgerufen werden kann. Simulationen gewährten also einen besseren Lerntransfer, so die Annahme (Alessi & Trollip, 2001, S.229; Fischler, 2006, S.35). Simulationen wird eine Scharnierfunktion, ergänzend zu realen Felderfahrungen, zum reibungslosen Übergang des universitären Kurssystems hin zur Praxis zugeschrieben (Billingsley & Scheuermann, 2014, S.256). Phänomene wie der berühmte Praxisschock sollen so vermieden und die Kluft zwischen Wissen und Handeln geschlossen werden (G. A. Blomberg, 2011, S.5; Ferry & Kervin, 2007, S.190). Ziel

ist eine Verzahnung von Theorie und Praxis (Bradley & Kendall, 2014, S.3; de Castell et al., 2014, S.336; Chieu & Herbst, 2011, S.105), indem die Simulationspartizipantinnen befähigt werden situationsangemessene mentale Modelle von Unterrichtshandeln zu entwickeln (Alessi & Trollip, 2001, S.214, 233; G. A. Blomberg, 2011, S.12 für den Bereich Video). Letztlich ist mit der umfassenden Digitalisierung der universitären Lehre auch die Hoffnung verbunden die Technikkompetenz angehender Lehrkräfte im Sinne einer *digital literacy* zu schulen und damit auch eine breitere, didaktisch sinnvolle Anwendung digitaler Methoden in die Klassenräume zu tragen (Egbert et al., 2000, S.23).

2.4.1 Empirische Studien

Empirische Studien zur Wirksamkeit von Simulationen speziell in der Lehrkräftebildung liegen bislang nur vereinzelt vor, bislang ist die Forschungslage uneindeutig (Fischler, 2006; Kervin et al., 2004, S.4; Dieker, Kennedy et al., 2014). In anderen Anwendungsbereichen konnte man positive Lerneffekte auf Simulationsnutzung zurückführen. Ein Beispiel dafür ist die große Metastudie von Hays et al. (1992) (kleine Effektstärken, $d=.02-.26$), die 247 Studien zu Flugsimulationen untersucht hat, welche Ward et al. (2006) zitieren und mit anderen (Meta-)Studien aus den Bereichen Sport und Medizin ergänzen (z.T. wesentlich höhere, mittlere bis große Effektstärken, z.B. $d=.5$, S.249). Insbesondere ist hier auf die Metastudie (N=109 Studien) von Issenberg et al. (2005) zu verweisen, welche die Effektivität von medizinischen Simulationen nahelegt, allerdings ohne Effektstärken zu berichten. Für das Gebiet Simulationen in der Lehrkräftebildung existieren bislang keine Metastudien, die tatsächliche Effektstärken berichten. Dies gilt auch für die noch am ehesten anzuführende Studie von Turan (2015), welche als systematisches Review aufgrund des geringen Literaturfundus nur 26 Studien berücksichtigt und eher theoretischer Natur ist. Diese Studie kommt zu dem recht unbefriedigenden Schluss, dass es zwar viele theoretisch fundierte Gründe zur Annahme gäbe, dass Simulationen handlungsnahe Lehrkraftkompetenzen schulen, die empirische Forschungslage aber stark verbesserungsbedürftig sei: „[There is] relatively little experimental or interventional research on the benefits of using simulations in teaching skills“ (Turan, 2015, S.36).

Ein generelles methodisches Problem von Simulationsstudien ist, dass die Effektivität von der didaktischen und zu einem kleineren Teil der technischen Ausgestaltung der Simulatoren abhängig ist. Anders formuliert: Nicht allein der Simulator trägt zu den Lerngewinnen bei, sondern die verwendete Simulation inklusive didaktischer Rahmung. Dies macht Metastudien in anderen Bereichen zusätzlich interpretationsbedürftig. Es könnte die These aufgestellt

werden, dass durch die Verwendung von weniger gut ausgestalteten Simulationen die potentielle Effektivität sogar eher unterschätzt wird. Allerdings ist eine Übertragung von berichteten Effektstärken aus anderen Lernkontexten in den Lehrkräftebildungsbereich nicht ohne konzeptionelle Probleme.

Ein wichtiger Faktor ist, dass für den Bereich Sport ausgeschlossen werden konnte, dass diese Wissenszuwächse in den handlungsnahen Kompetenzen nur auf eine *task familiarity*, also einer Vertrautheit mit der Aufgabe durch reine Beschäftigung mit dem Gegenstand, zurückzuführen sind (Ward et al., 2006, S.256). Dies ist deswegen von zentraler Bedeutung, weil es die theoretische Annahme, dass eine reale Handlungsdurchführung in einem situierten Lernumfeld tatsächlich besseren Wissenstransfer zu handlungsrelevanten Fertigkeiten ermöglicht (Effektivitätsargument), empirisch unterfüttert. Farrow et al. (1998) (zit. nach Ward et al., 2006, S.256) thematisieren dies in ihrer Studie. Anfänger im Tennissport wurden in drei Gruppen eingeteilt: Die Experimentalgruppe trainierte anhand eines filmbasierten Simulators die physische Reaktion auf einen vorgeführten Aufschlag, einer Placebogruppe wurden Tennisspiele vorgeführt und ihnen Fragen dazu gestellt und schließlich existierte eine Kontrollgruppe, welche keinerlei Schulung erhielt. Alle drei Gruppen absolvierten einen Pre- und Posttest, wobei sich zeigte, dass sich die Reaktionszeiten der Experimentalgruppe signifikant reduzierten, während die beiden anderen Gruppen keine Veränderungen zeigten. In einer frühen Studie aus dem Bereich der Lehrkräftebildung von Hershey et al. (1965) (N=280 Studierende) wurde jedoch ein gegenteiliges Ergebnis berichtet: Hier hatte eine Experimentalgruppe, welche einen analogen Simulator nutzt, den wir heutzutage eher als *microteaching* bezeichnen würden, keinen signifikanten Vorteil gegenüber der Kontrollgruppe, welche eine Unterrichtsbeobachtung durchlief. Die Effektivität wurde hier primär über verschiedene Multiple-Choice-Instrumente gemessen, jedoch auch durch Selbstauskünfte ergänzt.

In Bezug auf Simulationen in der Lehrkräftebildung ist diese beschriebene Effektivität der Übertragung ins Realweltliche bislang noch nicht ausreichend belegt (Dieker, Rodriguez et al., 2014, S.31). Verschiedene Studien (Chini et al., 2016, S.3; Andreasen & Hacıomeroglu, 2009; Christensen et al., 2011; Straub et al., 2014; Gibson, 2012a, S.5ff. M. Girod & Girod, 2006, S.489ff. Garland et al., 2012; Hopper et al., 2013; Y.-c. Yeh, 2002; Y.-C. Yeh, 2004; Y.-C. Yeh, 2006) berichten von beobachteten positiven Einflüssen auf handlungsnahen Lehrtätigkeiten, oft im allgemeinpädagogischen *classroom management* Bezug. Dabei ist allerdings zu beachten, dass die verwendeten Samplinggrößen sich primär im mittleren zweistelligen bis kleinen dreistelligen Bereich bewegen und daher nicht generalisierbar sind (Turan, 2015, S.40ff. nur für Second Life/*virtual worlds* F. Wang & Burton, 2012, S.364).

Ebenfalls werden vorrangig Selbstauskunftsbögen genutzt, um wahrgenommene Fähigkeitszuwächse zu quantifizieren (M. Girod & Girod, 2006). Für die Simulationen TeachLivE und simSchool liegen die meisten empirischen Untersuchungen zur Effektivität vor.

Relativ gut beforscht sind positive motivationale Effekte (Badiee & Kaufman, 2015; Ferry et al., 2005, S.90; Alessi & Trollip, 2001, S.229, 245; Cheong & Kim, 2009; Höntzsch et al., 2011; Y.-C. Yeh, 2004; IsaBella et al., 2006). Die Vermutung liegt nahe, dass diese motivationalen Effekte auf den verwendeten, von kommerziellen Spielen übertragenen Spielmechaniken wie Maximierung von Punkten oder Wahl des besten Simulationspfades beruhen. In Bezug auf VR-Simulationen, zu denen zum Zeitpunkt des Schreibens dieser Arbeit noch kaum Befunde vorliegen, ist auch ein Neuheitseffekt der verwendeten Technik anzunehmen.

Auch liegen einige Studien vor, die eine Steigerung der Selbstwirksamkeitserwartung nach Simulationsnutzung feststellen (Gundel et al., 2019; Hopper et al., 2013; Garland et al., 2012; McPherson et al., 2011; Billingsley & Scheuermann, 2014; Foley & McAllister, 2005, S.167, 174f. Christensen et al., 2011). Oft wird eine Simulation positiver beurteilt, je mehr mögliche Verhaltensweisen simuliert werden können, also umso höher die Modellfidelität ist (Fischler, 2006, S.198; Brown, 1999; Chini et al., 2016). Die ökologische Validität der Simulationen wird meist über Surveys der Simulationspartizipantinnen beforscht, tendenziell qualitativ. (z.B. Fischler (2006) und Christensen et al. (2011)) Hier zeigt sich erwartungsgemäß, dass Systeme, die eine hohe menschliche Interaktion erfordern, d.h. mit *microteachings* vergleichbar sind, in der Regel besser bewertet werden als Simulationen des Interaktionsmodellierungstyps 1–5.

Ob der Einsatz flankierender Materialien neben der Simulatornutzung für die Simulation ergiebig ist, ist weiterhin unklar. Während Dieker et al. (2017) zum Ergebnis kommen, dass zusätzlich bereitgestellte Informationen sogar die Effektivität der Simulation kontraintuitiv schmälern, betonen andere wie Kervin et al. (2006) stark deren wahrgenommene Nützlichkeit im Rahmen von Interviewaussagen von N=186 Studierenden. Es ist zu erwarten, dass eine pauschale Bewertung des Einsatzes flankierender Materialien auch in Zukunft nicht zu erwarten ist, da es sich hierbei um eine Mediendiskussion handelt. Auch hier gilt: Nicht die bloße Existenz von Rahmenmaterialien, sondern deren Konstruktion und Einbettung in das didaktische Simulationskonzept entscheiden über die Nützlichkeit. Es ist sehr wahrscheinlich, wenn man die Forschungserkenntnisse aus der Videoforschung hinzuzieht, dass Kontextinformationen zu den simulierten Situationen (Schulstunden, Tasks, etc.) als sehr hilfreich wahrgenommen werden können.

2.5 Bisherige Nutzung angrenzender Termini

2.5.1 *virtual classroom*

Der Begriff *virtual classroom* wird dazu eingesetzt, um Unterricht zu bezeichnen, der mit einer Lehr-/Lernplattform und verschiedenen Audio-/Videotechnologien bzw. in einem Web-Konferenzsystem in Echtzeit online stattfindet (Falck et al., 1997; Czerwionka et al., 2009, S.96). Ziel dieser Anwendungen ist insbesondere Unterricht in Situationen zu ermöglichen, in denen die klassische Präsenzlehre nicht oder nur schwer möglich ist. Ein Beispiel hierfür wäre die Unterrichtsimpementation in abgelegenen ländlichen Gebieten in Australien (Palloff & Pratt, 2013), Finnland (Falck et al., 1997) oder Kanada (Barbour, 2007; Barbour, 2015). Ein anderer klassischer Anwendungsfall ist das Fernstudium (McBrien et al., 2009). Lernende werden hier also nicht simuliert, sondern nehmen real am Unterrichtsgeschehen teil. Mit Aufkommen der Corona-Pandemie rückten diese *virtual classrooms* schlagartig in den gesellschaftlichen Fokus. Die Verwendung des deutschen Begriffes „Virtuelles Klassenzimmer“ ist mit der englischsprachigen Verwendung deckungsgleich. (Czerwionka et al., 2009; Mengel, 2011)

Bei *virtual classrooms* handelt es sich nicht um Simulationen, sondern um eine spezielle Form von Fernunterricht. Analog dazu bezeichnen die Begriffe *virtual schooling* (Barbour, 2007) den gesamten Prozess der Fernbeschulung, *virtual teaching* (Barbour, 2015) den Prozess des Unterrichtens und *virtual student* den realen Lernenden, der an dem Lehr-/Lernsetting online teilnimmt.

Der Begriff *virtual classroom* wird hier angeführt, da er gelegentlich mit anderer Bedeutung und damit missverständlich verwendet wird. So wird er beispielsweise in Ferry et al. (2004, S.300) und Ferry et al. (2005, S.87) genannt, um die Online-Komponente des Simulators ClassSim herauszustellen. Auch in den Veröffentlichungen um die Simulation TeachLivE wird diese teilweise als *virtual classroom* bezeichnet (Dieker et al., 2007, S.5; Garland et al., 2012, S.504), genauso wie die Simulation Tigerlake Project (Fischler, 2006, S.60). In den Veröffentlichungen zur Simulation simSchool findet sich die Erwähnung des Begriffes nicht; es ist davon auszugehen, dass es sich bei dem von Foley und McAllister (2005, S.161, 164) erwähnten Sim-School [sic!] um eine andere Simulation handelt, die wiederum als *virtual classroom* bezeichnet wird. Angesichts der bereits ausgeführten Doppelverwendung des Begriffes wird von einer synonymen Verwendung zur digitalen Simulation, wie sie gelegentlich in der Literatur praktiziert wird, abgeraten.

2.5.2 *virtual learning environment (VLE)*

Dieker et al. (2008, S.10) definieren *virtual environments* (VE) äußerst weit als „all systems in which a user interacts with virtual objects“. In ihrer Taxonomie zu *virtual environments* unterscheiden Bailenson et al. (2008) zwei spezialisierte Typen von VEs: Erstens das *immersive virtual environment (IVE)*, dass sich von einer regulären VE durch seine höhere *presence* (nähere Definition in Abschnitt 2.5.7, S. 30) unterscheidet. Bei dieser Form von VE haben die Autorinnen die Idee von HMD-Anwendungen (head-mounted displays) vor Augen. Zweitens definieren sie ein *collaborative virtual environment* als VE, in dem mehrere Personen miteinander interagieren können. Der Begriff scheint weitestgehend deckungsgleich mit MUVE (Multi User Virtual Environments) zu sein (Bradley & Kendall, 2014, S.7).

Unter VEs selbst verstehen Bailenson et al. (2008) „synthetic sensory information that leads to perceptions of environments and their contents as if they were not synthetic“. (Bailenson et al., 2008, S.103)

Virtual learning environments (VLE) beinhalten nach der häufig zitierten Definition von Dillenbourg et al. (2002) sieben Merkmale:

- (a) Ein *virtual learning environment* ist ein strukturierter Informationsraum.
- (b) Ein *virtual learning environment* ist ein sozialer Raum, in dem Lerninteraktionen zwischen verschiedenen Personen stattfinden.
- (c) Der virtuelle Raum wird explizit repräsentiert, entweder in Form von Text oder immersiver in Form von 3D-Welten.
- (d) Lernende sind nicht nur aktiv, sondern gestalten den virtuellen Lernraum mit.
- (e) *Virtual learning environments* umfassen nicht nur Distanzlernen.
- (f) *Virtual learning environments* integrieren heterogene Technologien und pädagogische Ansätze.
- (g) Die meisten *virtual learning environments* überlappen sich mit physischen Umgebungen.

Eine alternative Definition von Mikropoulos (2006) ist, dass VLEs *virtual environments* sind, die „ein oder mehrere Lernziele oder pädagogische Metaphern besitzen und Nutzerinnen mit Lernerfahrungen konfrontieren, die sie in der realen Welt nicht machen können und die zu spezifischen Lernoutcomes führen“ (Mikropoulos, 2006, S.198).

Anhand dieser Ausführungen wird deutlich, dass es sich bei *virtual environments* und *virtual learning environments* um Hyponyme handelt (Richards & Taylor, 2015, S.157), die vereinzelt auch als Selbstbeschreibungen für digitale Simulationen verwendet werden (z.B. Delamarre, Lunn et al. 2019). So wird auch die Simulation TeachME/TeachLivE von Andreasen und Haciomeroglu (2009) als *virtual learning environment* und von Dieker et al. (2017), als *virtual environment/virtual teacher environment* beschrieben – was zum einen korrekt ist, weil damit auf die kollaborative Form der Interaktionsmodellierung hingewiesen wird, aber zum anderen sehr verwirrend sein kann, da der Begriff so breit definiert wird. Häufiger wird der Begriff im Kontext der Arbeit mit virtuellen Welten (siehe Abschnitt 2.5.3, S. 27) wie Second Life verwendet (Kemp et al., 2009; Kallonis & Sampson, 2010). Dabei scheinen VLEs häufig einen anderen didaktischen Zweck zu erfüllen, der eher auf ein exploratives Nutzen digitaler Technologie gerichtet ist, um z.B. Lehrkräfte im Sinne der *digital literacy* zu schulen. Während in Simulationen die Schulung digitaler Kompetenzen mehr eine (häufig optionale) Nebenzielsetzung darstellt, ist eine solche Schulung in *virtual learning environments* oft das Hauptziel (Vgl. F. Wang & Burton, 2012).

2.5.3 *virtual worlds*

Virtuelle Welten oder *virtual worlds* können definiert werden als immersive dreidimensionale Umgebungsrepräsentation, in denen Nutzerinnen mittels Avataren interagieren können (Bulu, 2012, S.154; Gregory et al., 2012, S.5; Richards & Taylor, 2015, S.157; Gisbert et al., 2013, S.1991). Ein anderer Definitionsversuch ist, dass *virtual worlds* in Verwandtschaft zu Computerspielen gedacht werden müssen. *virtual worlds* basieren aber nicht wie Computerspiele auf regelbasierten Spielmechaniken, sondern sind persistente soziale Welten, die lose durch offene narrative Strukturen verbunden sind und in denen Nutzerinnen tun können, was sie möchten (Galarneau & Zibit, 2007, S.61).

Der Begriff überschneidet sich gerade in der ersten Definition so stark mit *immersive virtual environments* und *collaborative virtual environments*, dass er oft synonym verwendet wird (Gisbert et al., 2013). Im allgemeinen Forschungsdiskurs wird er verwendet um mehrspielerbasierte Online-Welten wie Second Life zu bezeichnen (Cho et al., 2015, S.70; Bulu, 2012, S.154). Oft werden *virtual worlds* als „empty spaces“ bezeichnet (Kallonis & Sampson, 2010, S.36), da sie im Gegensatz zu Simulationen über keine inhärente didaktische Anlage oder Zielsetzung verfügen, sondern, wie die zweite Definition deutlich macht, offen angelegt sind. Dass *virtual worlds* damit nur eine Art loses Fundament sind, auf das Simulationen aufgebaut werden können, aber

nicht zwangsläufig müssen, formuliert Aldrich (2009, S.xxxiv) treffend:

Most people lump Second Life and World of Warcraft into this area on their own. But it's not that simple. Virtual worlds can be a platform for Sims, much as Flash or commercial game engines can. If used for that purpose, they both increase the speed of development and shape the content of the product. But accessing a virtual world does not give one a Sim for free.

Damit aus einer *virtual world* eine Simulation werden kann, müssen folgende Komponenten erstellt werden: (a) Eine didaktische Rahmung (u.a. geskripte Story, Levels/Tasks, Interaktionsmodellierung), (b) eine Feedbackkomponente und (c) eine Benutzerinnenoberfläche, welche mit den didaktischen Zielen kompatibel ist (Aldrich, 2009, S.392).

2.5.4 *educational simulations*

Eine *educational simulation* gibt Nutzerinnen die Möglichkeit innerhalb von *virtual environments* durch Handlungsdurchführungen zu lernen, indem sie theoretische Konzepte praxisnah in realistischen, aber kontrollierten Umgebungen erproben (Fischler, 2006, S.1f.). Eine andere Definition schlagen Alessi und Trollip (2001, S.213) vor: „An educational simulation can be defined as a model of some phenomenon or activity that users learn about through interaction with the simulation.“ Der Begriff umfasst nicht nur die Simulationsnutzung zur Ausbildung von Studierenden, sondern beinhaltet auch die Anwendung von Simulationen innerhalb des Klassenkontextes, d.h. durch SuS.

Außerdem veranschaulicht die Verwendung des Begriffes die Abgrenzung zu technischen Simulationen, welche nicht entwickelt wurden um einen didaktischen Zweck zu erfüllen, z.B. Wettersimulationen. Eine ähnliche Abgrenzung findet sich in dem Begriff *social simulations*: Hier wird die Abgrenzung von präzisen Simulationsmodellen naturwissenschaftlicher Simulationen zu den eher interpretativen sozialwissenschaftlichen Simulationsmodellen herausgearbeitet. Alle Simulationen, auch naturwissenschaftliche, simplifizieren die Realität anhand eines Simulationsmodells (Alessi & Trollip, 2001, S.248; Gibson, 2009a, S.680; Gibson, 2012a, S.2). In einer sozialwissenschaftlichen Simulation ist eine grobe Approximation ausreichend, solange die für das Lernen notwendigen Variablen abgebildet werden (Alinier, 2010, S.16; Alessi & Trollip, 2001, S.214; Cruickshank, 1972, S.17). Es ist aufgrund der Unprognostizierbarkeit menschlichen Handelns weder möglich, noch aus didaktischen Gründen wünschenswert, die Realität innerhalb einer solchen Simulation zu replizieren. Eine Variablenreduktion ist oft sogar sinnvoll, um die

Nutzerinnen mit einer handhabbaren Anzahl an Variablen und Handlungsalternativen zu konfrontieren. Modelliert man eine überbordende Menge an Variablen, so ist die Gefahr groß, die Nutzerinnen damit eher zu verwirren und die didaktische Zielsetzung unabsichtlich zu verschleiern (Alessi & Trollip, 2001, S.228; Gibson, 2012a, S.2). Je realitätsnäher, d.h. fideler, das Simulationsmodell ist, umso mehr *cognitive load* (Sweller, 1988) wird bei den Nutzerinnen der Simulation erzeugt, was sich gerade bei Novizinnen im simulierten Lernfeld nachteilig auf den Lerneffekt auswirkt.

Durch die Verwendung des Begriffes *educational simulations* wird das bereits in [Abschnitt 2.2](#) (S. 7) angesprochene Problem der Assoziation mit hoher Prädiktorengenauigkeit wenn nicht aufgelöst, aber doch zumindest begrifflich thematisiert.

2.5.5 *classroom simulation / virtual classroom simulation*

Der Begriff der *classroom simulation* (z.B. verwendet bei Badiee & Kaufman, 2015, S.3) schildert auf den ersten Blick relativ exakt das Anwendungsgebiet der Simulation und ist auch aufgrund seiner Griffigkeit nützlich. Allerdings lässt der Begriff unbestimmt, welche Art von Klassenraum simuliert wird. So fallen unter diesen Begriff nicht nur Simulationen von Schulklassen, sondern auch Simulationen von Klassenräumen in der Erwachsenenbildung, z.B. in der Pflege (Goldenberg et al., 2005). Damit verwandt ist der Begriff des *simulated teachings* oder *teaching simulation*, welcher von Mukhtar et al. (2018) im Sinne eines erweiterten, zeitlich längeren *microteaching-peerteachings* verwendet wird.

Die Bezeichnung *virtual classroom simulation* (z.B. Skrødal, 2010) verweist auf die digitale Umsetzung der Klassenraumsimulation, wobei die Gefahr des Missverständnisses bezüglich des *virtual classroom* besteht. Man könnte dabei unter anderem an VR-Klassenraumsimulationen denken, auch wenn diese in der Benutzung von Skrødal (2010) damit nicht gemeint sind. Letztlich ist anzumerken, dass der Begriff nur Klassenraumsimulationen abdeckt. Eine Simulation anderer Handlungsorte wie Schulpausen oder Elterngespräche, die im Bereich der Lehrkräfteaus- und -fortbildung wichtig sein könnten, umfasst der Begriff nicht.

2.5.6 *computer simulation*

Der Begriff *computer simulation* zielt allein auf die technische Implementation der Simulation. Im Gegensatz zu den frühen analogen, medien- oder rollenspielbasierten Simulationen werden Computersimulationen auf einem

digitalen Endgerät präsentiert. Der Begriff umfasst allerdings auch Simulationen ohne jede Nutzerinneninteraktion wie beispielsweise physikalische Simulationen.

Computersimulationen wurden traditionell immer in Szenarien eingesetzt, in denen die reale Durchführung kompliziert oder mit hohen Risiken („high-risk“, Hochrisiko-Situationen) behaftet war. Beispiele dafür sind medizinische und militärische Simulationen, sowie Flugsimulationen (Bradley & Kendall, 2014, S.4; Brown, 1999, S.308; Bulu, 2012, S.154). Ein anderes Beispiel ist die Anwendung einer Simulation durch die NASA zur Schulung von Flight-Controllern im Kontext der Reparatur des Hubble Teleskopes (Brown, 1999, S.311).

2.5.7 *presence* und *immersion*

Physical presence (oder nur *presence* genannt), bezeichnet das subjektive Erleben sich in einer anderen Umgebung zu befinden als in der, in der man sich physisch gerade aufhält (Bulu, 2012, S.155; Cho et al., 2015, S.71; Mikropoulos, 2006, S.197; Persky et al., 2009, S.263; Psotka, 1995, S.405; Bailenson et al., 2008, S.109). Dies wird möglich durch *suspension of disbelief* (Hayes et al., 2013, S.144), d.h. durch die Willensanstrengung eine künstliche Welt als aktuelle Umgebung zu akzeptieren. Es ist wichtig zu erwähnen, dass dies keine VR-Techniken und noch nicht einmal Digitalisierung zwingend erfordert. Es ist möglich, sich in einem Buch oder einem Bild derart zu verlieren, dass man die eigene physische Umgebung subjektiv kaum oder nicht mehr wahrnimmt und sich an einen anderen Ort versetzt fühlt. Eine andere Form der *presence* ist *social presence*, was das Gefühl bezeichnet sich mit anderen in einer virtuellen Welt zu befinden (Bulu, 2012, S.155; Cho et al., 2015, S.71). Delamarre, Buche und Lisetti (2019, S.183) verstehen das Konzept von *social presence* in einem geringfügig anderen Sinne, was recht große konzeptuelle Konsequenzen hat: Für sie bedeutet es den Grad, wie computergesteuerte Agenten von Simulationspartizipantinnen als realistisch wahrgenommen werden. Dieses Verständnis würde sich mit der Fidelität des Simulationsmodells überschneiden.

Die Implikationen von *social presence* werden im Rahmen der Taxonomie von Simulatoren, insbesondere der Interaktionsmodellierung, im [Abschnitt 3.3](#) (S. 54), thematisiert.

Presence ist im Umgang mit Computersimulationen ein wichtiges Konstrukt, das von *Immersion* zu trennen ist (Hayes et al., 2013, S.143). Während im alltäglichen Sprachgebrauch der Begriff *presence* kaum vorkommt und häufig durch *Immersion* ersetzt wird, so bezeichnet der letzte Begriff doch lediglich die technische Realisierung der Präsentationsstufe der virtuel-

len Welt. Anders gesagt: Ob eine Welt mittels Text, 2D- oder 3D-Grafik in unterschiedlicher Qualität, via HMD oder CAVE abgebildet wird, bestimmt über die Immersionsstufe der virtuellen Welt, nicht aber über das subjektive Erleben. So können sich einige Personen in Textwelten relativ problemlos verlieren, erreichen hier also eine hohe *presence*, können dasselbe in hoch-immersiven VR-Welten nicht, da sie z.B. unter der Simulatorkrankheit leiden (Cho et al., 2015). Die sogenannte Simulatorkrankheit beschreibt das Aufkommen von Übelkeit und Unwohlsein beim Verwenden einer VR-Brille. Als personenbezogene Distinktionsmerkmale bezüglich *presence*-Wahrnehmung in der gleichen virtuellen Welt wurden Alter und epistemologische Überzeugungen identifiziert, während das Geschlecht keine Rolle spielt (Cho et al., 2015, S.71). Ebenfalls scheint eine technologische Affinität und der Grad des Konsums von Computerspielen die *presence* zu beeinflussen (Bulu, 2012), wobei aufgrund der kleinen Stichproben diese Ergebnisse keineswegs als gesichert angesehen werden können. Es ist zudem denkbar, dass die von Cho et al. (2015, S.71) identifizierten Distinktionsmerkmale lediglich mediierende Variablen darstellen.

Es gibt es Hinweise darauf, dass eine Steigerung der Repräsentationsfidelität, z.B. durch bessere Grafiken und Animationen für Simulationsagenten im Mittel über alle Versuchspersonen hinweg die *presence* steigert (Cho et al., 2015, S.71; Bulu, 2012, S.156). Zusätzlich steigern auch realistische Klassenraumumgebungen die wahrgenommene *presence* (Stavroulia et al., 2018, S.278).

Andere Forscher merken zu Recht kritisch an, dass der Fokus auf reiner Repräsentationsfidelität nicht ausreichend ist und fordern eine stärkere Auseinandersetzung mit den zugrundeliegenden Interaktionsmodellierungen, welche einen erheblich größeren Anteil an der wahrgenommenen *presence* hätten (Deale & Pastore, 2014, S.200). Dies kann man an einem Beispiel illustrieren. In einer hoch-immersiven Simulation seien Partizipantinnen mit aggressivem SuS-Verhalten konfrontiert. Sie haben dort jedoch nur die Wahl den Aggressor zu tadeln oder zu ignorieren. Als Outcomes sind entweder Erfolg oder Misserfolg möglich. Es ist unmittelbar ersichtlich, dass in dieser Simulation die *presence*-Werte der Simulationsnutzerinnen vermutlich niedriger sind als in einer weniger immersiven Simulation, die jedoch viele Handlungsalternativen und -outcomes beinhaltet. Außerdem schränken Cho et al. (2015) zusätzlich ein, dass es bei Steigerung der Simulationsfidelität zu unerwünschten Nebeneffekten kommen kann. Je realer und komplexer die virtuelle Welt ist, so die Theorie, umso mehr *cognitive load* (Siehe [Abschnitt 2.5.15](#),) wird erzeugt, was lernhinderlich ist (Cho et al., 2015, S.70). Weitere theoretische Überlegungen zu den veränderbaren Variablen der Repräsentationsfidelität werden im [Abschnitt 2.5.12](#) (S. 38) gemacht.

Der Zusammenhang zwischen *presence* und Lernoutcomes ist immer noch nicht geklärt. Während einige Studien von einem positiven Zusammenhang ausgehen (Lee & Ashton, 2012; z.T. Persky et al., 2009), sehen andere keine oder eine negative Korrelation (Merchant et al., 2012, S.562; Moreno & Mayer, 2004, S.170f.), was durch die Überlegungen hinsichtlich des *cognitive loads* zu erklären ist. *Presence*, insbesondere *social presence*, scheint allerdings ein Prädiktor für die Zufriedenheit der Nutzerinnen zu sein (Bulu, 2012, S.159).

2.5.8 *microteaching*

Microteaching bezeichnet ein Trainingsverfahren für Lehrkräfte, das Mitte der 1960er Jahre in Stanford entwickelt wurde (Fortune et al., 1967). Während des *microteachings* unterrichten angehende Lehrkräfte reale SuS und erhalten im Anschluss dazu Feedback von Supervisoren und anderen Studierenden (Allen, 1967). Oft gibt es die Möglichkeit die Resultate dieser Trainingssessions via Videoaufzeichnung erneut zu beobachten. Der zentrale Unterschied von *microteaching* gegenüber realen Hospitationen ist zum einen der schon erwähnte höhere Grad an diagnostischem Feedback, zum anderen eine Komplexitätsreduktion der Unterrichtssituation (Allen & Eve, 1968). So werden keine ganzen Stunden, sondern nur ca. fünf bis maximal zwanzig Minuten lange, thematisch und unterrichtsstrukturell fokussierte Unterrichtsausschnitte durchgespielt (Allen, 1967; Allen & Eve, 1968). Zudem ist die Anzahl der SuS auf eine Kleingruppe (genannt *microclass*) von ca. 5 SuS limitiert (Allen & Eve, 1968).

Es ist gerade in der Diskussion um Simulationen notwendig, das *microteaching-peerteaching* (Zifreund, 1966) vom Allen'schen Konzept abzugrenzen. Die gleiche Abgrenzung nimmt auch Klinzing (2002) in seiner Metastudie zur Effektivität von *microteachings* vor. *Microteaching-peerteaching* unterscheidet sich gegenüber klassischem *microteaching* dadurch, dass keine realen SuS partizipieren, sondern diese von Studien- und Lehr(ausbildungs)kolleginnen geschauspielert werden. Die Abgrenzung zu einem Rollenspiel ist beim *microteaching-peerteaching* fließend und schwer explizit festzumachen. Wenn man den Versuch unternehmen will die Begriffe auszudifferenzieren, unterscheidet sich das *microteaching-peerteaching* vom Rollenspiel durch das spezifische didaktische Setting und einen inhärenten Fokus auf Komplexitätsreduktion (Zeit, Klassengröße und Thema). Als letzte Komponente kommt das Feedback hinzu. *Microteaching-peerteaching* steht in enger Verwandtschaft zu *virtual puppetry* (Interaktionsmodellierung des Types 7) und lässt sich als analoge Simulation klassifizieren (Interaktionsmodellierung des Types 8). Klinzing (2002, S.201f.) analysiert in seiner Metaanalyse (N=10 Studien), ob *microteachings* stärkere Trainingseffekte hervorrufen als *microteaching-peerteachings*.

Dabei konnten keine statistisch signifikanten Effektivitätsunterschiede zwischen beiden Trainingsverfahren festgestellt werden. In aktuellen Einzelstudien werden theoretische Erwägungen angeführt, welche die Fidelität von *microteaching-peerteaching* bzw. den verwandten, z.T. computergestützten, Rollenspielen anzweifeln (Aldrich, 2009, S.45; Chini et al., 2016, S.1f. Brown, 1999, S.307; Gregory et al., 2012, S.29; M. Girod & Girod, 2006, S.484). So transportieren Studierende oder Lehrkräfte, die in *microteaching-peerteachings* SuS imitieren sollen, zum einen ihre eigenen biographisch-anekdoteschen Schulerfahrungen in ihre Rolle, zum anderen benötigt die Simulation einer Schülerin nicht zu unterschätzende schauspielerische Fähigkeiten. Während mit vorgegebenen Rollenkarten zumindest versucht wird, den Einfluss biographisch-anekdotescher Erfahrungen zu beschränken, so wird die genaue Umsetzung, auch im gestalterischen Rahmen, aus Zeitgründen so gut wie nie geschult.

Es ist zusammenfassend wahrscheinlich, dass *microteachings* unabhängig von der verwendeten Trainingsform effektiv sind. Zu diesem Schluss kommen 93 von den 113 untersuchten Studien bei Klinzing (2002, S.202).

2.5.9 *critical incidents*

Critical incidents bezeichnen Ereignisse, die als wichtige Wendepunkte in der professionellen Berufsinteraktion angesehen werden und damit ein hohes Potential besitzen didaktische Lerngelegenheiten für Fachpersonal zu ermöglichen (Tripp, 2011; Megawati et al., 2020, S.944). *Critical incidents* werden damit konstruiert und sind durch die Signifikanzinterpretation nicht vom Beobachter zu trennen (Tripp, 2011). Die Methode der Nutzung vom *critical incidents* als Reflexionsstimulus wurde in den 1940er Jahren im militärischen Zusammenhang der US-Luftwaffenforschung entwickelt und eingesetzt (Göbel, 2003, S.2). Durch ihre universelle Einsatzfähigkeit werden *critical incidents* in der Medizin, der Flugausbildung, Beratung, Organisationsentwicklung, im interkulturellen Training und erst seit vergleichsweise kurzer Zeit in der Lehrkräftebildung rezipiert (Göbel, 2003). Beobachterinnen von *critical incidents* sind im pädagogischen Umfeld oft die ausführenden Lehrkräfte, die selbst die Erfahrungen der *critical incidents* im eigenen Schulalltag gemacht haben. Es ist aber auch möglich, dass eine dritte Person die Ereignisse identifiziert. Die kritische Komponente ist dabei immer auf ein spezifisches Thema bezogen. Als Beispiele nennen Megawati et al. (2020) Probleme bei der Unterrichtsdurchführung, der Beteiligung der SuS, Sprachkompetenzprobleme oder die Unterrichtsvorbereitung. *Critical incidents* können sowohl problematisches Verhalten von Lehrkraft oder SuS darstellen, als auch *best-practice* zeigen (Göbel, 2003, S.1), wobei die Literatur eher über negative Beispiele be-

richtet (Megawati et al., 2020, S.945). *Critical incidents* sind eng verknüpft mit der Methode der reflexiven Praxis (Megawati et al., 2020, S.944) und des diagnostischen Unterrichtens (Tripp, 2011, S.8); sie sollen genutzt werden um effektiv Theorie und Praxis miteinander zu verknüpfen (Tripp, 2011, S.21). Tripp (2011) argumentiert, dass *critical incidents* eine ausgezeichnete Möglichkeit seien, zunehmendes Verständnis für und Kontrolle über das professionelle Urteilsvermögen und somit über die Praxis zu entwickeln. Außerdem dienen *critical incidents* als Mittel, Klassenrauminteraktionen wissenschaftlich zu erschließen. Eine reine Beschreibung oder Videoaufzeichnung der relevanten Ereignisse konstituiert nach Tripp (2011) strenggenommen nur ‚incidents‘. Erst unter dem Brennglas der Analyse und der diagnostischen Reflexion werden aus diesen Ereignissen *critical incidents*.

In einer anderen, älteren Lesart, werden *critical incidents* oder *critical events* als Ereignisse interpretiert, die radikale Änderungen in SuS und Lehrkräften erzeugen (Woods, 1993). In dieser Begriffsdefinition sind *critical incidents* konzeptionell Routineereignissen entgegengestellt. Da sich diese Begriffsdefinition jedoch als weniger gewinnbringend für die Lehrkräftebildung herausgestellt hat, unter anderem, da solche Ereignisse nur selten anzutreffen sind, wird der Begriff *critical incidents* in dieser Arbeit im moderneren Kontext verstanden. Die meisten *critical incidents* enthalten im jetzigen Verständnis demzufolge eher typische als dramatische Ereignisse.

In jedem Fall weisen *critical incidents* eine Nähe zu Simulationsszenarien oder Vignetten (Delamarre et al., 2017, S.401) auf, von denen anzunehmen ist, dass sie die Simulationsnutzerinnen mit einem didaktisch sinnvollen Handlungsrahmen konfrontieren (de Jong et al., 2012; Göbel, 2003). Diese Genese von *critical incidents* in videografischer Form eignet sich besonders als Grundlage zur Erstellung von Simulationsszenarien, da hier bereits Handlungsalternativen zu sehen sind, deren Auswirkungen visuell und auditiv betrachtet werden können. Im engeren Sinne wird unter Simulationsszenarien nicht nur die Kontextrahmung einer Situation (Alessi & Trollip, 2001, S.243f.) verstanden, sondern zusätzlich die Auswahl sinnvoller Handlungsalternativen zur Erreichung eines didaktischen Ziels. Analog zum Jargon aus der Computerspielwelt wäre ein Simulationsszenario nicht nur ein Intro, sondern ein Level. Ein Simulator kann technisch entweder auf ein Simulationsszenario explizit zugeschnitten sein oder mehrere Szenarien beinhalten.

Im Projekt Level/The Next Level wurden anhand von ca. 600 Videos ganzer Schulstunden verschiedene *critical incidents* identifiziert, die wiederum als Arbeitsgrundlage für die Szenarien in Clasivir 1.0 (erste Version einer VR-Simulation, siehe [Abschnitt 8.3](#), S. 238) genutzt wurden. Außerdem basiert das in dieser Arbeit verwendete Mockup-Video auf Daten einer videografisch erhobenen Schulstunde aus diesem Fundus (siehe [Abschnitt 6.3](#), S. 200).

2.5.10 (*intelligent*) *agents*

Unter dem Begriff Agent versteht man eine Entität – in der Regel ein Computersystem –, die in einer Umgebung situiert ist und autonome Entscheidungen ausführen kann, um seine vorgegebenen Ziele zu erfüllen (Wooldridge, 2002, S.5). Auf eine weitere Präzisierung der Definition über die Existenz von Autonomie innerhalb des Agenten konnte sich bislang nicht geeinigt werden. Selbst das Merkmal der Autonomie kann man kontrovers betrachten: Eine volle Autonomie in dem Sinne, dass die Software-Agenten völlig eigenständig reagieren ist natürlich nicht gegeben, sie sind abhängig von den vordefinierten Grenzen ihrer Programmierung. Im Kontext der KI-Forschung bezeichnet Autonomie jedoch nur das Treffen von Programmentscheidungen unabhängig von Benutzereingaben, insofern ist der Begriff in diesem Sinne korrekt (Wooldridge, 2002, S.5). Im Forschungsfeld um die Nutzung von Simulationen in der Lehrkräftebildung wird häufig eine simple, aber nützliche Unterscheidung zwischen *agents* und *avatars* gezogen: Agenten werden durch Algorithmen gesteuert, Avatare durch Menschen (Bailenson et al., 2008, S.105; Nagendran et al., 2013, S.144).

Wichtig zu erwähnen ist, dass der Begriff somit extrem weit definiert ist. Auch ein Thermostat kann nach Woolridge als Agent verstanden werden, denn das Thermostat ist in einer Umgebung situiert (Temperatursensoren in einem Raum), trifft autonome Entscheidungen (Raum kühlen oder Raum erhitzen) und erfüllt vorgegebene Ziele (z.B. Temperatur auf 20 Grad Celsius halten) (Wooldridge, 2002, S.7).

Im Gegensatz dazu ist ein *intelligent agent* ein Agent, der flexible autonome Entscheidungen ausführen kann, um seine vorgegebenen Ziele zu erreichen. Flexibilität kann als das Vorhandensein von drei Faktoren definiert werden:

- (a) Reaktivität: Der Agent kann die Umgebung wahrnehmen und zeitnah auf Veränderungen so reagieren, wie es die Ziele des Agenten vorsehen.
- (b) Pro-Aktivität: Der Agent zeigt zielorientiertes Verhalten und zeigt eigenständig Initiative, um diese Ziele zu erreichen.
- (c) Soziale Fähigkeit: Der Agent kann mit anderen Agenten oder Menschen zum Zweck der Zielerreichung interagieren (Wooldridge, 2002, S.8).

Intelligent agents sind nach dieser Definition recht einfach aufgebaut. Bereits ein Thermostat, der Wetterseiten im Internet abrufen, um eventuelle Temperaturanstiege oder -abfälle zu antizipieren, erfüllt das Kriterium der Pro-Aktivität. Die Schwächen der Definition von ‚Reaktivität‘ und ‚sozialer Fähigkeit‘ werden allerdings an dem von Woolridge selbst gewählten Beispiel deutlich: Bezüglich der Reaktivität wird nicht klar, wie stark ein Agent seine

Umgebung wahrnehmen muss. Es ist unklar, ob es bereits ausreicht, dass der Thermostat einen Temperatursensor hat, oder ob noch zusätzliche Informationen über die Umgebung gesammelt werden müssen, damit dieses Kriterium als erfüllt anzusehen ist.

Die soziale Fähigkeit lässt offen, ob bereits die Interaktion mit einem Menschen, der das Thermostat bedient, ausreichend ist, um das Kriterium zu erfüllen, oder ob dazu eine Interaktion verschiedener Thermostate nötig ist. Reiner Austausch von Daten mehrerer Geräte stellt auch nicht zwangsläufig soziale Fähigkeiten dar (Wooldridge, 2002, S.10).

Aus diesen Überlegungen folgt, dass *intelligent agents* sowohl simple als auch sehr komplexe Systeme sein können. Es ist von Bedeutung, dass die Agenten in komplexen, dynamischen Umgebungen agieren, die sich schnell ändern und häufig mit Unsicherheit behaftet sind. Zentrale Herausforderung in dem Agenten-Design ist es, eine Balance zwischen zielorientiertem und reaktivem Verhalten herzustellen. Ein *intelligent agent* muss in der Lage sein nicht nur zu reagieren, sondern Ziele auch irgendwann zu erreichen und sich ggf. neue Ziele setzen zu können.

In der Praxis wird oftmals keine spezifische Unterscheidung von verschiedenen Agententypen gemacht und die Begriffe werden synonym verwendet (Delamarre et al., 2017, S.400).

2.5.11 *pedagogical agents* und *intelligent tutoring systems (ITS)*

Intelligent tutoring systems (ITS) sind Computerprogramme, die ohne menschliche Intervention individualisierte Lernunterstützung leisten sollen. Oftmals wird hierzu eine Lehrkraft simuliert. Dieser Agent soll eine reale Schülerin beim Lernen unterstützen (Chou et al., 2003, S.256). Die Idee dahinter ist zum einen, dass eins-zu-eins Betreuungssituationen eine höhere Lerneffektivität zugesprochen wird (Chou et al., 2003, S.256). Außerdem sprechen administrativ-ökonomische Überlegungen für die Verwendung von *ITS*, die mit dem im [Abschnitt 2.4](#) (S. 16) beschriebenen administrativ-ökonomischem Argument für Simulationsnutzung weitestgehend identisch sind.

Bevölkert werden diese *ITS* durch einen oder mehrere *pedagogical* (oder auch *educational*) *agents*. Eine Abgrenzung zum Begriff *intelligent agents* ist, dass *pedagogical agents* zum Zweck des Unterrichts erschaffen wurden und üblicherweise eine menschenähnliche Repräsentation in grafischer Form besitzen (Baylor & Kim, 2005, S.2). *Educational agents* sind in ihrer Mensch-Maschine-Interaktion so konstruiert, dass sie menschliches Verhalten algorithmisch approximieren (Chou et al., 2003, S.258), jedoch nicht zwingend in einen *ITS*/Tutoring-Kontext eingebunden sind (Gibson, 2009a, S.673).

Traditionell wird von diesen *pedagogical agents* die Rolle einer Lehrkraft

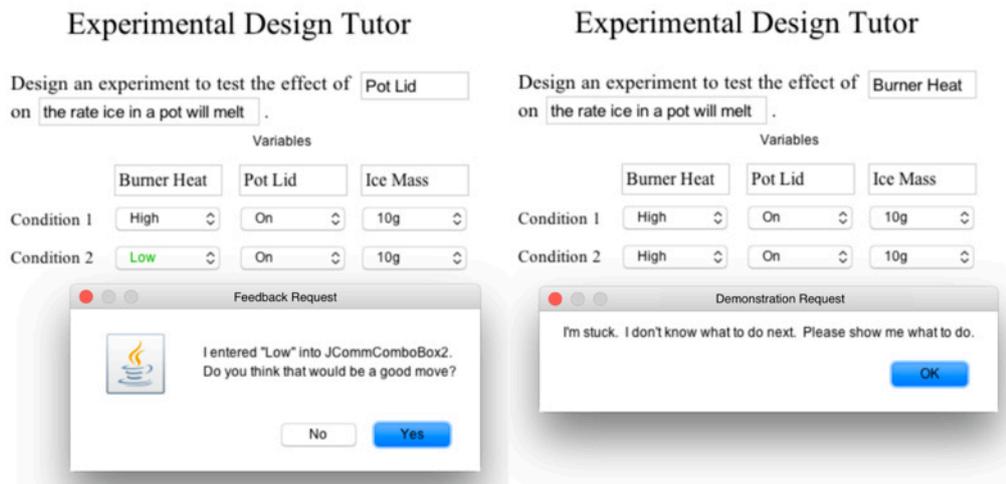


Abbildung 1: Screenshot eines *intelligent tutoring systems (ITS)*, Quelle: MacLellan (2017, S.35)

bzw. einer Expertin eingenommen, welche auf eine kognitive Wissensvermittlung ausgelegt ist. Allerdings sind auch andere Rollen erprobt worden, wie Mentorin, Lerngefährtin, Kollaborateurin, Wettbewerberin oder Unruhestifterin (Baylor & Kim, 2005, S.1f. Mohanan et al., 2017, S.7). *Pedagogical agents* sind instruktional, sie nehmen eine Lehrkräfte- oder Begleiterinnenrolle an. Die Rolle simulierter SuS – wie sie für die Erschaffung eines Simulationsmodells nützlich wäre – kommt in dem durch die Informatik geprägten Forschungsdiskurs nur vereinzelt vor (z.B. in VanLehn et al., 1994, S.2). Diese Form des Einsatzes wird auch *learning by teaching* oder *inverted ITS* genannt (Aïmeur et al., 1995), impliziert jedoch nicht den Kontext eines Klassenraums, sondern wird vielmehr als generelles *ITS*-Modell verstanden. In den letzten Jahren nahm das Interesse für dieses spezielle Forschungsgebiet tendenziell ab. Die verwendeten Simulationsmodelle des Lernens sind in der Regel stark domänenbezogen und task-orientiert (z.B. in MacLellan, 2017, S.7ff.), sie bilden keine holistischen Schülerinnenpersönlichkeiten ab, sondern sind Modelle, die ausschließlich Lernen abbilden. Diese Domänenspezifität in konkreter Modellierung und die feingranulare Lernoutput-Modellierung (z.B. in VanLehn et al., 1994, S.3) können aus Abb. 1 (S. 37) entnommen werden, welche eine Anwendung des SimStudents-Modelles zeigt.

Solche Modelle können zur Entwicklung von *ITS* verwendet werden, sie sind für die erforderliche Modellierung des in dieser Arbeit zu entwickelnden Simulationsmodells einer Schülerinnenpersönlichkeit aber ungeeignet. In

der Literatur ist deswegen vermutlich auch die Entwicklung von computer-gestützten Unterrichtssimulationen in der Lehrkräftebildung häufig vom Forschungsstrang der *ITS* und *educational/pedagogical agents* abgekoppelt.

2.5.12 Simulationsfidelität

Als Simulationsfidelität wird der Grad an Genauigkeit beschrieben, mit dem eine Simulation die Realität imitiert (Alessi & Trollip, 2001, S.233). Da Simulationsfidelität als aggregiertes Konstrukt aus verschiedenen Fidelitätskomponenten der Simulation konzeptionalisiert werden kann, ist der Begriff nicht unproblematisch (Vgl. auch Ward et al., 2006, S.258), was im folgenden kurz erläutert werden soll.

So ist beispielsweise die Fidelität der Simulationsrepräsentation von der Fidelität des Simulationsmodells zu unterscheiden (Alessi & Trollip, 2001, S.235, 257). Die Fidelität der Simulationsrepräsentation rekurriert auf die Immersionsform, bei grafischen Immersionsformen zusätzlich auf das Zusammenspiel konkreter grafischer Elemente. Grafische Elemente meinen dabei sowohl die Repräsentationsform der simulierten Figuren (z.B. Comiclook, Detailgrad) und der Umwelt, als auch die Anzahl, Form und Qualität der Animationen. Obwohl generell 3D-Umsetzungen eine höhere Fidelität der Simulationsrepräsentation zugesprochen wird (Bulu, 2012, S.156), so ist dies keine notwendige Bedingung. Eine stark ausgearbeitete 2D-Unterrichtssimulation mit vielen Animationen und individuell konstruierten Darstellungen der SuS besitzt eine höhere Fidelität der Simulationsrepräsentation als eine 3D-Simulation ohne Animationen mit lediglich einem 3D-Modell zur Darstellung aller SuS-Agenten.

Die Fidelität des Simulationsmodells hingegen beschreibt den an der Realität gemessenen Gütegrad der Approximierung des zugrundeliegenden algorithmischen Modells. Es ist demzufolge möglich eine Simulation zu entwickeln, die ein überaus komplexes Schülerinnenverhalten modelliert (hohe Fidelität des Simulationsmodells), in der Darstellung der Daten jedoch rein textgebunden ist (niedrige Fidelität der Simulationsrepräsentation). Die Fidelität des Simulationsmodells ist von der Präzision der Simulation abhängig. Präzision bedeutet in diesem Sinne wie weit die modellierten Prozesse verstanden und mathematisch exakt nachvollzogen werden können (Alessi & Trollip, 2001, S.239). Dies führt dazu, dass Modelle, die menschliches Verhalten beschreiben, nie wirklich hohe Fidelität des Simulationsmodells erreichen können, da sie auf Heuristiken angewiesen sind. Alessi und Trollip formulieren diese Problematik wie folgt:

The more chance is involved, the less precise the model is, and the harder it is to program the simulation. The extreme case is

when people are involved. Very little is really understood about individual human behavior, which makes predicting it almost impossible (Alessi & Trollip, 2001, S.239f.).

Allerdings zeigen Alessi und Trollip (2001) auch einen Ausweg aus diesem Dilemma auf, nämlich die Erschaffung einer Illusion von hoher Simulationsfidelität:

Because our knowledge of these real phenomena [menschliche Interaktionen, Anm. d. Vef.] is imprecise, it is not possible to have high fidelity models, although we may create the illusion among users that their fidelity is high (Alessi & Trollip, 2001, S.260).

Für den Zusammenhang zwischen *initial learning* (s.u.) und Simulationsfidelität gehen Alessi und Trollip (2001, S.234) immer von intrapersonell wahrgenommener Simulationsfidelität aus.

Eine weitere Komponente ist die Fidelität der Handlungsmethode (Alessi & Trollip, 2001, S.252). Diese beschreibt, inwiefern die in der Simulation ausgeführte Handlung der in der Realität durchgeführten Handlung ähnelt. Da diese Fidelitätskomponente auch ein relevantes taxonomisches Merkmal darstellt, wird sie an anderer Stelle ausführlicher beschrieben (siehe [Abschnitt 3.4](#), S. 68).

Letztlich beeinflusst auch der Grad des Feedbacks die Fidelität der Simulation. Feedback kann entweder natürlich (verzögert) oder künstlich (unmittelbar) sein. Eine Flugsimulation, die der Nutzerin mittels Hilfe-Popups die Bedeutung jedes Instruments anzeigt, besitzt weniger Fidelität, ist jedoch für eine Anfängerin besser geeignet (Alessi & Trollip, 2001, S.254ff.). Der Begriff der *psychological fidelity* ist keine abgrenzbare Fidelitätskategorie, sondern beschreibt eher summarisch, wie sehr die in der Simulation ausgeführte Tätigkeit die Anforderungen einschließt (Ward et al., 2006, S.244).

Dass diese unterschiedlichen Komponenten meist nur gebündelt verhandelt werden, führt nicht selten zu einer Überbetonung der Fidelität der Simulationsrepräsentation unter gleichzeitiger Vernachlässigung der anderen Fidelitätskomponenten, wobei es auch gegenteilige Entwicklungen gibt (Deale & Pastore, 2014).

Häufig wird ein positiver linearer Zusammenhang zwischen Simulationsfidelität und Simulationseffektivität postuliert (Ward et al., 2006; Bradley & Kendall, 2014; Hays et al., 1992). Als Erklärungsmodell dafür wird unter anderem ins Feld geführt, dass *high-fidelity* Simulationen den Transfer von trainierten Verhaltensweisen in reale Kontexte erleichtern würden (Bradley & Kendall, 2014, S.5). Zu berücksichtigen ist jedoch, dass beispielsweise Ward et al. (2006) sich ausschließlich mit Sportsimulationen, chirurgischen

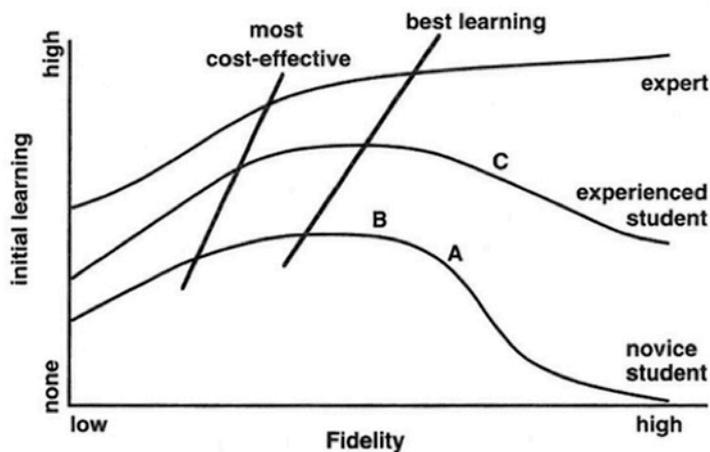


FIGURE 7.16

Hypothesized Relationship of Fidelity and Learning

By permission of the *Journal of Computer-Based Instruction* and the Association for the Development of Computer-Based Instructional Systems.

Abbildung 2: Hypothetischer Zusammenhang zwischen Simulationsfidelität und Lernoutcome, Quelle: Alessi und Trollip (2001, S.234)

Simulationen und Flugsimulationen beschäftigt, deren Simulationsziel in erster Linie ein Training motorischer Fähigkeiten umfasst. Zudem sind gerade im Bereich Simulation Metastudien mit Vorsicht zu interpretieren, denn die verwendeten Simulatoren unterscheiden sich in Features, Design und verwendeten Technologien stark; zusätzlich sind die pädagogischen Ziele häufig nicht klar definiert (Fischler, 2006, S.7). Kersh (1963) und Twelker (1967, S.202) postulieren, dass es keinen Zusammenhang zwischen Simulationsfidelität und Simulationseffektivität bzw. Lerntransfer gäbe. Diese Begründungen sind rein theoretischer Natur und rekurrieren auf die von ihnen verwendeten analogen Simulationen. Insbesondere in digitalen Simulationen ist jedoch ein starker Zusammenhang zwischen Effektivität und Fidelität zu vermuten.

Statt eines linearen Zusammenhangs zwischen wahrgenommener Simulationsfidelität und Simulationseffektivität scheint zumindest aus theoretischen Erwägungen ein komplexer Zusammenhang, abhängig von den Vorerfahrungen der Partizipantinnen, wahrscheinlicher. Alessi und Trollip (2001) schlagen die folgende Relation vor, die in Abb. 2 (S. 40) visualisiert ist.

Die Performanz innerhalb der Simulation (*initial learning*) muss von der Transferfähigkeit der Simulationserfahrung in die reale Welt (*transfer of learning*) getrennt werden (Alessi & Trollip, 2001, S.233). Während die Autoren *transfer of learning* tatsächlich noch in einem linearen Zusammenhang mit der Simulationsfidelität sehen, gehen sie bei dem *initial learning* von dem in Abb. 2 (S. 40) veranschaulichten Zusammenhang aus. Etwas klarer wird dies, wenn man die Funktionsweise an einem Beispiel veranschaulicht. Wenn eine Novizin einen *high-fidelity* Flugsimulator mit echten Bedienelementen

benutzt, so wird sie in dieser Simulation kaum etwas lernen, bestenfalls wie man Schub erzeugt und die Bremse betätigt (*initial learning* gering). Würde sie sich aber in einem realen Flugzeug befinden, so kann sie dort den Schub oder die Bremse wahrscheinlich bedienen (*transfer of learning* hoch). In einem Flugsimulator für den heimischen PC kann es sein, dass verschiedene Bedienelemente nicht vorkommen und der Abhebeprozess wesentlich vereinfacht ist. Außerdem zeigt das Programm vielleicht Hilfsfunktionen für alle Instrumente an und bietet ein Tutorial zum Start an. Es findet also ein höheres *initial learning* während der Simulation statt. Da allerdings Dinge wie Schub und Bremse in der Regel vereinfacht über die Maus oder eine Tastenbelegung gesteuert werden, findet sich die Novizin in einem realen Flugzeug vermutlich kaum zurecht und kann den Start, den sie in der Simulation erfolgreich absolviert hat, nicht in die Realität übertragen.

Für eine Novizin ist es demzufolge eher geeignet die Fidelität zu reduzieren, damit das *initial learning* steigt (Alessi & Trollip, 2001, S.234f. G. R. Girrod et al., 2007, S.215; Deale & Pastore, 2014, S.211). Dies empfiehlt sich auch für Unterrichtssimulationen, die ja menschliches Verhalten modellieren und deswegen ohnehin kein *high-fidelity* Simulationsmodell erzeugen können. Außerdem entsprächen Lehramtsstudierende als Simulationszielgruppe Novizinnen. Will man die Zielgruppe verbreitern, beispielsweise im Hinblick auf die zweite und dritte Ausbildungsphase von Lehrkräften, so schlagen Alessi und Trollip (2001, S.235) die Implementation dynamischer, adaptiver Fidelitätssysteme vor.

2.5.13 Simulationsmodelle und Konzeptionsmodelle

Um eine Computersimulation zu erstellen, benötigt man ein algorithmisches Modell, das spezifiziert, wie das System funktioniert. Dazu wird ein Modellierungsprozess durchlaufen, der Entitätstypen, deren Attribute und gegenseitige Beeinflussung beschreibt und dessen Resultat das sogenannte Simulationsmodell ist (Skrødal, 2010, S.39). Egal, ob das zugrundeliegende Modell hart deterministisch oder heuristisch/stochastisch (z.B. *hidden markov models*) arbeitet, es postuliert in jedem Falle Ursache-Wirkungs-Zusammenhänge. Dieses Simulationsmodell ist von einem Konzeptionsmodell zu unterscheiden. Konzeptionsmodelle bewegen sich auf einer höheren Abstraktionsebene, sie beinhalten eine Beschreibung eines Systems oder relevanter Teilsysteme (Skrødal, 2010, S.40), jedoch haben sie keinen algorithmisch-informations-technischen Bezug; Modellobjekte enthalten oft bidirektionale oder andere komplexe Verbindungen und Ursache-Wirkmechanismen sind – wenn überhaupt – nur so abstrakt beschrieben, dass man ihnen selten konkrete Zahlen zuordnen kann.

Es ist bei einer *educational simulation* mit ihren mannigfaltigen mit Unsicherheit behafteten menschlichen Interaktionen nicht verwunderlich, dass in der Literatur kaum Simulationsmodelle besprochen werden, allerdings ein Füllhorn an Konzeptionsmodellen existiert (Seidel, Tina, 2014). Erprobte, empirisch und theoretisch begründbare Simulationsmodelle bestehender Simulatoren aus dem [Abschnitt 4](#) (S. 71) sollen in die Konzeption des Simulationsmodells dieser Arbeit einfließen. Daneben sollen empirisch fundierte Konzeptionsmodelle herangezogen werden, um den Modellierungsprozess des interaktionalen Simulationsmodells zu realisieren.

Shannon (1998, S.10) merkt an, dass ein Großteil der Modellerstellerinnen ihre Modelle mit zu vielen trivialen Details füllen, was unnötige Komplexität und möglichen Verlust von signifikanten Ursache-Wirkungsverbindungen bewirkt. Es gilt, einen Mittelweg zwischen dieser überbordenden Komplexität (*bloated models*) und eines zu simplen, reduktionistischen Modells zu finden, was sich gerade in Hinsicht auf Modellierungen menschlichen Verhaltens als außerordentlich schwierig erweist. In einem sozialen System wie dem Klassenraum sind Entitäten, kausale Wirkmechanismen, deren Richtung und Gewichtungen der kausalen Effekte unklar (Skrødal, 2010, S.41).

In einem solchen Fall dient das Ziel der Simulation als Orientierungspunkt dafür, welche Entitätstypen ein Modell enthalten sollte. Skrødal (2010) beschreibt diesen Vorgang sehr anschaulich an einem Modell, welches die innere Funktionsweise einer Armbanduhr illustrieren soll:

Thus, if a model's purpose was to illustrate how the mechanics of the wristwatch [...] work, certain parts should already be excluded, such as the wristbands, front and back covers, the dial, and even its casing. In this instance, these components are not relevant to the inner mechanics, and would unnecessarily complicate the model if included. Although it is not always easy to distinguish the trivial parts of a system from the vital ones, early efforts in this area will be rewarded when formulating the simulation model and implementation (Skrødal, 2010, S.40).

Um nicht über- oder unterkomplexe Simulationsmodelle zu generieren ist es im Sinne des von Deale und Pastore (2014, S.208) propagierten modellzentrierten Ansatzes erforderlich, die Simulation sehr häufig durchzuspielen und iterativ zu verbessern.

Wie bereits in den Ausführungen zur Simulationsfidelität bemerkt, neigen komplexe Simulationsmodelle ebenso wie komplexe Darstellungsformen dazu, Simulationsnutzerinnen, insbesondere Novizinnen, kognitiv zu überfordern (Deale & Pastore, 2014, S.211; bereits Kersh, 1963, S.10 folgerte dies aus seinen empirischen Studien). Hierbei spricht man von dem sogenannten

cognitive overload (Sweller, 1988). Geringe Fidelität der Simulationsrepräsentation geht mit niedrigerer *presence* einher und wird als weniger motivierend wahrgenommen (Bulu, 2012, S.159; Alessi & Trollip, 2001, S.247; Psotka, 1995, S.409). Unterkomplexe Simulationsmodelle laufen Gefahr auf falschen oder zu verallgemeinernden Annahmen zu basieren und damit im schlimmsten Falle andere als die in der Realität benötigten Fähigkeiten zu schulen. Zudem basieren solche unterkomplexen Modelle häufig auf behavioristisch orientierten Reiz-Reaktions-Schemata.

2.5.14 *Entitäten, Entitätstypen und Entity-Relationship-Modelle (ERM)*

Skrødal (2010, S.8) benutzt in seiner Arbeit den Begriff Entitäten für eindeutig zu bestimmende Objekte, die im Rahmen seines Simulationsmodells Daten aufnehmen und miteinander interagieren. In seinem Fall nennt er hier die SuS, die Aufgabe und die Lehrkraft (Skrødal, 2010, S.42). Da Entitäten normalerweise auf individuelle Objekte rekurrieren (eine spezielle Schülerin, eine spezielle Aufgabe), ist hier korrekterweise von Entitätstypen oder Entitätsklassen zu sprechen. Obwohl es sehr viel Sinn macht von diesem Paradigma aus zu denken, so ist doch anzumerken, dass Begriff und Konzeption aus der informatischen Datenmodellierung kommen und somit schon den Blick auf den Gegenstand beeinflussen. Das zeigt sich unter anderem daran, dass man Entitätstypen relativ unverändert in Klassen aus der objektorientierten Programmierung übertragen kann. Auch der Autor dieser Arbeit verwendet diese paradigmatische Herangehensweise, begründet aus Ermangelung von pragmatisch gangbaren Alternativen.

2.5.15 *cognitive load*

Cognitive load bezieht sich auf die Menge an mentalen Ressourcen, die eine Person benötigt, um eine bestimmte Aufgabe zu bewältigen (Sweller, 1988; Sweller, 2011). Sweller (2011) geht davon aus, dass Lernen dann effektiver ist, wenn eine geringe kognitive Belastung vorhanden ist. Drei Arten kognitiver Belastung können unterschieden werden, nämlich intrinsische, extrinsische und germane Belastung. Die intrinsische Belastung ist die inhärente Schwierigkeit der Aufgabe, während die extrinsische Belastung durch äußere Faktoren wie die Art der Instruktion beeinflusst werden kann. Diese Art von kognitiver Belastung ist möglichst zu vermeiden. Die lernförderliche germane Belastung bezieht sich auf die mentalen Aktivitäten, die für das tiefere Verständnis und die Verarbeitung der Informationen erforderlich sind (Sweller, 2011; Sweller, 1988).

Nach dieser Cognitive Load Theory (CLT) ist es unerlässlich Simulationen komplexitätsreduziert zu gestalten, um die Partizipantinnen nicht zu überfordern. Damit wird der CLT folgend extrinsische Belastung vermieden und die Resilienz der Partizipantinnen bezüglich germaner Belastung gefördert. Um Unterrichtssimulationen in ihrer Komplexität zu reduzieren kann beispielsweise die Anzahl der SuS verkleinert werden. Eine weitere Möglichkeit ist es, bestimmte Unterrichtshandlungen nicht zu modellieren und damit auszublenken. *Cognitive load* ist bei unterschiedlich professionalisierten Personen unterschiedlich ausgeprägt. Eine langjährige Lehrkraft hat oft eine größere Menge an mentalen Ressourcen um ihre Aufgaben während des Unterrichts zu bewältigen. Nach diesen Überlegungen müssen Simulationen eine spezifische Zielgruppe in den Fokus nehmen oder bezüglich *cognitive load* adaptiv gestaltet sein.

3 Taxonomie der Unterrichtssimulationen

Während für Simulationen und Simulatoren bereits einige Taxonomien (z.B. Alessi & Trollip, 2001) existieren, so gibt es solche explizit für Unterrichtssimulationen bislang noch nicht. Die Taxonomie von Alessi und Trollip (2001) ist zudem aktualisierungs- und in einigen theoretischen Konstrukten verbesserungsbedürftig.

Nach reiflicher Abwägung wurde die Entscheidung getroffen eine dreidimensionale Skala einzuführen, mit den Dimensionen Granularität, Immersionsstufe und Interaktionsmodellierung. Die Begründung für die Auswahl dieser Skala, ergänzt durch in der Literatur vorgeschlagene alternative Taxonomien, wird in [Abschnitt 3.4](#) (S. 68) geliefert.

3.1 Nach Granularität

Granularität beschreibt, wie kleingliedrig oder umfassend eine Simulation gestaltet ist. Dies kann von der Modellierung einer einzelnen mathematischen Aufgabe bis zur Darstellung eines vollständigen Schulalltages, inklusive Pausen und Elterngesprächen, gehen. Bislang findet in der Forschungsdiskussion eine Kategorisierung anhand verschiedener Simulationsgranularitäten kaum statt. Während zugestanden wird, dass Granularität und Wirkungsbereich einer Simulation die präsentierten Handlungsalternativen maßgeblich beeinflussen (Aldrich, 2009, S.55), wird Granularität bislang als Taxonomiemöglichkeit nicht in Betracht gezogen. Eine Ausnahme bilden VanLehn et al. (1994, S.3ff.), die sich, wie bereits ausgeführt, mit intelligenten Agentensystemen beschäftigen und damit einem anderen Forschungsstrang zuzuordnen sind. Auch Gibson (2009a, S. 685) führt in der Erklärung seines algorithmischen Lernmodells die Wichtigkeit des Granularitätsniveaus (*grain level*) an; allerdings wird es in diesem Kontext nicht zu explizit taxonomischen Zwecken verwendet. Granularität ist deswegen so zentral, da unterschiedliche Granularitäten erheblich Struktur und Umfang des zugrundeliegenden Simulationsmodells mitbestimmen. Oft geben Simulationsgranularitäten schon einen Hinweis auf eine zu präferierende Interaktionsmodellierung.

3.1.1 *task simulation (TS)*

Task simulations sind die in ihrer Granularität am feinsten aufgelösten Simulationen und betrachten die Mikroebenen der Lehrkraft-Interaktion. Diese Klassifikation findet sich unter anderem in Ward et al. (2006, S.243f.) als *simulated task environment*. Wichtig zu erwähnen ist, dass dort *task* nicht immer als ‚Aufgabe‘, sondern als ‚Tätigkeit‘ zu übersetzen ist. Damit um-

fasst die Definition nicht nur klassenraumbezogene Interaktionen, sondern auch beispielweise Lehrkraft-Eltern-Gespräche. Die zu modellierende Tätigkeit ist inhaltlich klar abgegrenzt und umfasst nur eine modellierte Zeit von einigen Minuten. *Task simulations* sind oft stark domänenspezifisch und können Wissen sehr viel genauer als Simulationen anderer Granularitätsstufen modellieren. Dafür sind *task simulations* sogenannte ‚one-trick ponies‘, also nur für eine überaus spezifische Aufgabe geeignet und in der Regel kaum wiederverwendbar. Das zugrundeliegende Interaktionsmodell berücksichtigt keine Verhaltensweisen, die den Rahmen der modellierten Tätigkeit sprengen würden, weswegen bei der Modellierung von Unterrichtsaufgaben die psychologisch-emotionale Modellierung der Schülerinnenagenten im Gegensatz zur Wissensverarbeitung oft unterkomplex oder schlicht nicht vorhanden ist.

Bereits existierende *task simulations* sind beispielsweise die Anleitung zur Lösung einer Geometrie-Beweisführungs-Aufgabe (Chieu & Herbst, 2011) oder das Führen eines Schülerinnen-Lehrkraft-Gesprächs zur Prävention oder Früherkennung psychischer Störungen (Albright et al., 2013b; Albright et al., 2013a).

3.1.2 *school lesson simulations (SLS)*

School lesson simulations beziehen sich auf die Meso-Ebene der Lehrkraft-Interaktionen, sie haben zum Ziel eine Unterrichtsstunde oder einen logisch zusammenhängenden Lernblock zu modellieren. Wird diese Art der Granularität gewählt, so ist es wichtig, dass Abweichungen vom didaktischen Ziel der Unterrichtsstunde im Simulationsmodell berücksichtigt werden. Ansonsten wird ein wichtiger Aspekt einer Schulstunde, nämlich das *classroom management*, nicht berücksichtigt und die eingeübten Fähigkeiten wären nicht auf eine reale Klasse übertragbar. Damit wird das Simulationsmodell ausgeweitet, da es auch emotional-psychologische Faktoren berücksichtigt, allerdings wird es damit probabilistischer und somit ungenauer.

3.1.3 *school teacher simulation (STS)*

School teacher simulations beschreiben die Makroebene der Lehrkraft-Interaktionen. Von der Einstellung, über die Stundenvorbereitung, den eigentlichen Unterricht oder bis zum Gespräch mit der Schulleiterin können alle Interaktionen in der Simulation vorkommen. Aufgrund der eigentlich nicht modellierbaren Komplexität dieses Granularitätsniveaus verwenden die meisten makrogranularen Simulationen einfache deterministische Entscheidungsbaum (Typ 1). Die einzelnen Zustände und Verbindungspunkte des Entscheidungsbaums können nicht empirisch, sondern bestenfalls theoretisch, begrün-

det werden. Dies führt dazu, dass die Simulationen eher den Charakter einer ‚case study‘ haben.

Ihre Stärke entfalten diese Simulationen erst im diskursiven didaktischen Zusammenhang, wenn über Alternativhandlungen und Situationen reflektiert werden kann.

3.2 Nach Immersionsstufe

Die Immersionsform einer Simulation zielt auf die technische Ausgestaltung des Simulators ab. Tatsächlich ist für eine gelingende Simulationserfahrung die *presence* und die wahrgenommene Simulationsfidelität aber wesentlich entscheidender. Das Problem mit diesen beiden Konstrukten ist jedoch, dass *presence* sich auf intrapersoneller Ebene abspielt und genauso wie Simulationsfidelität schwer quantifizierbar ist. Eine Abgrenzung unterschiedlicher Immersionsformen ist aber problemlos möglich.

Alessi und Trollip (2001, S.247) verhandeln den Aspekt der Immersionsformen unter dem Schlagwort *modes of presentation*, geben ihm jedoch nicht allzu großen Raum in ihrer Taxonomie. Als Beispiele werden Text, Bild, Stimme, Animation, Video oder eine Kombination aus diesen Immersionsformen genannt. Für Alessi und Trollip (2001) beeinflusst der Modus der Präsentation in erster Linie die Motivation der Nutzerinnen, die Simulationsfidelität und die Userfreundlichkeit der Simulation. Eine fast identische Konzeption für webbasierte Lernplattformen findet sich bei Hsu et al. (2009, S.92) und auch Dieker, Kennedy et al. (2014) nutzen Immersionsformen als Distinktionsmerkmale, freilich ohne dass es in beiden Studien explizit um die Entwicklung einer Taxonomie geht.

3.2.1 (Hyper-)Text-based simulation

Diese Immersionsform nutzt primär Text zur Gestaltung der Simulationserfahrung. Die verwendete Technik ist hierbei nicht von Bedeutung, allerdings sind viele Simulationen webbasiert. Dies ermöglicht es, auch Bilder und andere audiovisuelle Medien einzubinden und mit den Textanteilen zu kombinieren. Reine textbasierte Unterrichtssimulationen ohne die Einbindung grafischer Elemente existieren kaum noch, so dass auf eine gesonderte Einordnung rein textbasierter Simulationen verzichtet werden kann. Textbasierte Simulationen bedingen nicht zwangsläufig eine narrative Struktur oder Ereignisbäume – auch tabellarische oder andere datenbasierte Simulationen (z.B. Cook School District Simulation, siehe [Abschnitt 4.5](#), S. 80) zählen zu den textbasierten Simulationen.

3.2.2 *2D (animated) graphical simulations*

Die Grenze zwischen den mit Bildern angereicherten text-basierten Simulationen und *graphical simulations* ist fließend. Das wesentliche Distinktionsmerkmal ist, dass grafische Simulationen visuelle Bedeutungsdimensionen transportieren, die für die Nutzung der Simulationen zentral ist – die verwendeten Grafiken können also nicht entfernt werden ohne die Simulation unbenutzbar zu machen. Dazu wird im Unterrichtskontext oft die Blickperspektive der Lehrkraft auf die Klasse wiedergegeben, welche entweder perspektivisch verzerrt dargestellt oder stark erhöht wird, um Überlappungen der abgebildeten Schülerinnenagenten zu vermeiden (Vgl. Aha! Classroom in [Abschnitt 4.1](#), S. 71, simSchool in [Abschnitt 4.16](#), S. 105 und [Abschnitt 4.17](#), S. 121). Alternative Klassendarstellungen existieren ebenfalls, beispielsweise die Abbildung von SuS-Portraits, welche je nach diskretem Zustand eine andere Mimik zeigen (Vgl. VCS in [Abschnitt 4.23](#), S. 134). Zeigen SuS-Agenten eine andere Mimik, so ist dies eine Bedeutungsebene, die nicht weggelassen werden kann, was den Simulator als *2D graphical simulation* klassifiziert. Über die Qualität der Animationen wird in dieser Taxonomie keine Aussage getroffen. Dominierende Umsetzung dieser Immersionsform sind statische Bilder für SuS und die Umgebung, welche sich bei Veränderungen der Simulationsvariablen in andere statische Bilder transformieren. Echte zweidimensionale Animationen im Sinne eines Trickfilms, nämlich, dass mehrere Keyframes und Inbetweens benutzt werden, um Bewegungen für das menschliche Auge zusammenhängend darzustellen, werden hingegen selten benutzt. Dies ist wenig verwunderlich, da eine solche Ausgestaltung nur mit einer marginalen Erhöhung der Simulationsfidelität einhergeht, aber große Kosten verursacht.

Simulatoren, die lediglich illustrative – und damit für die Simulationsnutzung nicht zwangsläufig erforderliche – Fotos von SuS im Rahmen von Schülerinnenbeschreibungen inkludieren, aber konkrete Situationen in textbasierter Form simulieren, zählen zu den textbasierten Simulationen.

3.2.3 *3D graphical simulations*

In einer *3D graphical simulation* werden Agenten als dreidimensionale Modelle in einer dreidimensionalen Umgebung dargestellt. Diese 3D-Modelle müssen im Vorfeld modelliert, textuiert und animiert werden, was erhebliches Know-How erfordert und durch seine Komplexität viele Arbeitsressourcen bindet. Dieser Prozess soll, weil er so wichtig für die Konstruktion dieser Simulationsform ist, kurz erläutert werden. Zunächst erfolgt der Modellierungsprozess, in dem in einem dreidimensionalen Koordinatenfeld Punkte, sogenannte Vertices gesetzt werden, die miteinander verbunden werden, um

Flächen, sogenannte Polygone, zu erzeugen. Mittels verschiedener Techniken der Computergrafik entstehen aus diesen Flächen komplexe Objekte. Anschließend wird via Projektion (*UV mapping*, *UV unwrapping*) eine zweidimensionale Grafik auf das ansonsten einfarbige 3D-Modell gelegt. Diesen Prozess kann man sich ähnlich wie das Einpacken eines Geschenkes vorstellen. Danach folgt das sogenannte *rigging*. Hierzu wird eine abstrahierte Knochenstruktur (*bones*) in dem modellierten Charakter erstellt und dann Schwellenwerte festgelegt, wie weit diese Knochen die Vertices der modellierten Figur beeinflussen. Letztlich folgt die Animation, die entweder aus dem manuellen Animieren der Knochenstruktur besteht oder aber mittels *motion capturing* arbeitet. *Motion capturing* ermöglicht es über klassischerweise zwei Kameras mit Markern¹ oder anderen Tiefensensoren Bildpunkte der Realität in ein digitales dreidimensionales Koordinatensystem zu überführen. Während früher dafür professionelles Equipment (wie spezielle Anzüge mit weißen Leuchtpunkten) benötigt wurde, ist heutzutage die Verwendung eines PrimeSense-Tiefensensors via Microsoft Kinect ausreichend, um *motion capturing* durchzuführen (Chini et al., 2016, S.4). Im Film- oder Computerspielsektor und in einigen wissenschaftlichen Anwendungen werden aufgrund ihrer Genauigkeit weiterhin Punktsysteme zum Tracking verwendet. Selbst in der Form eines *capturings* über Kinect ist der Animationsprozess sehr komplex.

Wenn die *motion capturing* Daten gesammelt worden sind, kann man sie auf die *bones*-Struktur übertragen, welche ihrerseits die Vertices der Figur steuert. Eine Animation von Gesichtsmuskulatur, um Emotionen darzustellen, funktioniert prinzipiell auf ähnliche Weise, bedingt aber die Verwendung von auf das Gesicht aufgeklebten oder gezeichneten Punkten. Es gibt mittlerweile Systeme, die durch Hilfe von *machine learning* nur eine Webcam benötigen, um Gesichtsanimationen auf 3D-Figuren zu übertragen, jedoch sind diese Anwendungen noch recht neu und verschiedenen Einschränkungen unterworfen. Es ist daher nicht verwunderlich, dass viele Unterrichtssimulationen sich in (Gesichts-)Animationen auf einen geringen Komplexitätsgrad beschränken.

Eine interessante Entwicklung neuerer Zeit ist die Möglichkeit der sogenannten prozeduralen Generierung von Figuren und Texturen. Hierbei entwirft ein Programm nach Vorgaben der Nutzerin automatisiert 3D-Modelle oder Texturen. Datengrundlage dieser Artefakte sind Uservorgaben, vordefinierte Basis-Modelle und Algorithmen.

Teams, die Simulationen entwickeln, sowohl in der freien Wirtschaft als

¹Neueste durch KI unterstützte *motion capturing* Methoden erfordern keine Marker und sogar nur eine Kamera um *motion capturing* durchzuführen, wobei die Technik weniger genaue Ergebnisse liefert und z.T. noch prototypisch ist. Näheres zu diesen Entwicklungen ist bei X. Wang et al. (2019) zu finden.

auch in der Forschung, lagern die Arbeit zur Erstellung von 3D-Modellen, Texturen und Animationen gerne an externe Dienstleister aus. Dafür müssen aber sämtliche Faktoren spezifiziert sein: Von der Anzahl der individuellen SuS-Agenten, über deren spezifische Ausgestaltung (beispielsweise Geschlecht, Kleidung, Frisur, etc.) bis hin zur Spezifikation der Handlungen, zu denen die Agenten fähig sein sollen. Es kommt selten vor, dass man diese Faktoren bereits am Anfang der Simulationsentwicklung vollständig spezifizieren kann, da eine iterative Verbesserung der Simulation bzw. des Simulationsmodells mit wechselnden Handlungsanforderungen einhergeht. Daher ist es wichtig, entweder sämtliche Simulationsmodelle vor Entwicklung bereits fertig ausgearbeitet zu haben oder in ständigem Kontakt mit externen Dienstleistern zu stehen, damit beispielsweise neue Animationen bei Bedarf ergänzt werden können.

Eine Verwendung von vorgerenderten *2D-Sprites* konstituiert keine grafische 3D-Simulation.

3.2.4 *interactive video-based simulations (IVBS)*

Interactive video-based simulations (IVBS) sind Simulationen, in denen Teile der Simulationserfahrung über eine Videoeinspielung präsentiert werden. Üblicherweise wird den Partizipantinnen dabei über ein Einstiegsvideo das Szenario geschildert. Danach folgt eine Präsentation von Handlungsalternativen, wobei der Ausgang zunächst unklar ist. Wie die Simulation im folgenden gestaltet ist, hängt von der Interaktionsmodellierung ab, die aufgrund des Fakts, dass man für jedes modellierte Verhalten ein eigenes Video benötigt, begrenzt ist.

Eine der ältesten Simulationen, nämlich das Inner-City Simulation Laboratory nutzt diese Immersionsform und kombiniert u.a. Videos als Stimulus mit anschließender realer Handlungsdurchführung (Interaktionsmodellierung des Typs 8) als Rollenspiel (siehe [Abschnitt 4.9](#), S. 92).

Eine andere Form der Interaktionsmodellierung ist für jede mögliche Handlungsalternative ein Ergebnisvideo bereitzustellen, so dass ein deterministischer Entscheidungsbaum entsteht (Interaktionsmodellierung des Typs 1). Ein Beispiel dafür ist die Simulation Classroom Discipline, die von der University of Alberta produziert wurde und als Trägermedium eine Laserdisk verwendet. Es gibt vier Simulationsszenarien im Sinne von *critical incidents* bei denen per Multiple-Choice eine präferierte, vorgegebene Handlungsalternative gewählt werden kann. Geht man alle Handlungspfade ab, ergeben sich 148 mögliche Outcome-Szenarios. (Overbaugh, 1995)

Theoretisch denkbar sind auch Interaktionsmodellierungen des Types 2 und 3, sie konnten allerdings in der untersuchten Literatur nicht gefunden wer-

den. Letztlich ist wichtig zu erwähnen, dass das reine Anschauen von Videos mit anschließender Reflexion oder Aufgabenbearbeitung (Video-Annotation, Beantwortung von Multiple-Choice Fragen) ein wichtiges Mittel in der Aus- und Weiterbildung von Lehrkräften darstellt (G. Blomberg et al., 2014), aber keine Simulation konstituiert.

3.2.5 *Virtual Reality (VR) / Augmented Reality (AR) simulations*

Virtual reality (VR) wird ein großes Potential im Simulationseinsatz zugeschrieben, speziell in der Lehrkräftebildung (Billingsley & Scheuermann, 2014, S.270). Wird von VR gesprochen, so gibt es zwei konkurrierende Verständnisebenen, zum einen die theoretische, zukünftige, noch nicht vorhandene Funktionen imaginierende, zum anderen die technisch-praktische, die auf die Anwendung existierender Systeme abzielt.

Die technisch-praktische Verständnisebene deckt sich weitestgehend mit dem umgangssprachlichen Begriff von VR, nämlich der Verwendung von sogenannten *head-mounted displays (HMD)*. Diese HMD besitzen die Form von großen, abgedichteten Brillen, in die kleine Monitore integriert sind. Sie bestehen aus einer abdunkelnden Polsterverkleidung und werden mit einem elastischen Spanngurt am Gesicht befestigt. Setzt man die Brille auf, kann man, anders als bei *augmented reality*, seine reale Umgebung nicht mehr wahrnehmen.² Die Monitore erzeugen für jedes Auge ein passendes Bild, welches durch ein Linsensystem an die Augen weitergeleitet wird. Damit entsteht eine stereoskopische Bildwahrnehmung. Kopfbewegungen werden durch verschiedene Techniken (z.B. Infrarot) sehr akkurat in ein 3D-Modell übertragen, so dass man sich in einer virtuellen Welt umsehen kann. Diese HMD-Technik wird bis zu einem gewissen Grad synonym zu VR verwendet, was den Begriff aber zu stark einengt. Auch andere technische Darstellungsformen wie die DOME/CAVE sind als VR zu bezeichnen.

DOME ist eine Kurzform von *fulldome* („Ganzkuppel“). Bei dieser Technik wird eine Deckenkuppel mit mehreren Projektoren geschickt so angestrahlt, dass der Eindruck einer virtuellen Umgebung entsteht. Die Technik wird seit Mitte der 1980er Jahre in Planetarien eingesetzt, hat jedoch auch andere Anwendungsszenarien. Das Akronym CAVE steht für *Cave Automatic Virtual Environment* („Höhle mit automatisierter, virtueller Umwelt“). CAVE-Systeme unterscheiden sich von DOME-Systemen im Wesentlichen dadurch,

²Einige neuere Modelle von VR-Brillen enthalten auch Kameras, welche die reale Umgebung auf den Bildschirmen der VR-Brille anzeigen können und somit in gewisser Hinsicht eine VR/AR Mischform darstellen.

dass keine Kuppel, sondern einzelne Wände angestrahlt werden oder mittels wandfüllenden LED-Bildschirmen die Illusion einer virtuellen Umwelt erzeugt wird.

Weder bei den HMD noch bei DOME/CAVE handelt es sich um neue Techniken. Was sich allerdings geändert hat, ist, dass diese Techniken nun günstiger herzustellen sind und auf den Massenmarkt drängen, da sie ausgereift sind (Lugrin et al., 2016, S.6). Die meisten Probleme in Bezug auf Kinderkrankheiten, wie die sogenannte Simulatorkrankheit, wurden durch Techniken wie akkurates Verfolgen der Kopfbewegungen, ausreichende Pixeldichte und Framerate weitestgehend gelöst, so dass man von einer qualitativen Verbesserung sprechen kann (Bailenson et al., 2008, S.105; Lugrin et al., 2016, S.18).

Wird proklamiert, dass VR „alle menschlichen Sinne“ stimulieren kann (Brown, 1999, S.310), so wird von einer Anwendung von VR ausgegangen, die so vollumfänglich nicht existiert bzw. noch in den Kinderschuhen steckt. VR ist ein Medium, das auf Sehen und Hören fokussiert. Taktile Wahrnehmung wird nur sehr rudimentär über Force-Feedback simuliert. Die gängige Umsetzung in VR-Headsets ist, dass die Controller, die die Nutzerin verwendet, bei bestimmten Aktionen vibrieren. Ein echtes Anfassen von Gegenständen ist dies freilich nicht. Zum Zeitpunkt dieser Arbeit liegen schon vereinzelt Prototypen vor, die dies über einen ‚Haptischen Handschuh‘ (<https://haptx.com>) zumindest für den sensorischen Bereich der Hände realisieren. Zur Zeit können solche Prototypen noch nicht erworben werden, sie haben noch keine Marktreife erreicht. Ein Patent, das sich mit Ganzkörperanzügen für haptisches Feedback via Exoskelett beschäftigt, wurde 2015 eingereicht (Rubin & Crockett, 2015), allerdings stehen kommerzielle Anwendungen bis heute aus. Propriozeption, also Körperempfindung, kann mittels *off-the-shelf* HMD-Technologie nur eingeschränkt durch Tracking der Hände bzw. Handcontroller erreicht werden. Eine freie Bewegung im Raum ist entweder nicht oder nur in einem bestimmten Tracking-Bereich (typischerweise maximal 2x3 Meter) möglich. Es gibt vereinzelt schon Ansätze dies über sogenannte *virtual-motion*-Plattformen zu erweitern. Diese Plattformen sind Käfige, in denen die Anwenderin mittels Gürteln oder in einem Außenring fixiert ist und somit nur auf der Stelle läuft. Aktuell populäre marktreife, aber nur auf Nachfrage zu erwerbende, *virtual-motion*-Plattformen sind WalkOVR (<http://www.walkovr.com/>) und Virtuix Omni (<http://www.virtuix.com/>). Beide *virtual-motion*-Plattformen werden über eine konkave Bodenplatte gesteuert, über die man mehr hinübergleiten muss, als dass man tatsächliche Schritte durchführt. Da man ständig fixiert ist und nicht die Möglichkeit hat sich in anderen Positionen fortzubewegen – z.B. zu kriechen – ist auch diese Lösung nur eine, wenn auch durchaus beeindruckende, Approximati-

on von Propriozeption. Echte 360 Grad Laufbandlösungen, wie Omnifinity (<http://omnifinity.se/>), sind zum Zeitpunkt der Erstellung dieser Arbeit noch nicht marktreif.

Andere Wahrnehmungsformen wie olfaktorische, gustatorische Wahrnehmung, sowie Nozizeption und Thermozeption sind entweder nicht vorhanden oder Randerscheinungen in Prototypform.

Augmented Reality unterscheidet sich von Virtual Reality dadurch, dass die reale Umgebung mit virtuellen Entitäten angereichert wird. Im Gegensatz zu den abschirmenden, hermetisch geschlossenen VR-Brillen, sind AR-Brillen so konstruiert, dass durchsichtiges Material, ähnlich wie bei einer normale Brille, verwendet wird. Computergenerierte Bilder überlagern die Realität, so dass der Eindruck der Verschmelzung von Realität und Computergrafik entsteht.

3.2.6 Analoge Simulationen

Eine analoge Simulation bedient sich keinerlei digitaler Hilfsmittel. Es existiert eine Überschneidung bis hin zur Synonymie zum Rollenspiel und zum *microteaching/microteaching-peerteaching*, falls die Interaktionsmodellierung vom Typ 8 ist. Für eine genauere Beschreibung sei an [Abschnitt 2.5.8](#) (S. 32) verwiesen.

Ein Beispiel für eine analoge Simulation mit einer Interaktionsmodellierung des Typs 8, die kein Rollenspiel oder *microteaching/microteaching-peerteaching* darstellt, könnte das bei de Jong et al. (2012) geschilderte Szenario sein. In diesem Szenario wurden Studierende mit einem inszenierten Streit zwischen einem Dozierenden und einem Kollegen konfrontiert. Während der Dozierende versuchte sein Seminar zu halten, wurde er mehrfach von seinem Kollegen unterbrochen, der ihn aufforderte wichtige Dokumente zu unterzeichnen. Daraufhin kam es zu heftigen Beschimpfungen durch den Dozierenden, der anschließend mit dem Kollegen den Raum verließ, um die Situation angeblich zu klären. Die Studierenden, die nicht eingeweiht waren, dass es sich um eine Inszenierung handelte, mussten nun Fragen zu ihren emotionalen Zuständen beantworten (de Jong et al., 2012, S.37). Es ist durchaus fraglich, ob man die geschilderte Begebenheit als Simulation klassifizieren sollte, selbst wenn dies de Jong et al. (2012) tun, begründet dadurch, dass es sich um ein inszeniertes Szenario handelt. Es scheint sinniger diese Begebenheit als psychologisches Experiment oder Versuch zu klassifizieren, da in einer Simulation alle Partizipantinnen darüber informiert sind, dass sie an einer Simulation teilnehmen.

Theoretisch sind auch Interaktionsmodellierungen des Typs 1 oder des Typs 2 vorstellbar, solche Simulationen kamen in der untersuchten Literatur

aber nicht vor.

3.3 Nach Interaktionstyp

Jede Simulation, egal welchen Typus', schränkt verfügbare Handlungsalternativen ein und bevorzugt bestimmte Aktionen als relevant, um das System zu modellieren (Aldrich, 2009, S.52). Jede Simulation benötigt ein Modell, anhand dessen die Interaktion modelliert wird (Gibson, 2009a, S.674). Am deutlichsten treten diese Modelle innerhalb digitaler Simulationen zutage, da sie offen als Algorithmen erkennbar sind. Bevor speziell auf diese Modelle eingegangen wird, sollen zunächst die nur scheinbar ‚modellfreien‘ rollenspielähnlichen Interaktionsmodellierungen durchdacht werden.

Selbst bei Interaktionsmodellierungen des Typs 6 bis 8, die menschliche Interaktionen erfordern, gehen die Simulationen von einem Modell der Wirklichkeit aus. Dieses Modell kann entweder explizit sein, indem Rollenkarten oder Skripte verwendet werden, welche spezifizieren, welches Verhalten oder welche Art von SuS dargestellt werden soll. Zudem können Personen, die in dem jeweiligen Modell als human-in-the-loop (siehe [Abschnitt 3.3.6](#), S. 64) agieren, geschult werden. Findet dies statt, so werden diese Personen dazu angehalten, sich so zu verhalten, wie es das entwickelte Modell vorsieht. Werden hingegen keine Vorgaben gemacht, so ist nach wie vor ein Modell enthalten, allerdings ist es nur ein implizites, das die an der Simulation beteiligten Personen anhand eigener Erfahrungen individuell konstruieren. Dieses naive Modell ist deswegen problematisch, weil es intrapersonell differenziert und damit nicht kontrollierbar ist. Mit anderen Worten auf den Unterrichtskontext bezogen: Jeder versteht etwas anderes darunter, wie sich SuS verhalten. Im schlimmsten Fall kann es zu Verhalten innerhalb der Simulation führen, die den intendierten didaktischen Zielen entgegenlaufen. Eben aus diesem Grunde arbeiten fast alle rollenspielähnlichen Simulationen mit verschiedenen Vorgaben, um die Interaktionen zu beeinflussen (Alinier, 2010, S.18). Letztlich lässt sich das Grundproblem der intrapersonellen Denk- und Handlungsmodelle damit nur abmildern und nicht ganz lösen.

Bei digitalen Simulationen, die Interaktionsmodellierungen des Typs 1–5 nutzen, stellt sich zunächst die Frage, inwieweit menschliche Handlungen überhaupt algorithmisch darstellbar sind. Menschliches Verhalten ist inhärent mit Unsicherheit behaftet und damit nicht in eine kausale Form übertragbar. Die Lösung dieses Dilemmas besteht darin nicht zu proklamieren echtes menschlichen Handeln zu modellieren, sondern lediglich eine Approximation zu suchen, die sich real genug anfühlt, um die didaktischen Ziele der Simulation zu erreichen.

Außerdem ist zu diskutieren, welche und wie viele Handlungsalternativen

ausgewählt werden können. In Simulationen ist immer ein Spagat zwischen zu vielen und zu wenigen Handlungsalternativen erforderlich. Zu viele Handlungsmöglichkeiten überfordern die Nutzerinnen und verkomplizieren das Erreichen der Lernziele (Alinier, 2010, S.16; Gibson, 2007), zudem müssen die Information über Handlungsmöglichkeiten auch durch die Simulationsnutzerinnen verarbeitet werden. Bei zeitkritischen Simulationen, wie es die Simulation einer Klassenrauminteraktion ist, führt dies im Extremfall dazu, dass die für diese Interaktionen typische Unmittelbarkeit in der Simulation nicht mehr zum Ausdruck kommt. Sind die Handlungsalternativen jedoch zu beschränkt, verliert die Simulation an wahrgenommener Fidelität, sie wird nicht mehr als authentisch erlebt (Aldrich, 2009, S.56; Alessi & Trollip, 2001, S.260). Zudem müssen die Handlungsmöglichkeiten je nach didaktischem Ziel beschränkt werden (Aldrich, 2009, S.420).

Bislang gibt es in der Literatur noch keine explizite Taxonomie von Interaktionsmodellierungen. Die einzig auffindbare recht grobe Einteilung nimmt Alinier (2010) vor, der zwischen „*programmed scenarios*“ und „*on-the-fly*“-Steuerungen unterscheidet. Damit wird bezeichnet, inwieweit die Versuchsleiterinnen der Simulation in die Interaktionsstruktur während der Simulation eingreifen (Alinier, 2010, S.20).

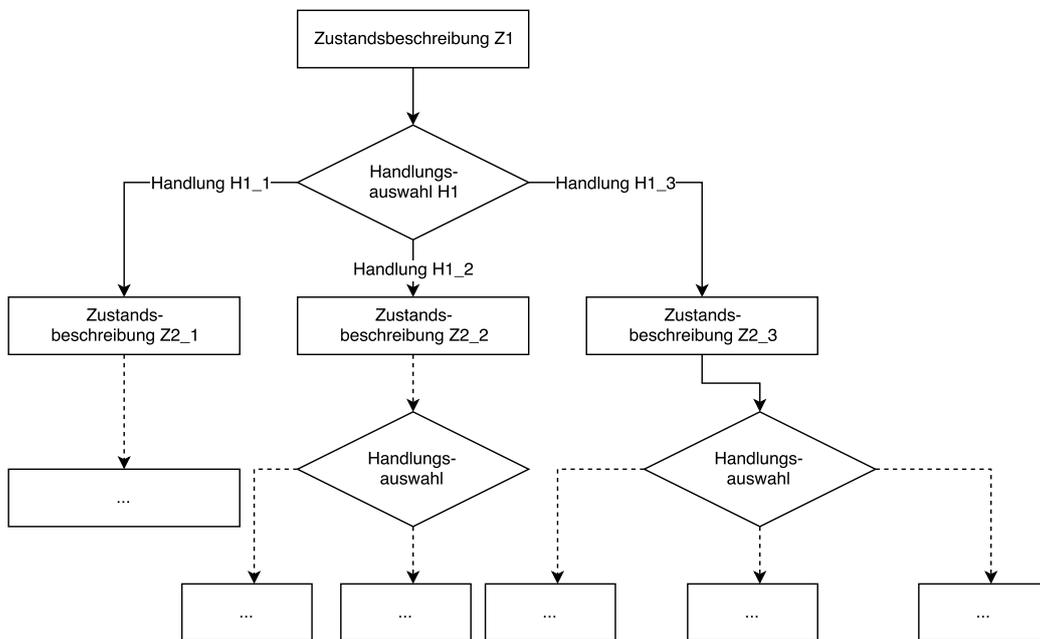


Abbildung 3: Beispielhafte Veranschaulichung eines deterministischen Entscheidungsbaums, Quelle: Autor, adaptiert von Deale und Pastore (2014, S.199)

3.3.1 Typ 1 – Deterministischer Entscheidungsbaum

Ein deterministischer Entscheidungsbaum ist eine Form der Interaktionsmodellierung in der jede Handlung immer zu einem im Vorfeld festgelegten Ergebnis führt. Synonyme Begriffe sind „*deterministic simulation*“ (Skrødal, 2010, S.28), „*branching stories*“ (Aldrich, 2009, S.15) und „*logic simulations*“ (Alessi & Trollip, 2001, S.238). Fischler (2006) beschreibt diese Art der Modellierung durch die Analogie der „Choose your own adventure“-Bücher:

In these novels, readers were given the choice to, for example, turn to page 102 if they wanted the protagonist to drive to Jersey to look for his twin brother, or turn to page 78 if they wanted him to stay in Missouri and marry his high school sweetheart. In effect, the reader had some control over the plot when given choices at pivotal points in the storyline (Fischler, 2006, S.80).

Ein deterministischer Entscheidungsbaum unterscheidet sich nicht von diesem Verfahren, er ist lediglich digital umgesetzt. Eine grafische Darstellung der Funktionsweise ist in Abb. 3 (S. 56) zu finden.

Der größte Vorteil dieser Implementation ist, dass die Simulation programmiertechnisch einfach umzusetzen ist. Außerdem sind auch alternative Sze-

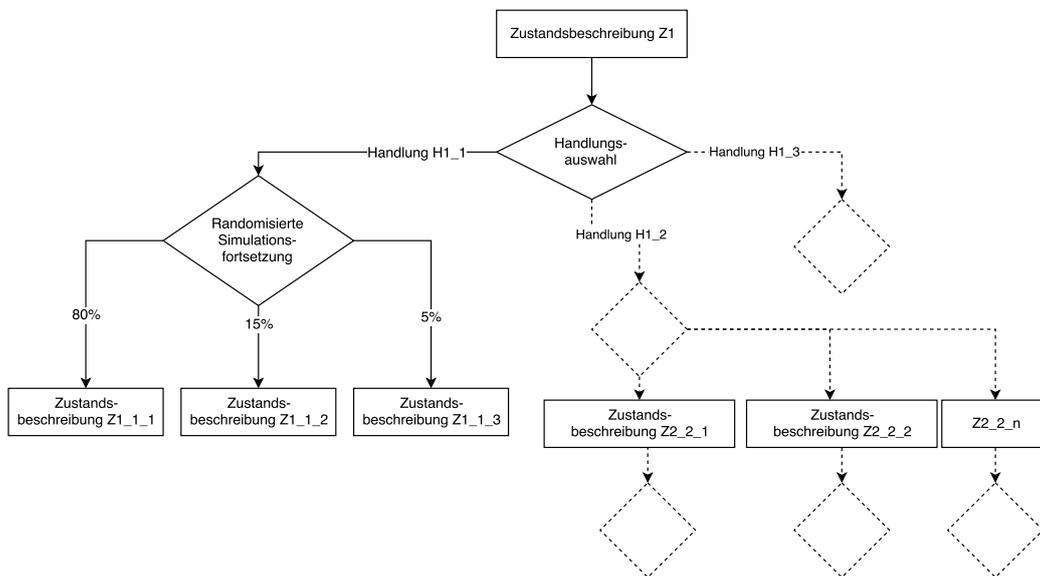


Abbildung 4: Beispielhafte Veranschaulichung eines randomisierten Entscheidungsbaums, Quelle: Autor

narien relativ einfach zu entwickeln, die stark auf fachspezifische Probleme zugeschnitten werden können. Editoren sind einfach zu konstruieren, in denen eine große Spanne an Szenarien entwickelt werden können. Sämtliche Granularitäten können von deterministischen Entscheidungsbäumen modelliert werden.

Größtes Problem ist jedoch die ökologische Validität. Das Verhalten von SuS wird aufgrund empirisch nicht gesicherter Vorannahmen und anekdotischer Evidenz ad hoc konstruiert. Lerntheoretisch fragwürdig ist, dass behavioristische Annahmen zugrunde liegen und ein Reiz stets dieselbe Reaktion hervorruft. Ein zentrales Merkmal von Unterricht, nämlich dessen Unvorhersagbarkeit, wird nicht berücksichtigt (Kunter & Trautwein, 2013, S.115).

3.3.2 Typ 2 – Entscheidungsbaum mit einfacher Randomisierung (Stochastische Simulation)

Ein Entscheidungsbaum mit einfacher Randomisierung ist ähnlich aufgebaut wie die Interaktionsmodellierung des Typs 1, mit dem Unterschied, dass es mindestens eine Stelle im Entscheidungsbaum gibt, an der das Ergebnis einer ausgewählten Handlung vom Zufall abhängig gemacht wird. Üblicherweise werden einige Handlungsergebnisse präferiert, andere kommen seltener vor. Ein synonym verwendeter Begriff zu dieser Interaktionsmodellierung lautet „Stochastische Simulation“ (Alessi & Trollip, 2001).

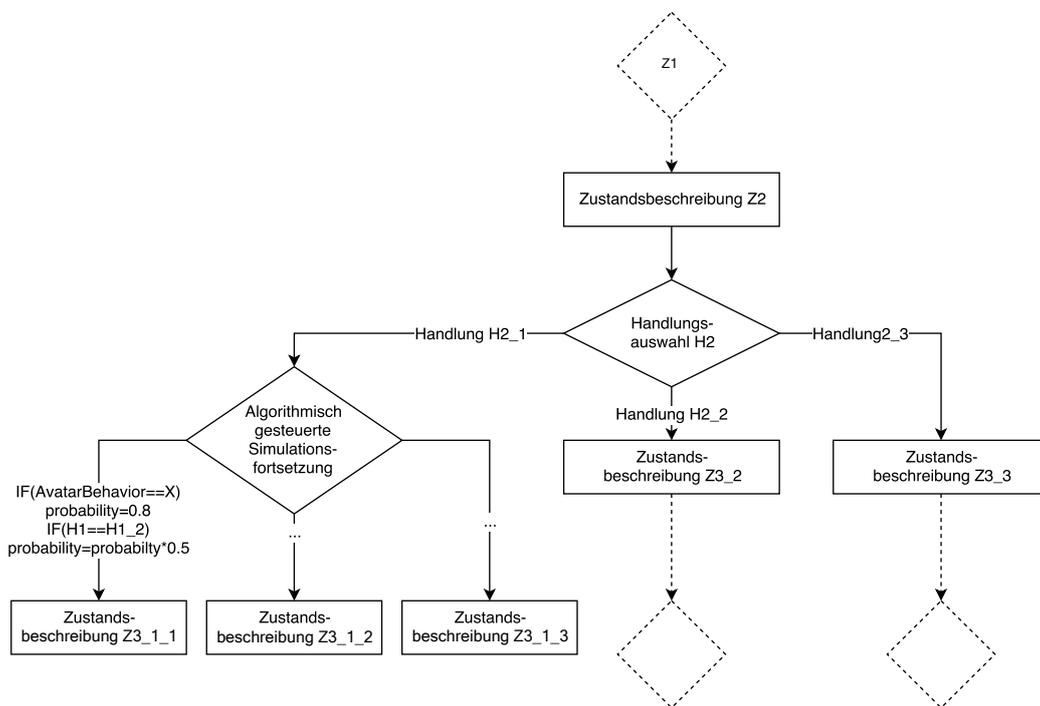


Abbildung 5: Beispielhafte Veranschaulichung eines komplexen Entscheidungsbaums mit algorithmischer Steuerung, Quelle: Author

Ein kleines Beispiel soll schildern, wie solche Simulationen funktionieren: Angenommen, eine Simulationsnutzerin benutzt einen Unterrichtssimulator, in dem sie mit dem Störverhalten einer Schülerin konfrontiert ist. Sie kann zwischen n Handlungsalternativen wählen, z.B. die Schülerin ermahnen oder die Störung ignorieren. Ermahnt sie die Schülerin, beträgt die Wahrscheinlichkeit, dass die Schülerin mit dem Störungsverhalten aufhört, 80%, dass sich das Verhalten fortsetzt 20%. Wählt sie aus die Schülerin zu ignorieren, so beträgt in diesem Szenario die Wahrscheinlichkeit, dass sie ihr Störverhalten beendet, 40%, dass sie es fortsetzt 60%. Dies ist der erste Typ von Interaktionsmodellierung, der die Nichtlinearität menschlichen Verhaltens (Skrødal, 2010, S.37; Alessi & Trollip, 2001, S.224) berücksichtigt, wenn auch nur über einfache Randomisierungen.

3.3.3 Typ 3 – Entscheidungsbaum mit komplexer algorithmischer Steuerung

Ein Entscheidungsbaum mit komplexer algorithmischer Steuerung ist die logische Fortsetzung der bereits vorgestellten Varianten von Interaktionsmodellierungen. Statt fester Wahrscheinlichkeiten, steuert in mindestens einem Ent-

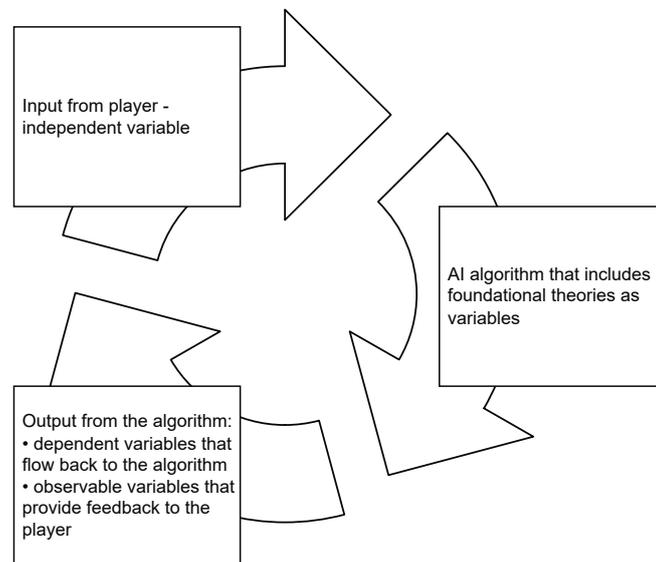


Abbildung 6: Interaktionen in einer Simulation mit komplexer algorithmischer Steuerung, Quelle: Deale und Pastore (2014, S.199)

scheidungsknoten ein Algorithmus, welches Ergebnis eine Simulationshandlung erzeugt. Hierbei kann der Algorithmus beispielsweise frühere, von der Simulationsnutzerin getroffene Entscheidungen oder Persönlichkeitsfaktoren der SuS berücksichtigen, um die Verzweigungswahrscheinlichkeit zu beeinflussen. Der Algorithmus selbst kann äußerst komplex oder relativ simpel sein, insofern existiert eine große Spannweite an möglichen Umsetzungen. Benutzt man eine solche Interaktionsmodellierung (oder Typ 4–5) wird es wichtig, Simulationsmodelle zu generieren, die sich an physische, emotionale, kognitive, soziale und organisationale Theorien bzw. Konzeptionsmodelle anlehnen (Christensen et al., 2011, S.203). Diese Simulationsmodelle müssen anschließend angewendet, getestet und nachjustiert werden (Christensen et al., 2011, S.203). Die generelle Funktionsweise einer solchen Simulation wird nach Deale und Pastore (2014) in Abb. 6 (S. 59) veranschaulicht.

Der verwendete Begriff „komplexe Modellierung“ wird auch bei Deale und Pastore (2014, S.210) verwendet und wurde, in seiner Bedeutung leicht angepasst, für diese Taxonomie adaptiert.

3.3.4 Typ 4 – Komplexe algorithmische Steuerung ohne Entscheidungsbaum

Bei dieser Art von Simulation existiert kein festes Szenario, abgesehen von der Situierung des Szenarios (bei *lesson simulations* beispielsweise ein Klassenraum). Sämtliche simulierten Handlungen und Handlungsergebnisse werden durch die Simulationsnutzerin oder durch die den Agenten steuernden Algorithmen hervorgerufen. Diese Interaktionsmodellierung ist die abstrakteste (Gibson, 2009b, S.142) und verlangt am deutlichsten nach einem komplexen Simulationsmodell. Der Abstraktionsgrad der Modellierung ermöglicht eine generalisierbare Anwendung und Weiterverwertung über einzelne spezifische Fächer oder *critical incidents* hinaus. Daraus folgt jedoch der Nachteil, dass eine Schulung spezifischer Szenarien schwierig bis unmöglich wird und die Simulation auf einem höheren Abstraktionsniveau liegt. Aus diesem Grund eignet sich dieses Modell zur Erstellung von Simulationen mit höherer Granularität. Simulationen, die diese Interaktionsmodellierung verwenden, besitzen große Ähnlichkeit mit Experimenten. Das didaktische Ziel ist, durch gezielte Veränderung von Input- auf Outputvariablen durch die Simulationsnutzerinnen das dahinterstehende Modell zu verstehen (Hopper, 2014, S. 27f.). Während Entscheidungsbäume so konzipiert sind, dass ein oder zwei Pfade dazu auserkoren werden, ‚richtiges‘ bzw. erwünschtes Verhalten zu markieren, ermöglichen Interaktionsmodellierungen des Typs 4 durch ihre experimentelle Anlage multivariante Lösungen (Hopper, 2018, S. 91; Gibson, 2012a, S. 41; Kruse & Gibson, 2011).

Als Beispiel könnte ein Simulator wie VCS genannt werden (Skrødal, 2010, S.5), in dem virtuellen SuS Fragen gestellt werden können. Im Simulationsmodell werden allen Fragen drei Variablen zugeordnet, nämlich Schwierigkeit, Neuheitsgrad und notwendige Zeit um die Frage zu beantworten. Die Variablen sind intervallskaliert, mit Ausprägungswerten von minimal 0 bis maximal 1. Die Lehrkraft-Frage wird im Simulator nie ausformuliert, sie wird nur durch diese drei numerischen Ausprägungen beschrieben. Änderungen an den Ausprägungen der Inputvariablen führen zu unterschiedlichen Ergebnissen. Im konkreten Falle werden bei jeder Lehrkraft-Handlung über einen Algorithmus fünf SuS-Outputvariablen verändert, nämlich Gemütszustand, Ablenkbarkeit, Selbstwirksamkeitserwartung, Wissen und Metakognition. Diese Art von Simulation ist aufgrund ihres Abstraktionsgrads wesentlich weniger immersiv, die Simulationsnutzerin wird eher in die Rolle einer Forscherin oder Experimentatorin hineinversetzt. Sie versucht eher die Funktionsweise des modellierten Systems zu verstehen, als dass sie sich situativ in die Rolle einer Lehrkraft vor einer Klasse hineinversetzt fühlt. Das didaktische Lernziel ist die Aneignung des Simulationsmodells.

Dies ist in einer physikalischen Simulation sehr sinnvoll, da die zugrunde liegenden Modelle meist mathematisch genau vorhersagbar sind. So ist die Simulation eines schiefen Wurfes auf gerader Fläche durch Abwurfhöhe, Geschwindigkeit und Winkel vollständig definiert (Inputvariablen) und die Outputvariablen von zurückgelegter Strecke, Maximalhöhe oder benötigter Zeit bis zum Aufprall sind unumstritten. Ein Interaktionsmodell zwischen zwei oder mehreren Menschen, wie es für die Simulation unterrichtlicher Lehrkraft-SuS-Interaktionen nötig ist, folgt einer anderen Modellierungslogik, die auch *qualitative simulation modelling* genannt wird (Skrødal, 2010, S.37). Der Grund dafür ist, dass menschliche Interaktionsprozesse inhärent mit Unsicherheit behaftet und damit schwieriger zu modellieren und zu validieren sind (Skrødal, 2010, S. 37, 40; Alessi & Trollip, 2001, S. 255). Hier muss ein Modell zwangsläufig eine radikale(re) Vergrößerung sein.

Die Interaktionsmodellierung vom Typ 4 ist die einzige Interaktionsmodellierung ohne menschliche Operatoren, die kontinuierliche Eventmodellierung möglich macht und nicht auf diskrete Eventmodellierung angewiesen ist (Skrødal, 2010, S. 44). Bei einer diskreten Eventmodellierung läuft die Simulation in abgrenzbaren Schritten ab, bei einer kontinuierlichen Eventmodellierung verändern sich die Outputvariablen der Simulation ständig (Skrødal, 2010, S. 78).

Diese Trennung ist nicht ganz trennscharf, denn aus technischen Gründen aktualisiert sich eine Computersimulation immer in bestimmten – wenn auch sehr geringen – diskreten Zeitintervallen. Darüber hinaus verstehen andere Autorinnen unter kontinuierlicher und diskreter Modellierung gänzlich andere Konzepte (Alessi & Trollip, 2001, S.237f.). Eine Definition für diese Arbeit ist, dass die Simulationsnutzerin bei kontinuierlicher Eventmodellierung den Eindruck hat, dass ihr Verhalten in Echtzeit von den Agenten verarbeitet wird, während bei diskreter Eventmodellierung der Eindruck überwiegt die Simulation in klar abgegrenzten ‚Simulations-Runden‘ zu absolvieren.

Da dieser Typ von Interaktionsmodellierung sehr stark von der wahrgenommenen Güte der verwendeten Algorithmen abhängt, soll an dieser Stelle kurz thematisiert werden, ob künstliche neurale Netzwerke (engl. *ANN – artificial neural networks*) oder künstliche Intelligenz (*AI – Artificial intelligence*) hierbei nützlich sein können, vor allem vor dem Hintergrund, dass dies häufig in der Literatur als eine Art Königsweg angepriesen wird (Brown, 1999, S.311; Chieu & Herbst, 2011, S.105; Gibson, 2009a, S.673; Gibson, 2012a, S.42).

Aktuelle Fortschritte in den Bereichen *machine learning* und *deep learning* tragen dazu bei, auf dem Feld der Simulationen unter Umständen unrealistische Erwartungen zu wecken. Es liegt außerhalb des Zieles dieser Arbeit die genauen Funktionsweisen dieser Algorithmen zu beschreiben. Es ist aber

wichtig künstliche Intelligenz von dem darin enthalten Teilbereich *machine learning* logisch zu trennen, da dies Auswirkungen auf die Diskussion um Simulationsmodelle hat. Außerdem soll kurz erklärt werden, welche Fortschritte von *machine learning* im Simulationskontext zu erwarten und welche Hoffnungen eher unbegründet sind.

KI ist nach der offiziellen Definition der Association for the Advancement of Artificial Intelligence (AAAI) sehr offen: „Artificial intelligence is the method of making computer programs intelligent or to act conditionally based on stimuli.“ In diesem Sinne wird künstliche Intelligenz auch in der Diskussion um Simulationen verwendet. Häufig verstecken sich hinter diesem Oberbegriff logikbasierte Modellierungsansätze: Vereinfacht gesprochen handelt es sich um komplexe IF-THEN-ELSE Statements. Während auch diese Arbeit innerhalb dieser logikbasierten Ansätze operiert, so ist festzuhalten, dass diese Algorithmusform seit Anbeginn der Digitaltechnik existiert und wenig mit der aktuellen Euphoriewelle um künstliche Intelligenz zu tun hat.

Machine learning hingegen funktioniert nach dem Prinzip, dass einem System verschiedene Inputvariablen gegeben werden, die anschließend eine bestimmte Outputvariable erzeugen sollen, wobei der dazu nötige Algorithmus eine Art Black Box darstellt. Dazu ist es nötig ein neuronales Netzwerk mit einer gewaltigen Masse an (oft) menschengenerierten Zuordnungen von Input- zu Outputvariablen zu trainieren. Um dies an einem klassischen Beispiel des *machine learnings*, der Bilderkennung, zu illustrieren: Zunächst wählt man die Inputvariablen aus, was in diesem Falle recht einfach ist, nämlich sämtliche Pixel mit deren RGB-Farbausprägungen, die ein Bild enthält. Nun erstellt man einen riesigen Datenkorpus an Bildern, der nach einer Outputvariable kodiert wird, z.B. ob auf dem Bild ein Hund oder eine Katze zu sehen ist. Es ist nötig den Umfang der Outputvariablen zu kontrollieren: Würde man alle möglichen Dinge beschreiben, die auf einem Bild zu sehen sind oder hypothetisch zu sehen sein könnten, erhält man eine annähernd unendliche Datenmenge. Legt man hingegen alle möglichen Tiere als Outputvariable fest, so muss der Datenkorpus gewaltig sein, da sonst das System nicht genügend Trainingsdaten für jedes einzelne Tier enthält. Anschließend stellt das künstliche neuronale Netzwerk einen Zusammenhang zwischen Input- und Outputvariablen her. Man kann einem so trainierten Netzwerk Bilder vorlegen, mit denen es nicht trainiert wurde und es wird identifizieren können, ob sich darauf z.B. ein Hund oder eine Katze befinden. Es ist für den Menschen nicht direkt aus dem Quellcode nachvollziehbar, welche Entscheidungskriterien dazu führen die Outputvariable zu bestimmen. Die Netzwerkknoten sind zwar in ihren Werten nachvollziehbar, aber Menschen können nicht genau sagen, was sie bedeuten. Aktuell findet verstärkt Forschung statt, die sich darum bemüht mittels Heatmap-Bildern oder kurzen textbasierten Statements

der KI zu begründen, warum die KI den Inputvariablen eine bestimmte Outputvariable zugeordnet hat (Samek et al., 2019).

Der Einsatz von *machine learning* in der Interaktionsmodellierung von Unterrichtssimulationen ist mit drei Problemen behaftet, einem administrativen und zwei logischen. Das administrative Problem hängt damit zusammen, dass *machine learning*-Ansätze riesige Mengen an Trainingsdaten benötigen, damit sie funktionieren. Geeignete Trainingsdaten existieren für den Schulunterricht nicht. Vorhandene Daten haben einen zu geringen Umfang um sie im Rahmen von *big data* Ansätzen zu verwenden. Außerdem wäre häufig eine Umkodierung der vorhandenen Rohdaten erforderlich. Zudem stellt sich die Frage, wie geeignete Trainingsdaten aussehen müssen, was direkt zum ersten logischen Problem führt. Anders als in dem geschilderten Beispiel zur Bilderkennung ist bei Interaktionsprozessen im Unterricht nicht klar, welche Inputvariablen auszuwählen sind. Von vornherein ist klar, dass nicht alle Inputvariablen der Realität berücksichtigt werden können. Es ist nicht möglich alle Inputvariablen (z.B. Motivation) kontinuierlich innerhalb einer Stunde zu messen. Außerdem würde eine Berücksichtigung aller Faktoren es erfordern die Gehirnstruktur eines Menschen zu modellieren, was die Erschaffung einer sogenannten starken KI bedingen würde. Letzteres ist bisher nicht möglich und es gibt viele Expertinnen aus dem Gebiet der Informatik, die dies generell für unmöglich halten. Demzufolge bleibt nur die theoretisch begründete Auswahl einiger Inputvariablen, die man über den Verlauf von tausenden von Unterrichtsstunden messen müsste. Die Outputvariablen sind durch Beobachtung/Rating zu erzeugen. Hier stellt sich eine ähnliche Frage wie bei dem Beispiel der Bilderkennung von Tieren. Sämtliche Interaktionen ergeben einen unendlich großen Datenkorpus, also müssen diese – ebenfalls theoretisch begründet – in verschiedene Typen geclustert werden. Gehen wir davon aus, dass diese Daten generiert wurden, kann ein Simulator konstruiert werden, der SuS simuliert, die bei ähnlichen Inputvariablen in den Simulationen ähnliches Verhalten zeigen wie in der Realität.

Das zweite logische Problem ist, dass der Algorithmus auch den Simulationskonstrukteurinnen unbekannt bliebe. Das macht es schwierig die Simulation so einzusetzen, dass damit bestimmte Fertigkeiten geschult werden, weil man eben nicht weiß, wie eine Lehrkraft handlung nun genau mit dem Verhalten einer Schülerin korreliert.

Es ist keineswegs ausgeschlossen, sondern sogar sehr wahrscheinlich dass *machine learning* Simulatoren ermöglicht, die als realitätsnäher als Simulatoren mit logikbasierten Ansätzen wahrgenommen werden. Dennoch existieren die beschriebenen methodologischen Probleme, welche auch in Zukunft nicht aufgelöst werden können, da sie logisch-grundsätzlicher Natur sind.

Aus diesen Gründen wird in dieser Arbeit auf einen *machine learning*-basier-

ten Ansatz der Interaktionsmodellierung verzichtet. Zuletzt erwähnenswert ist, dass bisher noch nicht einmal Ansätze, z.B. theoretisch-basierte Begründungen von relevanten Input- und Outputvariablen, vorliegen, die zur Erstellung einer *machine-learning* basierten Unterrichtssimulation dienen könnten.

3.3.5 Typ 5 – Komplexe algorithmische Steuerung mit eingeschränktem Entscheidungsbaum

Bei dieser Art der Simulation werden Entscheidungsbäume durchgespielt und um Handlungsstränge ergänzt, die durch Simulationsnutzerinnen oder Simulationsagenten hervorgerufen werden. Größtenteils ähnelt dieser Typ von Simulation damit dem Entscheidungsbaum mit komplexer algorithmischer Steuerung, kann jedoch im Gegensatz zu dieser Interaktionsmodellierung von einer rein deterministischen Abfolge abweichen. Beispielsweise könnte ein Simulator nach dieser Interaktionsmodellierungslogik folgendermaßen implementiert sein: Das geskriptete Szenario umreißt den Umgang mit Störverhalten einer Schülerin im Moment Z_0 . Es gibt mehrere Möglichkeiten mit der Störung umzugehen, die in verschiedene Zustandsbeschreibungen münden (Z_1) und schließlich mit einem Abschluss der Situation in Z_2 enden. Nehmen wir ferner an, die Simulationsnutzerin entscheidet sich die Störung zu ignorieren. Dies verändert verschiedene interne motivationale Faktoren im Schülerinnenpersönlichkeitsmodell einer anderen Schülerin. Diese beginnt daraufhin motorisch unruhig zu werden. In der Simulation wird nun ein Handlungsstrang hinzugefügt, der zwischen Z_1 und Z_2 liegt und sich damit beschäftigt auf die motorische Unruhe zu reagieren oder diese zu ignorieren. Je nach Komplexität des Simulators können diese Abweichungen von der Haupthandlungskette auch einen größeren Umfang als nur einen Handlungsstrang umfassen, allerdings kehrt der Simulator nach Abschluss dieser Nebenhandlungsstränge immer wieder zum Haupthandlungsstrang zurück. Die Idee hinter der Implementierung einer solchen Interaktionsmodellierung ist es, die Reduktion zu einem reinen Reiz-Reaktions-Schema, wie es in Typ 1–3 modelliert wird, zu vermeiden, aber im Gegensatz zu Typ 4 in der Lage zu sein spezifische Szenarien nach *critical incidents* zu simulieren.

3.3.6 Typ 6 – *human-in-the-loop*

Der Begriff *human-in-the-loop* (Chini et al., 2016, S.4; Nagendran et al., 2013, S.143; Alessi & Trollip, 2001, S.237, hier „*operator-in-the-loop*“; Dawson, 2016, S.35) beschreibt einen Simulator, der zwangsläufig die Beteiligung mindestens einer weiteren menschlichen Person, Operatorin genannt, erfordert. Eine Möglichkeit der Implementierung ist die Verwendung von Client-

Server-Architekturen, wie beispielsweise in der Simulation Breaking Bad Behaviors. In dieser Simulation steuert eine Operatorin einen Server, mit Hilfe dessen sie sämtliche Computeragenten steuern kann. Vorteil an dieser Lösung ist, dass hier detailliert auf die Handlungen der Simulationsnutzerin eingegangen werden kann. Während bei *virtual puppetry* mehrere Operatorinnen denkbar sind, ist in der *human-in-the-loop* Interaktionsmodellierung nur eine Person erforderlich. Von Nachteil ist, dass eine Operatorin nicht gleichzeitig alle SuS-Agenten steuern kann und somit ein zentraler Punkt unterrichtlichen Geschehens, die Simultaneität, verloren geht. Strenggenommen operiert Breaking Bad Behavior mit einer Hybrid-Modellierung aus Typ 4 und 6, da einige Verhaltensweisen der SuS-Agenten auch automatisiert ausgeführt werden können (siehe [Abschnitt 4.3](#), S. 74).

Eine andere mögliche Umsetzung ist die textbasierte Einsendung von Simulationshandlungen innerhalb eines beschriebenen Szenarios an eine Expertin. Diese wertet mithilfe ihres Expertinnenwissens die Simulationshandlungen aus und beschreibt einen von ihr als plausibel empfundenen Fortgang der Simulation (Vgl. The Tigerlake Projekt, [Abschnitt 4.20](#), S. 130).

3.3.7 Typ 7 – *digital/virtual puppetry*

Digital/Virtual puppetry ist strenggenommen eine Sonderform des *human-in-the-loop*-Ansatzes. *Digital/Virtual puppetry* weist jedoch einige Eigenheiten auf und wird in der weitverbreiteten Simulation TeachLivE verwendet, die häufig im Forschungsdiskurs erscheint. Deswegen wird dieser Ansatz als gesonderter Punkt in die Taxonomie aufgenommen.

In einem *virtual puppetry* Simulator werden SuS-Agenten, also graphische Repräsentationen von SuS, durch mehrere menschliche Operatorinnen, auch Interaktantinnen genannt (Andreasen & Haciomeroglu, 2009; Dieker et al., 2008, S.11), gesteuert. Es findet keine algorithmisch gesteuerte Modellierung der Agenten statt, stattdessen sind die Simulationsmodelle implizit. Bei der Interaktionsmodellierung via *virtual puppetry* werden die Bewegungen der Schauspielerinnen auf ein 3D-Modell einer Schülerin via *motion capturing* (zu genaueren Erklärung dieses Prozesses siehe [Abschnitt 3.2.3](#), S. 48) übertragen. Um für einen *virtual puppetry* Simulator geeignet zu sein, muss das *motion capturing* in Echtzeit ablaufen. Die Interaktantinnen sind von den Simulationspartizipantinnen räumlich getrennt, zur Kommunikation werden Mikrofone und Lautsprecher/Kopfhörer verwendet.

Im Kontext der Simulorkonstruktion im Bildungsbereich wird oft der Begriff *virtual puppetry* verwendet (Bradley & Kendall, 2014, S.3), der gängige Begriff in der kommerziellen Nutzung in Filmen oder Videospiele lautet *digital puppetry*. Dieser Begriff wird auch bei Barmaki und Hughes (2015,

S.531) und Dieker et al. (2012, S.97) benutzt, um den Simulator TeachLivE zu beschreiben. Die Begriffe werden demzufolge synonym verwendet.

Diese Art der Interaktionsmodellierung ist eng verwandt mit dem *micro-teaching-peerteaching*, in denen auch SuS durch Lehrkräfte oder Lehrkräfte in Ausbildung verkörpert werden. Vereinzelt wird behauptet, dass *microteachings* weniger effektiv seien als *virtual puppetry* Simulationen; allerdings ist die Forschungslage hierzu nicht eindeutig (Dieker, Rodriguez et al., 2014, S.30; Straub et al., 2014, S.45). Die von Dieker, Rodriguez et al. (2014) angeführte Nicht-Wiederholbarkeit von *microteachings* ist in einer *virtual puppetry* Simulation ebenso gegeben. Es existieren vereinzelt empirische Studien aus anderen Forschungsbereichen, wie der klinischen Psychologie, die einen Vorteil von *VR-virtual puppetry* Anwendungen gegenüber klassischen Rollenspielen andeuten (Park et al., 2011).

Der große Vorteil dieser Interaktionsmodellierung ist, dass er den höchsten Freiheitsgrad aller Interaktionsmodellierungen aufweist, denn Menschen können flexibel auf neue Handlungssituationen reagieren und dennoch Szenarien oder Skripten folgen. Nachteil dieser Methode ist jedoch, dass dies die Skalierbarkeit der Simulationsanwendung stark einschränkt (Bradley & Kendall, 2014, S.7; Lugin et al., 2016, S.6), auch wenn vereinzelt anderes behauptet wird (Dieker et al., 2007, S.4). Die Übertragung der Animationen erfordert eine komplexe technologische Infrastruktur, außerdem müssen Personen, die als Schauspielerinnen fungieren sollen, geschult werden (Chini et al., 2016, S.3). Diese Schulung wiederum zerfällt in drei Einzelkompetenzen: Zum einen muss eine ausreichende Kenntnis der allgemeinen Pädagogik/pädagogischen Psychologie vorhanden sein, um die virtuellen SuS überzeugend darzustellen, zum anderen muss die interaktive Performance trainiert werden, die zwar schauspielerische Komponenten beinhaltet, jedoch auch auf die Spezifität der technologischen Rahmung eingehen muss, da z.B. Mimik nur eingeschränkt ausdrückbar ist (Dieker et al., 2007, S.6; Dieker et al., 2008, S.11). Zuletzt muss die Nutzung der technischen Komponenten an sich geschult werden (Chini et al., 2016, S.3). Dies bedeutet einen nicht unerheblichen Zeitaufwand, bei bestenfalls mittlerer wahrgenommener ökologischer Validität. In diesem Kontext ist auch eine Analyse eines Werbevideos für TeachLivE nicht uninteressant, in dem man eine Interaktantin sieht, die eine Schülerin schauspielert (Enterprise Florida, 2009). Es ist deutlich zu erkennen, dass die Interaktantin ihre Bewegungen in unnatürlicher Weise überdramatisiert. Es kann natürlich nur spekuliert werden, warum dieses Verhalten so gezeigt wird, aber es gibt zwei plausible Theorien. Die erste Theorie ist techniksoziologischer Natur und lautet, dass sich die Interaktantin so verhält, weil nur so ausreichende Thresholds übertragen werden können, damit die Figuren animiert wirken. Die zweite, psychologische Theo-

rie ist interessanter und auch wahrscheinlicher: Die Interaktantin verhält sich so, weil sie – eingezwängt in einen Motioncapture-Anzug – das Gefühl hat, sie müsse sich so verhalten. Da im Versuchsaufbau implizit angenommen wird, dass SuS-Körpersprache im Unterrichtskontext bedeutsam ist, werden große Anstrengungen unternommen, diese technisch zu übertragen. Ausgeblendet wird, dass die Körpersprache nur in bestimmten Interaktionen wichtig ist. So kann es zu Überkompensation kommen, die tendenziell unreal wirkt.

Ein großes Problem ist, dass der Verteilungsschlüssel von Interaktantinnen auf Partizipierende ungünstig ausfällt; es ist eine de facto eins zu eins Betreuung erforderlich, was einen erheblichen administrativen Aufwand bedeutet (Lugrin et al., 2016, S.6). Es ist prinzipiell denkbar, SuS von mehreren Interaktantinnen spielen zu lassen. Diese Art der Anwendung von TeachME/TeachLivE war jedoch in der Literatur nicht zu finden (Andreasen & Haciomeroglu, 2009; Chini et al., 2016; Dieker et al., 2008). Übernimmt eine Person die Kontrolle über alle SuS, hat das Auswirkungen auf die Fidelität der Simulation: Die erreichte Fidelität kann unter die eines regulären *microteaching-peerteachings* fallen. Der Grund dafür ist, dass gleichzeitige Handlungen der SuS so nicht ausgeführt werden können und auch die Stimme der Schauspielerin sich beim Wechsel zwischen den kontrollierten Agenten nicht ändert.

Eine Simulationseinheit ist nur dann wiederholbar, wenn alle beteiligten Interaktantinnen für eine weitere Simulationsrunde bereitstehen, was nur unter großem personell-administrativen Aufwand möglich ist. Es stellt sich die Frage, inwieweit solche Simulationen einen Vorteil gegenüber traditionellen *microteaching-peerteaching*-Settings darstellen. Dieker et al. (2008, S.14) postulieren diesen Vorteil, aber der relevante Unterschied zwischen *virtual puppetry* und *microteaching-peerteaching* liegt einzig in der Wahrnehmung der SuS durch die Simulationspartizipantinnen. Ob es wirklich einen Vorteil bietet, dreidimensionale Figuren von SuS statt der schauspielenden Simulationsteilnehmerinnen zu betrachten, ist keineswegs ausgemacht. Hierbei ist noch erschwerend zu berücksichtigen, dass in aktuellen Simulationsimplementierungen die Mimik der Interaktantinnen nur sehr rudimentär – falls überhaupt – übertragen wird. Feinheiten der Schauspielerinnenmimik gehen im Prozess der Übertragung auf die dreidimensionale Figur verloren (Chini et al., 2016, S.2), so sie denn überhaupt übertragen werden.

Ihre Stärke hat dieser Typ von Simulation insbesondere in der Sonderpädagogik, wenn es darum geht, gezielt spezielle Verhaltensweisen bei virtuellen SuS darzustellen und so z.B. den Umgang mit Autismus zu schulen (so wird auch TeachLivE bei Dieker et al., 2008 und Garland et al., 2015 verwendet). Die Wahl dieser Interaktionsmodellierung empfiehlt sich bei Adaption eines konstruktivistischen Lernparadigmas (Bailenson et al., 2008, S.112).

3.3.8 Typ 8 – Reale Handlungsdurchführung

Handelt es sich bei der Simulation um eine analoge, so können Simulationenutzerinnen und Interaktantinnen reale Handlungen durchführen. Diese Interaktionsmodellierung ist mit *virtual/digital puppetry* nicht deckungsgleich. Grund dafür ist nicht das verwendete Medium (digital versus analog), sondern die Steuerungsgranularität der Handlungen. Hierzu ein Beispiel: Während es in der Realität ohne weiteres möglich ist, als geschauspielerte Schülerin einen Papierflieger zu falten und zu werfen oder aber eine Karikatur der Lehrkraft zu zeichnen, so sind diese speziellen Handlungen im Bereich *virtual/digital puppetry* immer an die Programmierung gebunden. Sämtliche Handlungen, die simuliert werden sollen, müssen vorher explizit programmiert werden, sonst sind diese nicht durchzuführen. Theoretisch können sich Schauspielende auf einem ca. 2x3 Meter großen Raum frei bewegen, in der konkreten Implementierung von TeachLivE sind Gestensteuerungen allerdings nur in sitzender Pose möglich. Mimik kann oft nur holzschnittartig wiedergegeben werden und von der Programmierung nicht vorgesehene Handlungsweisen können nicht initiiert werden, was die Fidelität der Simulation einschränkt. Diese Art der Interaktionsmodellierung wird auch *authentic assessment* genannt (Fischler, 2006).

3.4 Andere Möglichkeiten der Taxonomie

Neben den bisher präsentierten Kategorien werden in der Literatur noch andere taxonomische Kategorien vorgeschlagen. Diese Kategorien wurden ausortiert, da sie sich nicht spezifisch mit Simulatoren in der Lehrkräftebildung beschäftigen, sondern global für alle Simulatoren entwickelt wurden. Da sich diese Arbeit jedoch ausschließlich mit Unterrichtssimulationen beschäftigt, interessieren z.B. physikalische Simulationen nur insofern, wie sie einen Kontrast zu Unterrichtssimulationen bilden.

Ein Beispiel dafür ist die Abgrenzung von „*about something simulations*“ (Unterkategorien *physical* und *iterative*) und „*how to do something simulations*“ (Unterkategorien *procedural* und *situational*) von Alessi und Trollip (2001, S.214).

Physical simulations sind Simulationen, die sich mit einem physikalischen Gegenstand oder Phänomen beschäftigen. Man lernt, indem man Objekte und Variablen manipuliert. Eine Simulation, die eine elektrische Schaltung simuliert, wäre ein Beispiel dafür (Alessi & Trollip, 2001, S.215).

Iterative simulations sind Prozess-Simulationen, bei denen man unveränderliche Eingangsvariablen festlegt und dann nicht mehr mit der Simulation interagiert; die Simulation berechnet ein festes Ergebnis (Alessi & Trollip,

2001, S.217f.). Ein Beispiel hierfür ist eine Wettersimulation. Wird mit einer solchen Simulation gelernt, so geht es darum sich durch mehrfache Anwendung die Simulationsmodelle zu erschließen.

procedural simulations beschäftigen sich damit optimale Handlungssequenzen zu vermitteln (Alessi & Trollip, 2001, S.221). Häufig werden solche Simulationen in Laborszenarien eingesetzt, um Partizipantinnen zu schulen, bevor sie sich in die reale Situation begeben. In seltenen Fällen sollen solche Simulationen aus ethischen Gründen die reale Situation ersetzen, wie im Falle des Sezieren/Obduzieren von Tieren/Menschen (Psotka, 1995, S.421f.). Ein weiteres Beispiel für *procedural simulations* sind Flugsimulationen.

Situational simulations beschäftigen sich mit Verhalten oder Einstellungen von Menschen oder Organisationen in verschiedenen Situationen. Alessi und Trollip (2001, S.224) sehen *situational simulations* als einen Spezialfall von prozeduralen Simulationen an, mit dem Unterschied, dass das Verhalten von Menschen wesentlich schwieriger zu modellieren ist, weswegen *situational simulations* einen gewissen Grad an Probabilistik zeigen müssen. Zudem ist die Simulationsteilnehmerin hier selbst Teil der Simulation und nicht davon abgekoppelt (Alessi & Trollip, 2001, S.255). Alessi und Trollip nennen als Beispiel für diese Art von Simulation den Klassenraumsimulator.

Es ist unschwer zu erkennen, dass Unterrichtssimulationen im engeren Sinne nur *situational simulations* sein können, mit Ausnahme eines Simulators, der über einen deterministischen Entscheidungsbaum (Interaktionsmodellierung vom Typ 1) strukturiert ist, da dieser nicht die Unsicherheit menschlichen Handelns in Betracht zieht. Aus diesem Grund ist diese Taxonomie für Unterrichtssimulationen ungeeignet.

Eine andere taxonomische Möglichkeit ist die technische Kategorie der Eingabemethode, mit einer Unterscheidung zwischen Motion Controllern, Maus/Tastatur, Gamepad/Joystick und Stimmerkennung (de Castell et al., 2014, S.335; Psotka, 1995, S.421). Zusätzlich unterscheiden einige Autorinnen auch noch Eingabeart, also z.B. Slider oder Multiple Choice (Aldrich, 2009, S.56f.). Eingabemethoden und -arten sind in einer Simulationstaxonomie immer dann sinnvoll, wenn die zu schulende Fertigkeit stark mit der motorischen Umsetzung zusammenhängt. Die Hoffnung auf haptische Controller, die über Force-Feedback hinausgehen (Psotka, 1995, S.421, 424) hat sich bislang für den Massenmarkt nicht erfüllt.

de Castell et al. (2014) sehen durch die Verwendung von realitätsnahen Eingabemethoden einen Paradigmenwechsel von der Simulation hin zur Imitation, was sie generell begrüßen.

Motorische Fähigkeiten können sich mit gewissem Aufwand also stark dem Paradigma der Imitation annähern. Da jedoch Unterrichtssimulationen für gewöhnlich keine Schulung von motorischen Fertigkeiten beinhalten, ist

die Kategorie der Eingabemethode für eine Taxonomie der Unterrichtssimulationen ungeeignet. Die Umsetzung erfolgt überwiegend über Tastatur- und Mauseingaben, eine Spracherkennung findet selten statt. Die Spezifika der Interaktion mit Menschen ist durch die Interaktionsmodellierungen bereits abgedeckt. Daher wird auf eine Aufnahme der Eingabemethoden in die Taxonomie der Unterrichtssimulationen verzichtet.

4 Simulatoren in der Lehrkräftebildung

4.1 Aha! Classroom Sim

Interaktionsmodellierung:	Typ 5 (Komplexe algorithmische Steuerung mit eingeschränktem Entscheidungsbaum)
Immersionsform:	2D (animated) graphical simulation
Granularität:	Teacher simulation
Noch in Benutzung:	Unklar, aber erwerbbar. Wird auf der Webseite des Herstellers nicht mehr angeboten (https://www.ahaprocess.com), die Webseite für das Produkt ist eingestellt und nicht mehr verfügbar (http://www.classroomsim.com/SIM6-8.html) bzw. nur über das Archivierungssystem Waybackmachine (https://web.archive.org/web/20130824032925/http://www.classroomsim.com/SIM6-8.html). Die Simulation ist aber nach wie vor als Beigabe der Monographie „Working with Students: Discipline Strategies for the Classroom“ von Ruby K. Payne erhältlich. Eine kostenfreie Demoversion ist zum Zeitpunkt des Verfassens dieser Arbeit noch auf https://classroom-sim-6-8.software.informer.com/download/ verfügbar.
Vorgänger:	–
Nachfolger:	–

Bei der Aha! (Process) Classroom Sim handelt es sich um eine kommerzielle Simulation bzw. ein *serious game* der von Ruby K. Payne gegründeten Bildungsgesellschaft Aha! Process, Inc. mit dem ausschließlichen Fokus auf *classroom management*. An der Entwicklung der Software waren diverse Programmiererinnen, Grafikerinnen und Simulations-Designerinnen unter der Projektadministration von Bethanie Hamlett Tucker, Ed.D. und Nancy Sturm, M.s.Ed beteiligt. Aha! (Process) Classroom Sim ist als Addendum zu der Monographie „Working with Students: Discipline Strategies for the Classroom“ von Payne (2006) konzipiert.

Es existieren insgesamt vier verschiedene Spiele als eigenständige Programme, je nach simulierter Jahrgangsstufe: 2, 3–5, 6–8, 9–12 (Chuang &

Shih, 2014). Die Applikation selbst ist in einem Comicstil gehalten, Schülerinnenbewegungen und verschiedene Interaktionen sind animiert. Der Klassenraum ist aus einer erhöhten Egoperspektive der Lehrkraft dargestellt, um Verdeckungen der Schülerinnenagenten zu vermeiden. Die Lehrkraft selbst wird nicht dargestellt.

Das *serious game* zerfällt in drei Teile: Eine Planungsphase vor und nach dem Unterricht, eine Unterrichtsphase und ein zusammenfassendes Feedback. Während der Planungsphase können Schülerinnenakten eingesehen werden und Klassenregeln konfiguriert werden. In der Unterrichtsphase wird der Lernprozess selbst nicht thematisiert, die Partizipantinnen bekommen nur Gelegenheit auf zwei Arten von randomisierten Ereignissen zu reagieren: *classroom-management*-bezogene (z.B. eine Schülerin hat ihre Hausaufgaben vergessen) und andere, die nicht lernbezogen sind (z.B. teilt eine Schülerin mit, dass sie den Unterricht gut findet). Reagieren Partizipantinnen via Mausclick auf eine nicht *classroom-management*-bezogene Aussage, interpretiert das die Applikation als Zeitverschwendung und zieht Zeit von der verbleibenden Stunde ab. Auf eventuell während dieser Zeit stattfindende Ereignisse können die Partizipantinnen nicht mehr reagieren. Wenn hingegen auf eine *classroom-management*-bezogene Aussage reagiert wird, erhält man 4-6 Handlungsvarianten, die man per single-choice Verfahren auswählen kann. Ähnlich einem 50/50-Joker, kann man mit dem Item „Kaffee“ Handlungsalternativen eliminieren, die das Spiel für schlecht geeignet hält. Nach der Stunde erhält man die Möglichkeit Kaffee und andere Items wie Bücher mit Punkten, die man für ‚gute‘ Handlungsalternativen zugeschrieben bekommt, einzukaufen. Wird eine Handlungsalternative ausgewählt, hat dies einen Einfluss auf die am unteren Bildschirmrand immer dargestellten Klassenfaktoren „*happiness*“, „*learning*“ und „*behavior*“. Zusätzlich erhält man bei jeder Handlungsalternative Feedback in Form eines Elternbriefs, der direkt nach der Auswahl vom virtuellen Schulleiter überreicht wird.

Tendenziell weist Aha! (Process) Classroom Sim eine größere Nähe zu *serious games* und *game based learning (GBL)* (Chuang & Shih, 2014, S.344) als zu einer Simulation und *simulation based learning* auf. In der Literatur wurde Aha! (Process) Classroom Sim nur am Rande wahrgenommen (Bradley & Kendall, 2014, S.8; Frick & Galindo, 2008, S.11; Gibson, 2009a, S.672; Chuang & Shih, 2014, S.344; McPherson et al., 2011, S.95), Studien scheinen mit der Applikation nicht durchgeführt worden zu sein.

4.2 Behavior Breakthroughs

Interaktionsmodellierung:	Typ 1 (Deterministischer Entscheidungsbaum) / Typ 3 (Komplexe algorithmische Steuerung mit Entscheidungsbaum)
Immersionsform:	3D graphical simulation
Granularität:	Task simulation
Noch in Benutzung:	Nein
Vorgänger:	–
Nachfolger:	–

Behavior Breakthroughs ist eine vom Southwest Research Institute, eine in San Antonio, Texas ansässige NGO, entwickelte iOS App, die den Umgang mit einem verhaltensauffälligen Kind (ADHS, Autismus) simuliert – allerdings nicht im Kontext von Schule und Unterricht. Die Applikation orientiert sich stark an einem behavioristischen Lernparadigma (Southwest Research Institute, 2011). Dies zeigt sich unter anderem daran, dass B. F. Skinner als gerahmtes Foto in der Simulation dargestellt wird. Zielgruppe sind Eltern und weitere Betreuerinnen des Kindes (Pritchard, 2013, S.85).

Mit diesem Spiel soll operante Konditionierung gelernt werden. Dafür gibt es sechs Levels, die jeweils Verstärkung und Extinktion, *Shaping*, Korrektur, *Chaining*, *Prompt Fading*, *Errorless Teaching* und *Discrete Trails* zum Thema haben (Pritchard, 2013, S.85; Lowdermilk et al., 2012, S.31). Laut eigener Aussage sollen mit der App ABA-Techniken (Applied Behavior Analysis) eingeübt werden (Lowdermilk et al., 2012, S.29).

Die Interaktionsmodellierung scheint, ähnlich wie die Kognito Interactive Simulations, zwar auf eine Variable abzu zielen, nämlich das „Agitation Meter“, allerdings scheint die Simulation in ihrer Anlage eher auf einen deterministischen Entscheidungsbaum hinzuweisen als auf eine komplexe algorithmische Modellierung. Die Handlungsmöglichkeiten sind stark reduziert: So kann man in einem Szenario, in dem das Kind Asa eine Lampe zerstört hat, etwa sagen „Stop That“, „What’s wrong with you?“, „Good quiet“ [sic!] und „You know better than that.“; außerdem kann man dem Kind eine Süßigkeit geben oder es ignorieren. Wird die ‚korrekte‘ Antwort gegeben erklingt ein Glockengeläut, gibt man die falsche Antwort, hört man einen Buzz-Sound (Pritchard, 2013, S.85).

Behavior Breakthroughs erschien 2011 und wurde bis 2013 betreut, ist aber heutzutage nicht mehr im Apple Store gelistet und demzufolge nicht

mehr erhältlich. Eine Desktop-Version wurde früher für \$50-99 USD verkauft (Pritchard, 2013, S.85).

Pritchard (2013) stellte der Software im Produktreview-Teil des Journals Behavior Analysis in Practice mit 1.5 von 5 Punkten ein vernichtendes Urteil aus und fürchtete, dass die Übersimplifizierung der App die angewandte Verhaltenstherapie in ein schlechtes Licht rückt. In der Studie von Lowdermilk et al. (2012) hingegen stellten die N=89 Studierenden, die die App nutzten, dieser ein eher positives Zeugnis aus, außerdem berichteten sie im Rahmen einer Selbstauskunft über Kompetenzzuwächse im Umgang mit autistischen Kindern (Lowdermilk et al., 2012, S.32f.).

4.3 Breaking Bad Behaviors

Interaktionsmodellierung:	Typ 6 (human-in-the-loop)
Immersionsform:	VR simulation
Granularität:	Task simulation, School lesson simulation
Noch in Benutzung:	Unklar, Entwicklung eingestellt.
Vorgänger:	–
Nachfolger:	–

Breaking Bad Behavior (auch B3 von Latoschik et al., 2016, S.317; oder BBB Lugin et al., 2016, S.2 abgekürzt) ist ein HMD-basierter VR-Simulator, der von der HCI Group der Universität Würzburg entwickelt wurde. Ideengeber des Projekts war Marc Erich Latoschik (Mathematik, Informatik und Human-Computer-Interaction), der zusammen mit dem aus dem gleichen Forschungsumfeld stammenden Jean-Luc Lugin den Simulator technisch entwickelte. Die pädagogischen Anteile gehen auf Christian Seufert (Studienrat) and Silke Grafe (Lehrstuhl für Schulpädagogik) zurück, welche ebenfalls an der Universität Würzburg arbeiteten. Das Pädagogikteam tauschte sich alle 2–4 Wochen mit dem technischen Entwicklungsteam aus (Lugin et al., 2016, S.3). Beim Design und der Evaluation waren zudem Michael Habel und Daniel Roth von der Universität zu Köln beteiligt. 13 Bachelor-Studierende arbeiteten an der technischen Realisation des Projektes mit (Lugin et al., 2016, S.3).

Breaking Bad Behaviors basiert auf einem Client-Server-Modell: Zum einen gibt es eine sogenannte Instruktorin, welche die Handlungen der virtuellen SuS über eine zweidimensionale Desktop-Anwendung steuert, zum anderen gibt es die Simulationspartizipantin, welche eine HMD-Brille des

Typs Orculus Rift DK2 verwendet, um mit den SuS in VR zu interagieren (Latoschik et al., 2016, S.317; Lugin et al., 2016, S.6). Der simulierte Klassenraum umfasst 24 SuS, wobei insgesamt 30 unterschiedliche 3D-Figuren von SuS mit einem mittleren Detailgrad (7.500-12.500 triangles³) erstellt wurden (Latoschik et al., 2016, S.317; Lugin et al., 2016, S.7f.). Es wird randomisiert gezogen, welche der Figuren in der konkreten Simulation welchen Platz besetzt. Um realistische Animation zu erstellen, wurde mit *motion capturing* gearbeitet (Lugin et al., 2016, S.8). Die simulierten SuS sind der Jahrgangsstufe 10–13 zuzuordnen (Lugin et al., 2016, S.12), der Grafikstil ist naturalistisch. Spätere Versionen von Breaking Bad Behaviors unterstützen bis zu 40 virtuelle SuS (Lugin, Charles et al., 2018, S.2236).

Technisch umgesetzt wurde die VR-Komponente durch die Nutzung der Unreal Engine 4. Zusätzlich diente eine Microsoft Kinect V2 dazu die Körperbewegungen der Simulationspartizipantinnen aufzuzeichnen und auf einen virtuellen Avatar zu übertragen (Latoschik et al., 2016, S.318; Lugin et al., 2016, S.7, 14). Dies sei nötig, weil *classroom management*, als das für die Simulation zentrale Konzept, viele non-verbale Komponenten wie Blickkontakt, Ausdruck von Affekten, Proximität, Handgesten und Kopfbewegungen beinhaltet, welche mit dieser Technik einer Analyse zugänglich seien (Lugin et al., 2016, S.3f.). Partizipantinnen können sich auf einer Fläche von 2.5 x 2.5m frei bewegen (Lugin et al., 2016, S.7), sie benutzen Kopfhörer, mit denen sie drei Arten von Klängen hören können: (1) Umgebungsklänge außerhalb der Schule wie Vogelzwitschern, (2) vorab aufgezeichnete Aussagen der virtuellen SuS und (3) Audio-Feedback der Instruktorinnen (Lugin et al., 2016, S.6, 9). Physiologische Daten der Partizipantinnen wurden über ein E4 Wristband der Firma Empathica aufgezeichnet. Die Kosten für sämtliche verwendeten Geräte beliefen sich auf 2500-4000 Euro (Latoschik et al., 2016, S.318; Lugin et al., 2016, S.6).

Die Instruktorin hat die Möglichkeit

- (a) den Unterrichtsraum aus drei Perspektiven zu beobachten: (a) Lehrkraft-Perspektive der Partizipantin, (b) vorne im Klassenraum, (c) hinten im Klassenraum (Lugin et al., 2016, S.7f.)
- (b) die SuS sechs verschiedene Animationen ausführen zu lassen
- (c) zwei Arten von verbalen SuS-Rückmeldungen zu initiieren
- (d) synchrones und asynchrones Feedback zu geben

³Unter triangles/Dreiecke versteht man die Dreiecke im virtuellen Raum, aus denen jedes Polygon besteht. Ein quadratisches Polygon besitzt immer 2 Dreiecke, während N-Gons beliebig viele Dreiecke enthalten können.

Die Instruktorinnen bei Breaking Bad Behaviors wurden nicht aus wissenschaftlichem Personal rekrutiert, sondern sind selbst Studierende, die ein zwölfwöchiges Videoanalyse-Seminar besucht haben und eine ca. 5 Minuten lange Schulung zur technischen Benutzung des Client-Systems erhielten (Lugrin et al., 2016, S.14). Instruktorinnen sind dazu angehalten ein Skript zu verwenden. Es existieren zwölf ca. 35-sekündige Phasen der Simulation, die exakt vorschreiben, was von der Instruktorin zu initiieren ist, beispielsweise wird die Phase 2 folgendermaßen beschrieben:

Create blue [Farbkodierte Störungskategorie, Anm. d. Verf.] disturbing behavior in the front of the classroom on a pupil outside teacher’s focus (20 Sekunden) Wait for the teacher to react (7 Sekunden) Don’t forget to give feedback (7 Sekunden) (Lugrin et al., 2016, S.13)

Die gesamte Simulation dauert sieben Minuten. (Lugrin, Oberdörfer et al., 2018, S.626) Die virtuellen SuS-Charaktere können per Drag und Drop dazu gebracht werden, eine der sechs Animationen abzuspielen. Dabei handelt es sich um die nicht-kursiv gedruckten Handlungen in der [Tabelle 1](#) (S. 76).

Tabelle 1: Störungsverhalten in Breaking Bad Behaviors, Quelle: Lugrin et al. (2016, S.4)

Bad behavior levels				
Null	Low	Mid	High	Extreme
Listening-and-quiet	Sleeping-on-table	Laughing	Making-funny-noises	Dancing
Raising-arm-for-question	Typing-on-mobile	Poking-neighbors	Receiving-phone-call	Screaming
<i>Head-Scratching</i>	<i>Whispering</i>	<i>Playing-with-pen</i>	<i>Throwing-object</i>	<i>Fighting</i>
<i>Leg-over</i>	<i>Looking-around</i>	<i>Talking-neighbors</i>	<i>Refusing-to-work</i>	<i>Making-phone-call</i>
<i>Writing</i>	<i>Leaning-back</i>	<i>Playing-with-pen [sic!]</i>	<i>Leaving-classroom</i>	<i>Breaking-object</i>

Bei den kursiv gedruckten Handlungen handelt es sich um konzipierte, aber noch nicht in den Simulator implementierte SuS-Handlungen. Wie aus der Tabelle ersichtlich, ist jede Handlung einem Level von Unterrichtsstörung zugeordnet, was theoretisch mit den Untersuchungen von Borich (2016) erklärt wird (Lugrin et al., 2016, S.4).

Neben der direkten Zuweisung von Störverhalten durch die Instruktorin gibt es zwei Algorithmen, die über das Verhalten der SuS bestimmen. Zunächst existiert eine ständig aktive Routine, die vorschreibt, dass SuS immer in die Richtung schauen aus der auditiv wahrnehmbares Störverhalten kommt. Eine weitere sogenannte ‚Idle-Routine‘ wird bei einer höheren Einstellung als ‚Null‘ auf dem Slider für den generellen Unruhegrad der Klasse über die GUI der Instruktorin aktiv (Lugrin et al., 2016, S.7). Ein einfacher

FSM-Algorithmus (*finite state machine*)⁴ wartet eine zufällige Zeit und wählt dann für jeweils alle SuS eine zufällige Animation aus der aktiven Liste aus (Lugrin et al., 2016, S.10). Die aktive Liste wird durch den Schwerheitsgrad der Unterrichtsstörung bestimmt, welcher auf dem Slider ausgewählt wurde.

Es gibt zwei Arten von Audio-Rückmeldungen von SuS, bezeichnet als ‚*simple*‘ und ‚*advanced*‘. Simple Audio-Rückmeldungen bestehen aus kurzen Satzfragmenten wie „Huh?“, „I don’t understand“, „Why?“ oder „Sorry“, während sich die erweiterten Audio-Rückmeldungen auf das konkrete Szenario beziehen und Aussagen wie „Is mobile phone usage authorized during the trip?“ beinhalten (Lugrin et al., 2016, S.7).

Beim synchronen Feedback wird jede Phase der Simulation mit einer Zustimmung oder einer Ablehnung des Verhaltens der Simulationsteilnehmerin bewertet. Dieses positive oder negative direkte Feedback ist mit Audio-Klängen gekoppelt. Das asynchrone Feedback enthält eine detaillierte fünfstufige Likert Skala von jeweils sieben Faktoren, nämlich (1) *time to react*, (2) *time to resolve*, (3) *coping strategy*, (4) *appropriate tone*, (5) *self-control*, (6) *classroom management* und (7) *teaching content* (Lugrin et al., 2016, S.7f.). Das asynchrone Feedback wird nach Ende der Simulation von der Instruktorin ausgefüllt.

Inhaltlich geht es in der Simulation um die Schulung von *classroom management* und speziell um den Umgang mit Störungsverhalten. Das Szenario ist nicht-fachbezogen. Die Teilnehmerinnen sollen die virtuellen SuS über einen anstehenden Klassenausflug informieren (Lugrin, Oberdörfer et al., 2018, S.626; HCI Group Würzburg, 2017). Es soll sowohl die Information der geplanten Aktivitäten als auch die Schulregeln – keine Zigaretten, Alkohol oder Drogen, sowie feste Bettzeiten – kommuniziert werden (Lugrin et al., 2016, S.12). Außerdem musste auf Rückfragen der SuS reagiert werden. Die Information der geplanten Aktivitäten wurde den Teilnehmerinnen vier Wochen im voraus zur Vorbereitung bereitgestellt (Lugrin et al., 2016, S.12).

Während sich die Teilnehmerinnen in der Simulation befinden, reden sie in natürlicher Weise mit den virtuellen SuS. Ein wichtiges Kriterium für die Forscherinnen ist die Erzeugung von Stress bei den Teilnehmerinnen, weil sie nach eigener Aussage nur so einen geeigneten Stimulus für Ihre Forschungsfrage erzeugen können (Latoschik et al., 2016, S.317; Lugrin et al., 2016, S.2ff.).

Zielgruppe sind zum einen L1–L3 Lehramtsstudierende, aber auch Studierende der Pädagogik oder der Educational Science (Lugrin, Oberdörfer

⁴Eine *finite state machine (FSM)* (auch endlicher Automat) ist ein mathematisches Modell, das zur Beschreibung des Verhaltens eines Systems verwendet wird. Es besteht aus einer begrenzten Anzahl von Zuständen, Übergängen zwischen den Zuständen und Aktionen, die bei bestimmten Übergängen ausgeführt werden.

et al., 2018, S.625; Lugin et al., 2016, S.15f.). Es war geplant in Zukunft auch praktizierende Lehrkräfte als Zielgruppe zu erschließen (Lugin et al., 2016, S.18f.).

Mit dem Simulator wurden bislang drei Studien durchgeführt: Eine mit N=53 LA-Studierenden, die einen Vorteil von VR-Simulationen gegenüber einem Video-Treatment erkennen lässt (Lugin, Oberdörfer et al., 2018), eine mit N=22, die keinen Einfluss von synchronem Feedback auf Performanz, *cognitive load* und Grad der intuitiven Nutzung erkennen konnte (Lugin et al., 2016), sowie eine mit N=148 Partizipantinnen, welche Untersuchungen zur Teilautomatisierung der virtuellen SuS durchführte. Zeigte eine Schülerin in dieser letzten Studie ein Störungsverhalten, so gab es nun drei verschiedene Algorithmen, die das Verhalten der anderen SuS bestimmten. In jedem Fall reagierte ein Teil der anderen SuS nun ihrerseits mit Störungsverhalten. Im ‚heuristischen‘ Ansatz war dieses Störungsverhalten abhängig von der Distanz zur Lehrperson, im ‚regionsbasierten‘ Ansatz abhängig von dem Abstand zur Störungsquelle. Der ‚stochastische‘ Ansatz, der randomisiert zuwies, welche SuS Fehlverhalten zeigten, wurde von den meisten Partizipantinnen als am realistischsten empfunden (Lugin, Charles et al., 2018). Diese Studie untersuchte ähnlich wie die hier vorliegende Studie die Auswirkungen und den Fidelitätsgrad des Simulationsmodells.

Breaking Bad Behaviors wurde trotz des relativ neuen Erscheinungsdatums bereits national und international rezipiert (Wiepke & Paulicke, 2018; Delamarre, Lunn et al., 2019; Delamarre, Buche & Lisetti, 2019; Stavroulia et al., 2019). Die Autorinnen planen in Zukunft auch voll-autonome SuS-Agenten zu entwickeln (Latoschik et al., 2016, S.318).

4.4 ClassSim

Interaktionsmodellierung:	Typ 1 (Deterministischer Entscheidungsbaum)
Immersionsform:	(Hyper-)Text-based simulation
Granularität:	Lesson simulation
Noch in Benutzung:	Nein
Vorgänger:	–
Nachfolger:	–

Bei ClassSim handelt es sich um einen australischen Simulator einer Kindergartenklasse mit 26 SuS im Alter von 5 bis 6 Jahren, entwickelt von

der University of Wollongong (Ferry et al., 2004, S.294). Einen alternativen Namen verwenden Christensen et al. (2011, S.214); sie bezeichnen den Simulator als „Virtual Kindergarden Classroom“. Da dies jedoch nicht die Eigenbezeichnung der beteiligten Forscherinnen ist, wurde auf das Hinzufügen des Alternativnamens in diesem Fall verzichtet.

Innerhalb der Simulation finden randomisiert gezogene Events statt, die in zwei Kategorien aufgeteilt sind: (*classroom*) *management* Entscheidungen und Lernentscheidungen, auf die die Simulationspartizipantinnen via Freitexteingaben oder Multiple/Single-Choice Auswahl reagieren müssen. Freitexteingaben können über den sogenannten ‚Thinking Space‘, eine Art Forum, mit anderen Lernenden und mit dem Urteil von Expertinnen verglichen werden (Ferry et al., 2004, S.297f. Ferry & Kervin, 2007, S.197). Der Fokus lag mehr darauf, die Simulationspartizipantinnen zur Reflexion über ihr Handeln einzuladen und weniger auf einer immersiven Simulation von Unterricht.

Um das Simulationsmodell von ClassSim zu entwickeln, wurde die Software Stella verwendet, welche die Form eines interaktiven Flussdiagramms besitzt. Für einen visuellen Eindruck dieses Modelles, welches den graphischen Darstellungen des Simulationsmodells von Clasivir (siehe [Abschnitt 8.3](#), S. 261) ähnelt, sei auf Kervin et al. (2004, S.6) verwiesen. Als Grundlage zur Modellierung wurden eigene Lehrerfahrungen der Forscherinnen herangezogen, sowie auch nicht spezifizierte weitere Datenquellen: „We drew upon our own classroom teaching experiences and the data we had collected from the many teachers we had observed and documented“ (Ferry & Kervin, 2007, S.193).

Der Simulator wurde entwickelt, um Studierende zu motivieren den Lehrkraftberuf zu ergreifen und die hohen Studienabbruchraten von Lehrkräften in Australien abzufedern (Ferry et al., 2005, S.84). Auch die Verbindung von Theorie und Praxis war eines der Ziele, allerdings ohne einen Fokus auf die Schulung handlungsnaher Kompetenzen (Ferry et al., 2004, S.295; Carrington et al., 2006, S.1). Zielgruppe waren Lehramtsstudierende der unteren Semester (Carrington et al., 2006, S.6). Inhaltlich lag der Fokus auf der Vermittlung von *classroom management*-Kompetenzen (Ferry et al., 2004, S.294; Ferry & Kervin, 2007, S.194). Die Studienpartizipantinnen sollten den virtuellen SuS Lese- und Schreibfähigkeiten vermitteln (Ferry et al., 2005, S.85; Carrington et al., 2006, S.3). Die Autorinnen betrachteten ihre Software als einen „walk-through of a classroom with some interactivity rather than an actual simulation“ (Ferry & Kervin, 2007, S.203f.).

Mit einer Dauer von 2x90 Minuten pro Simulationssession (Ferry et al., 2005, S.89) – was später auf 2x60 Minuten reduziert wurde (Ferry & Kervin, 2007, S.203; Carrington et al., 2006, S.6) – zählt ClassSim zu den ungewöhnlich langen Simulationen.

Im Feedback von Studierenden, die den Simulator im Rahmen von Lehrveranstaltungen benutzen durften, wurde häufig zurückgemeldet, dass die Entscheidungsmöglichkeiten nicht ausreichend seien (Ferry & Kervin, 2007, S.201).

ClassSim wurde auch in einigen neueren Forschungsarbeiten rezipiert (Collum, 2015, S.39f. Kaufman & Ireland, 2016, S.264; Fluck & Fox, 2011, S.400; Manburg et al., 2017, S.129; Deale & Pastore, 2014, S.204).

4.5 Cook School District Simulation

Interaktionsmodellierung:	Typ 4 (Komplexe algorithmische Steuerung ohne Entscheidungsbaum)
Immersionsform:	(Hyper-)Text-based Simulation
Granularität:	School lesson simulation
Noch in Benutzung:	Nein
Vorgänger:	–
Nachfolger:	–

Die Cook School District Simulation (auch Cook Simulation oder Cook School District) ist ein webbasierter Simulator, der von Gerald Girod, Mark Girod und – in programmiertechnischer Leitung – Jeff Denton an der Western Oregon University entwickelt und um das Jahr 2006 fertiggestellt wurde (M. Girod & Girod, 2006, S.497; G. R. Girod et al., 2007, S.209). Während der Jahre 2002–2004 befand sich der Simulator in der Entwicklungsphase, gefördert im Rahmen der ‚Preparing Tomorrow’s Teachers to use Technology (PT3)‘-Förderung. An der Konstruktion des Simulators waren Programmierinnen, Bildungswissenschaftlerinnen und Lehrkräfte beteiligt (M. Girod, 2009, S.211).

Die Cook School District Simulation benutzt keine Bilder und erinnert in der Umsetzung an eine Tabellenkalkulation. Diese Ähnlichkeit ist nicht zufällig, denn Jeff Denton erwähnt, dass die webbasierte Umsetzung auf einem Prototypen basierte, der in Oracle PL/SQL programmiert worden ist. Technisch basiert der Simulator auf PHP, Javascript und MySQL und wurde von Jeff Denton, Will M. Drury, Michael J. Ellis, Bryan T. Hockett, Cadee L. Hale und Robert Minato programmiert. Die Konzeption des Algorithmus’ geht hingegen auf Gerald Girod zurück (Denton, 2013).

Im Cook School District Simulator wird Unterricht in dem Schuldistrikt von Cook Country, Georgia, USA simuliert. Loggt sich eine Simulationspartizipantin auf der Webseite ein, kann sie ökonomische, bildungsbezogene und

kommunale Informationen über den Distrikt abrufen, um einen Eindruck über die Rahmenbedingungen des Unterrichts zu bekommen (M. Girod, 2009, S.213). Nach der Eingabe, welches Fach und welche Jahrgangsstufe unterrichtet werden soll, zeigt die Simulation eine tabellarische Zusammenstellung aller simulierten SuS an (Abschnitt 8.3, S. 245). Ebenso wie simSchool ist die Simulation auf algorithmischer Ebene fachunspezifisch, d.h. es können zwar zehn Fächer gewählt werden, aber diese beeinflussen die Funktionsweise der Simulation nur insofern, dass sie andere Freitexteingaben durch die Simulationsnutzerin provozieren (Denton, 2013). Eine Klasse besteht aus 20 (M. Girod & Girod, 2006, S.488) bis 30 (Denton, 2013) SuS, es können nur Klassen der Jahrgangsstufe 3 oder 9 unterrichtet werden (M. Girod, 2009, S.213; Denton, 2013). Es existiert eine Datenbank mit 200 vorgefertigten SuS, aus denen 20–30 SuS randomisiert gezogen werden (M. Girod, 2009, S.213; M. Girod & Girod, 2006, S.485). Mit einem Klick auf die Namen der SuS werden genauere Informationen über diese angezeigt. Allen SuS wurden viele Variablen zugewiesen (für eine genaue Aufschlüsselung siehe Abschnitt 4.5, S. 84) und es wurde definiert, wie sich die SuS zu konkreten Aufgaben verhalten sollen. Bei dieser Festlegung halfen Lehrkräfte durch ihre Erfahrungen mit realen SuS, auf denen die virtuellen SuS basieren (M. Girod & Girod, 2006, S.485). M. Girod (2009, S.211) erwähnen, dass die Datenbank dieser sogenannten privaten Variablen insgesamt 70,000 Einträge hat. Die Schülerinnenvariablen beinhalten unter anderem:

- (a) Adresse, Muttersprache, Ethnizität, Gender
- (b) Informationen über Namen, Beruf und Beziehungsstatus der Eltern
- (c) *ESL (english as second language)* und *IEP (individualized education program)* Informationen
- (d) 13 standardisierte Testresultate und 11 Noten
- (e) Schulan- und abwesenheiten
- (f) Gesundheitsreporte
- (g) Typische Verhaltensweisen der Schülerin (15–20 Verhaltensweisen) (G. R. Girod et al., 2007, S.211f. Denton, 2013)

Diese Variablen interagieren mit den Usereingaben über einen komplexen Algorithmus (M. Girod, 2009, S.211), der von den Autoren nicht offengelegt wird.

Der Simulator ist in drei Teile aufgeteilt, in welchen Usereingaben möglich sind: (a) Pretest, (b) der Unterricht an sich und (c) Posttest (G. R.

Girod et al., 2007, S.211). Während des Pre- und Posttests stellt die Simulationsnutzerin einen virtuellen Test nach verschiedenen Kriterien zusammen. Anschließend simuliert der Algorithmus die Ergebnisse der virtuellen SuS. Weder während des Pre- und Posttests, noch in der Simulation des Unterrichts wird das Verstreichen von Zeit modelliert. Sind alle Eingaben gemacht, präsentiert der Algorithmus einfach das errechnete Ergebnis und konstituiert damit eine *iterative simulation* (siehe [Abschnitt 3.4](#), S. 68).

Bevor Einstellungen zum Pre- und Posttest gemacht werden können muss festgelegt werden, welche Instruktionsziele während der Simulation verfolgt werden sollen. Dazu sind User aufgefordert zunächst die Instruktionsziele via Freitexteingabe zu beschreiben und sie anschließend einer Domäne und einer Wissensdimension zuzuordnen. Als Domänen können kognitive, verhaltensbezogene oder physische Instruktionsziele ausgewählt werden. Als Domänendimension steht im kognitiven Bereich *remembering, understanding, applying, analyzing, evaluating* und *creating* zur Auswahl. Damit werden auf die Wissensdimensionen der Bloom'schen bzw. Anderson-Krathwohl'schen Aufgabentaxonomie rekurriert (Anderson et al., 2000), allerdings ohne dies explizit in der Literatur anzugeben. Zu physischen und verhaltensbezogenen Instruktionszielen wird leider nicht erwähnt, welche Domänendimensionen verwendet werden. Es können beliebig viele Instruktionsziele eingefügt werden.

Anschließend wird man aufgefordert die Instruktionsstrategien festzulegen. Diese Instruktionsstrategien sind im wesentlichen gleichbedeutend mit Unterrichtsmethoden. Zur Auswahl stehen:

- (a) *discussing*
- (b) *group assessment*
- (c) *group work assigned*
- (d) *independent assessment*
- (e) *independent information gathering*
- (f) *independent problem solving and exploration*
- (g) *lecturing/explaining*
- (h) *modelling or models provided*
- (i) *practice provided*

Für jede dieser Instruktionsstrategien kann man eine Bearbeitungszeit in 5-Minuten Schritten auswählen (bis zu 100 Minuten). Eine Stunde kann beliebig viele Instruktionsstrategien enthalten, außerdem ist es möglich mehrere simulierte Unterrichtsstunden zu erzeugen. Jeder Stunde können auch Hausaufgaben zugeordnet werden, deren Bearbeitungszeit über drei Zeitintervalle festlegbar ist.

Danach wird im Pretest bestimmt, auf welches Instruktionsziel sich der Pretest bezieht. Kriterien für den Test können per Freitexteingabe durch die Partizipantinnen ergänzt werden. Außerdem wird der Typ des Assessments festgelegt, wobei (a) mündlich, (b) schriftlich und (c) Performance/Konstruktion zur Auswahl stehen. Zuletzt wird ausgewählt, wie viele Punkte im Test maximal erreichbar sind. Es ist über den Button ‚*make adaptations*‘ möglich bestimmten SuS einen adaptierten Test auszuhändigen. Wurden alle notwendigen Einstellungen gemacht, errechnet der Algorithmus, wie viele Punkte die SuS im Pretest erreicht haben, was wiederum auch tabellarisch dargestellt wird. Auch in dieser Ansicht ist es möglich über den Namen der SuS detailliertere Informationen über sie zu erhalten. Es ist Aufgabe der Partizipantinnen darüber zu reflektieren, warum einige SuS bei dem spezifischen Test besser und andere schlechter abgeschnitten haben und ob man ggf. individuelle Adaptionen machen sollte. Zur Reflexion kann zum einen ein ständig vorhandenes Notizfeld auf der linken Bildschirmseite verwendet werden, zum anderen tauchen regelmäßig eingeblendete ‚*Cues & Prompts*‘ mit Fragen und Anregungen auf (M. Girod, 2009, S.215). Nachdem man die Cues beantwortet hat, kann man sich die Antworten der anderen Partizipantinnen anzeigen lassen.

Abschließend wird der Unterricht an sich modelliert, wobei der Algorithmus nur eine boolesche Outputvariable für alle SuS berechnet (M. Girod, 2009, S.214; M. Girod & Girod, 2006, S.485). Bei dieser Variable handelt es sich um die On- und Off-Task Variable. Zudem existiert noch eine Anwesenheitsvariable, welche aber rein randomisiert gesteuert ist. Jedes Instruktionsziel wird separat modelliert und angezeigt.

Die Generierung und Darstellung des Posttests entspricht vollständig der des Prätests (Denton, 2013).

Theoretisch basiert der Simulator auf der, ebenfalls an der Western Oregon University entwickelten, Teacher Work Sample Methodology (TWSM), die das Ziehen von Verbindungen zwischen Lehrverhalten und Lernverhalten der SuS didaktisch thematisiert. Für eine genauere Beschreibung der TWSM sei an G. R. Girod et al. (2007, S.208f.) oder M. Girod und Girod (2006, S.483) verwiesen. Es wurden zwei Studien mit der Cook School District Simulation durchgeführt. Eine davon (N=3 x 24–28 Studierende) beschäftigte sich damit mittels Selbstauskunftsbögen zu eruieren, ob das Benutzen der

Simulation die Fähigkeit steigert Verbindungen zwischen Lehr- und Lernverhalten zu ziehen. Außerdem wurde geprüft, ob die Simulation als nützlich wahrgenommen wurde (M. Girod & Girod, 2006). Die andere mit N=33 Studierenden, LiVs und Lehrkräften mit drei Jahren Berufserfahrung und einer Vergleichsgruppe mit N=38 Personen untersuchte, ob Unterschiede zwischen Lehrkräften unterschiedlicher Erfahrungsgrade durch die Simulationsnutzung festzustellen sind und falls ja, wie sich diese manifestieren (M. Girod, 2009, S.218). Beide Studien weisen tendenziell lange Zeiten für die Simulationsnutzung von 60–120 Minuten auf (M. Girod & Girod, 2006, S.488; M. Girod, 2009, S.218).

Während des Verfassens dieser Arbeit ging die Webseite zur Cook School District Simulation vom Netz. Davor konnte man nach einem simplen Registrierungsprozess die Simulation nach Belieben nutzen. Einige Informationen zur genauen Funktionsweise speisen sich aus der Benutzung der Simulation.

Die Cook School District Simulation wurde zwar häufig in der Literatur erwähnt (Turan, 2015; Kaufman & Ireland, 2016, S.263; Manburg et al., 2017, S.129; Deale & Pastore, 2014, S.204; Frick & Galindo, 2008, S.11; Gibson, 2009a, S.672) jedoch scheint der Simulator weder weiterentwickelt, noch über die zwei erwähnten Studien hinaus wissenschaftlich verwendet worden zu sein.

Tabelle 2: Verwendete Variablen in der Cook School District Simulation

Independent Variables	Specific Elements
Test Sequence	Pretest, formative, post-test
Item type	Orgal, written, performance, or constructed
Curriculum area	Math, science, health, wood shop, and so forth
Instructional strategy	Independent assessment, group assessment, group work assigned, independent information gathering, independent problem solving and explanation, discussing, practice provided, lecturing/explaining, modeling or models provided, homework assigned
Domain and level	Attitude - topical and self-confidence; Cognitive - remember, understand, apply, analyze, evaluate, and create; Physical - simple and complex

4.6 Der Simulierte Klassenraum (SKR) / Simulated Classroom

Interaktionsmodellierung:	Typ 2 (Randomisierter Entscheidungsbaum)
Immersionsform:	(Hyper-)Text-based simulation
Granularität:	Task simulation
Noch in Benutzung:	Nein, nur in Form von VSC und V.C.R.
Vorgänger:	–
Nachfolger:	Virtual School Class (VSC) / Virtual Simulated Classroom (V.C.R.)

Der Simulierte Klassenraum (SKR) (Südkamp et al., 2008) (auch Simulated Classroom in Fiedler et al. 2007, S.666 genannt) ist eine der ersten, wenn nicht die erste digitale Unterrichtssimulation im deutschsprachigen Bereich. Die Bezeichnung Virtual School Class (VSC) wird von Fiedler (2012, S.5) zwar für die älteren Versionen verwendet, bezeichnet aber zusätzlich eine Weiterentwicklung des Programms. Es wurde daher aus Übersichtlichkeitsgründen beschlossen die Begriffe Virtual School Class (VSC) und Virtual Simulated Classroom (V.C.R.) nur für die Weiterentwicklungen des Simulierten Klassenraums (SKR) zu verwenden und getrennt zu behandeln. Entwickelt wurde das Programm in einer Arbeitsgruppe unter Klaus Fiedler, Anna Südkamp und Jens Möller (Kaiser & Südkamp, 2017, S.260).

In der ersten Version des simulierten Klassenraums wurden sechzehn SuS der sechsten Jahrgangsstufe modelliert, acht davon männlich, acht weiblich (Fiedler et al., 2002, S.535); in einer späteren Version waren zehn SuS (Südkamp & Möller, 2009, S.165) der Jahrgangsstufe 9 abgebildet. Die Namen der SuS gehen auf die beliebtesten Vornamen des Jahres 1992 zurück (Südkamp & Möller, 2009, S.165).⁵

Das mögliche Verhalten der SuS besteht aus zwei Handlungen, nämlich der Meldung und der Antwort, die richtig oder falsch sein kann. Bestimmt werden diese Handlungen durch zwei Variablen: Die Motivationsvariable (von 0 bis 100) legt die Wahrscheinlichkeit fest, zu der sich eine Schülerin meldet, die Fähigkeitsvariable legt die Wahrscheinlichkeit für eine korrekte Antwort fest (Fiedler et al., 2007, S.666; Fiedler et al., 2002, S.529). Zum Teil wurden Antworten der Schülerinnen in Textform dargeboten (Südkamp et al., 2008,

⁵Das Jahr 1992 wurde gewählt, da die simulierten SuS mit dem Geburtsjahr 1992 zum Zeitpunkt der Datenerhebung der Studie (2007) die 9. Jahrgangsstufe besuchen würden (Südkamp & Möller, 2009, S.165).

S.264; Südkamp & Möller, 2009, S.166), wobei mehrere Variationen von falschen Antworten existierten, um eine Dopplung bei gleicher Fragestellung zu vermeiden (Südkamp & Möller, 2009, S.166). In anderen Studien, die die Simulation verwendeten, wurde nur angegeben, ob die Antwort korrekt war oder nicht (Fiedler et al., 2002, S.538). Es ist nicht möglich eine Schülerin um eine Antwort zu bitten, die sich nicht gemeldet hat (Fiedler, 2012, S.6).

Die Verteilung der Variablenausprägung unter den SuS wurde innerhalb der Studien zum Teil manipuliert. Thematisch wurde in einer Studie mit acht verschiedenen Unterrichtsfächern (Fiedler et al., 2002, S.529) gearbeitet, in einer anderen wurde eine Physikstunde modelliert (Fiedler et al., 2007, S.667), wobei die Simulationsnutzerinnen in beiden Fällen nach Abschluss der Simulation gebeten wurden, zu schätzen, wie oft sich eine Schülerin prozentual zu den Fragen gemeldet und wie oft sie die richtige Antwort gegeben hat (Südkamp & Möller, 2009, S.163). Der modellierte Handlungsspielraum der simulierten SuS ändert sich nur in einem Versuchsaufbau zur Simulation einer Politikunterrichtsstunde. Hier konnten simulierte SuS einer politischen Aussage (a) zustimmen, (b) sie ablehnen oder (c) unentschlossen sein (Fiedler et al., 2007, S.670). Eine Simulationssession dauerte ca. 17–20 Minuten (Kaiser & Südkamp, 2017, S.260; Fiedler et al., 2002, S.538; Südkamp & Möller, 2009, S.166). Die Fragen und Antworten, die implementiert worden sind, stammen zum einen aus Textbüchern für die sechste Jahrgangsstufe, zum anderen wurden sie von den Autorinnen selbst erstellt (Fiedler et al., 2002, S.537).

Es ist durch die Anlage der Studien klar erkennbar, dass der Simulierte Klassenraum dazu eingesetzt wird, diagnostische Kompetenz bei Lehrkräften zu messen. Deswegen sind die Handlungsmöglichkeiten so stark eingeschränkt (Südkamp et al., 2008, S.263). Anders formuliert: Es ist nicht Ziel der Simulation der Simulationspartizipantin das Gefühl zu vermitteln, sie unterrichte tatsächlich eine reale Klasse. Es geht vielmehr darum zu eruieren, ob Lehrkräfte Meldeverhalten und korrekte Antworten richtig einschätzen können oder ob systematische Verzerrungen auftreten, eventuell sogar ein Bias gegen bestimmte Namen oder das Geschlecht, auch abhängig vom Fach (Fiedler et al., 2002, S.528, 536). Die Simulationsfidelität ist aus diesen Erwägungen heraus eher niedrig gehalten und die Interaktion findet entweder vollständig auf Textebene oder durch zugeordnete Fotos statt (Fiedler et al., 2002, S.537; Südkamp & Möller, 2009, S.165).

Die Versionen des Simulierten Klassenraums ab 2008 wurden in JAVA programmiert und werden in Form des Virtual Classroom (V.C.R.) (Fischer et al., 2017) noch heute verwendet. International wurden weder der Simulierte Klassenraum (SKR), noch die Nachfolgesimulationen rezipiert.

4.7 Diversity Simulation

Interaktionsmodellierung:	Typ 1 (Deterministischer Entscheidungsbaum) / Typ 6 (Human-in-the-loop)
Immersionsform:	(Hyper-)Text-based Simulation
Granularität:	Task simulation
Noch in Benutzung:	Unklar
Vorgänger:	ClassSim (Inspiration)
Nachfolger:	–

Bei der Diversity Simulation von Manburg et al. (2017) handelt es sich um einen hypertextbasierten Simulator, der in Zusammenarbeit mit der privaten Nova Southeastern University und den Firmen Tata Interactive Systems und TOPSIM entwickelt wurde (Manburg et al., 2017, S.130). Das Budget für die Entwicklung der Simulation betrug 100.000 \$, wobei Tata Interactive Systems für die Programmierung der Weboberfläche zuständig war, TOPSIM für das „overall custom simulation design“ und zehn Vollzeit Mitarbeitende der Universität für die Erstellung des Contents bzw. der Szenarien (Manburg et al., 2017, S.130).

Es wurden insgesamt sechs Simulationsszenarien erstellt, die alle dem Themenfeld Diversity zuzuordnen sind. Die Szenarien basieren auf realen Ereignissen und *critical incidents*, über die in verschiedenen Medien berichtet wurde (Manburg et al., 2017, S.130). Unter anderem wird die Diskriminierung einer muslimischen Schülerin, einer Englisch-als-Zweitsprache-Schülerin und von LGBTQIA+ SuS im Schulumfeld thematisiert (Manburg et al., 2017, S.50). Die Szenarien finden in verschiedenen Schultypen wie Kindergartenklassen (vergleichbar mit Vorschulklassen), Middle Schools und High Schools statt.

Technisch handelt es sich bei dem Simulator um eine dynamische HTML-Seite, z.T. wurde auch ergänzend Flash benutzt. Es gibt insgesamt vier inhaltliche Unterseiten: (a) Profile der SuS (*,student profiles‘*), (b) Beschreibung des Szenarios (*,scenarios‘*) (c) Mögliche Interventionen (*,types of intervention‘*) (d) Ausgang der Interventionen (*,reports‘*) (Manburg et al., 2017, S.134).

Unter (1) finden die Simulationspartizipantinnen Fotos, Namen und eine umfangreiche Biographie der virtuellen SuS, getrennt nach Schule, Familie, Ethnizität, akademischen Leistungen, Persönlichkeit und Hobbies. Die SuS basieren aus forschungsethischen Erwägungen nicht auf realen SuS, sondern wurden für die Simulation entwickelt. Unter (2) wird das Szenario im Fließtext beschrieben. Auf der Interventionsseite (3) können nun bis zu zehn mög-

liche, vorgefertigte Interventionen ausgewählt werden, um auf das unter (2) genannte Szenario zu reagieren. Außerdem kann per Freitext eine weitere, individuelle Intervention vorgeschlagen werden. Die Entscheidung nach Auswahl der Intervention wird nicht individuell, sondern in der Gruppe gefällt, nach einer einwöchigen Online-Diskussion. Die Ergebnisse der Intervention (4) sind in den sogenannten ‚reports‘ verfügbar. Diese werden automatisch erstellt, falls eine der vorgefertigten Interventionen gewählt wird. Wird hingegen eine individuelle Intervention präferiert, entwirft die Seminarleiterin einen passenden Ausgang der Situation (Manburg et al., 2017, S.135ff.).

Ziel der Simulation war es Gruppen- und Selbstreflexion anzustoßen; forschungsparadigmatisch ordnen sich die Forscherinnen dem *problem-based learning (PBL)* zu (Manburg et al., 2017, S.130). Zielgruppe sind Lehramtsstudierende. In einer Studie, die den Simulator mittels Selbstauskunftsbögen und Essays der Teilnehmerinnen evaluierte, nahmen N=193 Studierende, verteilt über drei Jahre, teil (Manburg et al., 2017, S.130, 137).

4.8 First Class

Interaktionsmodellierung:	Typ 1 (Deterministischer Entscheidungsbaum)
Immersionsform:	3D/mixed reality graphical simulation
Granularität:	Task simulation
Noch in Benutzung:	Vermutlich nein
Vorgänger:	–
Nachfolger:	–

Bei First Class handelt es sich um einen *mixed reality* Simulator, der 2015–2016 an der Pennsylvania State University entwickelt wurde und dessen Entwicklung gemeinhin der Musikdidaktikerin Ann Clemens zugeschrieben wird, wobei eine Beteiligung der Educational Gaming Commons (EGC), einer Gruppe innerhalb der universitätsinternen Einheit Teaching and Learning with Technology, erwähnt wird. Ann Clemens fungierte als Ideengeberin und Repräsentantin, die eigentliche programmiertechnische, grafische und zum größten Teil auch die konzeptionelle Arbeit am Simulator führte jedoch Zac Zidik, Bachelor of Fine Arts, Design & Visual Communications durch, der seit 2008 als Multimediaspezialist an der Penn State U arbeitet. Die Entwicklungsschritte beschreibt Zac Zidik ausführlich auf seinem Entwicklungsblog, dem annähernd vollständig die Beschreibung dieses Simulators entnommen

ist (Zidik, 2015). Es existieren zwar noch einige journalistische Artikel zum Simulator (Werni, 2016; Oberdick, 2016; Ravipati, 2016), diese bieten jedoch kaum nennenswerte Informationen, die Zidik nicht auch liefert. Wissenschaftliche Studien, die unter Verwendung von First Class durchgeführt worden sind, existieren nicht, ebensowenig wie eine akademische Rezeption. In der Selbstbeschreibung handelt es sich bei First Class um eine *immersive experience platform*, Zidik nennt es einen *teaching simulator* (Zidik, 2015). Analysiert man die Funktionsweise, liegt die Klassifikation als *serious game* nahe. Zielgruppe waren Lehramtsstudierende, wobei es auch Bestrebungen gab, die Simulation über das Feld der Lehrkräftebildung hinaus zu entwickeln (Werni, 2016; Oberdick, 2016). Es gibt keine Hinweise darauf, dass diese Pläne weiterverfolgt wurden.

First Class wurde in der Game Engine Unity entwickelt. Die Simulationsteilnehmerin sieht eine 3D-Darstellung einer virtuellen Klasse auf einem großen Monitor, wobei die Sensoren der Microsoft Kinect 2.0 benutzt werden, um die Position der Simulationsteilnehmerin im Raum zu bestimmen. Anhand dieser Daten wird die 3D-Darstellung verändert: Geht die Teilnehmerin näher zum Monitor, zoomt die Darstellung der SuS heran, geht sie nach links scrollt auch der Raum leicht nach links. Es ist nicht ganz klar, warum die Entscheidung getroffen wurde, dieser sogenannten *mixed reality* Darstellung, die auch TeachLive verwendet, den Vorzug gegenüber einer Implementierung mittels VR-Brille zu geben. Während der Simulation verwenden die Teilnehmerinnen ein Tablet als zweiten Monitor.

First Class ist die einzige moderne Simulation in der Lehrkräftebildung, welche eine automatisierte Stimmerkennung verwendet, die allerdings extrem rudimentär ist. Zidik, der vor dem Projekt noch nicht mit Stimmsteuerungen gearbeitet hat, nutzte zur Implementierung ein Unity Plugin von Rumen Filkov, welches auf der Microsoft Speech Platform Runtime V11 basiert. Als Eingabegerät wurde das interne Mikrofon der Kinect verwendet. Es gibt einige wenige Schlagworte, wie die Namen der SuS und zehn verschiedene Handlungen, wie z.B. „raise your hand“ oder „answer question“. Es wurde programmiert, dass First Class Aussagen erkennt, unabhängig davon ob Adressat oder Handlung zuerst genannt werden. Es ist, in anderen Worten egal, ob man sagt „Darren, raise your hand“ oder „Raise your hand, Darren“. Trotzdem müssen Befehle klar durch Pausen voneinander getrennt werden und das Hinzufügen von Wörtern in Befehle führt zu Fehlern. „Please answer question five, Darren“ würde verstanden werden; „Could you please answer the question, Darren“ hingegen nicht. Es ist zumindest diskutabel, ob die Sprachsteuerung mit ihrem engen Befehlskorsett der Simulation eine wertvolle Komponente hinzufügt.

In First Class werden sechs virtuelle SuS dargestellt, welche von Zidik in

einem Comicstil modelliert wurden und ein Alter von ungefähr dreizehn Jahren haben sollen. Zum Thema, warum Zidik genau sechs SuS implementiert hat, berichtet er folgendes: „Why 6? I don't know. It seemed like a manageable amount of characters to create without spending too much time.“ Die Modelle der SuS basieren nur auf einem 3D-Grundmodell, angepasst durch leichte geometrische und texturale Änderungen. Die SuS werden in zwei Dreierreihen dargestellt, was nach Zidik damit begründet wird, dass so möglichst wenige Verdeckungen entstünden. Es gibt insgesamt acht separate Körper-Animationen: (1) Motorische Unruhe, Fingertippen auf dem Tisch, (2) Motorische Unruhe mit den Füßen, (3) Eine Position, bei der sich die Schülerin tief im Stuhl hängen lässt, (4) Abstützen des Kopfes mit der Hand auf dem Tisch, (5) Hand heben zur Meldung, (6) Nicken, (7) Kopfschütteln (8) und ruhig sitzen (keine Animation). Daneben wird die Blickrichtung der SuS dynamisch gesteuert. In regelmäßigen Zeitabständen wird eine Augenblinzel-Animation durchgeführt und bei freundlichen Signalworten eine Lächel-Animation abgespielt.

Inhaltlich beschäftigt sich die Simulation zunächst mit Überprüfung der Anwesenheit, danach mit einer Lektion zum Thema Bullying. Jedes der vier Szenarien hat eine Länge von zwei bis fünf Minuten.

Das Verhalten der virtuellen SuS ist in First Class algorithmisch gesteuert. In vielen journalistischen Artikeln wird von einer AI gesprochen (Werni, 2016; Oberdick, 2016; Ravipati, 2016), jedoch sind die Routinen rein deterministisch gestaltet (Zidik, 2015).

Allen SuS wird eine prozentuale Aufmerksamkeitsvariable zugewiesen. Besteht unverdeckter Blickkontakt zwischen SuS und Lehrkraft, was über die Kinect ermittelt wird, so kann dieser Wert nicht unter 25% fallen. Befindet sich die Lehrkraft in räumlicher Nähe zur jeweiligen Schülerin (ebenfalls basierend auf Kinect-Daten), so steigert dies die Aufmerksamkeitsvariable um 50%. Spricht die Lehrkraft eine Schülerin an, so wird die Aufmerksamkeitsvariable unabhängig von den anderen Bedingungen auf 90% gesetzt. Eine Aufmerksamkeit von 100% wird erreicht, wenn in der Ansprache der Schülerin die freundlichen Signalworte „Good Morning“, „Thank you“ oder „Please“ verwendet werden. Die Aufmerksamkeitsvariable wird bei Blickkontakt zu einer Schülerin über eine Statusanzeige in Form eines Bleistifts in der rechten oberen Ecke des Bildschirms visualisiert. Zunächst bleibt die erreichte Aufmerksamkeit für eine bestimmte Zeit auf dem Ausgangswert. Diese Zeit ist abhängig von einer Variablen für die Aufmerksamkeitsspanne, die allen SuS fest zugewiesen ist. Die hintere Reihe der SuS hat eine Aufmerksamkeitsspanne von 30 Sekunden, die vordere Reihe von 60 Sekunden. Nach dieser Zeit fällt der Wert der Aufmerksamkeit quadratisch ab. Wie stark der Wert abfällt wird über eine jedem SuS zugeordnete Verfallsvariable determiniert.

SuS in den hinteren Reihen wird ein stärkerer Aufmerksamkeitsverfall zugewiesen. Fällt die Aufmerksamkeitsvariable auf einen Wert von unter 50% werden zufällig die Animationen (1), (2), (3) oder (4) abgespielt, oder der Blickkontakt von der Lehrkraft abgewendet. Damit sollen die virtuellen SuS *off-task* Verhalten signalisieren. Übersteigt die Aufmerksamkeitsvariable 50%, kehren die SuS in einen *on-task* Zustand (Animation 8) zurück. Dieser Wechsel geschieht nicht direkt bei dem Überschreiten des Grenzwertes von 50%, sondern wird durch eine Signifikanz-Variable mediiert, deren Funktion nicht detailliert geschildert wird.

Am Ende der Simulation wird ein Quiz durchgeführt. Ob die SuS die Fragen korrekt beantworten, hängt vom Durchschnitt ihrer Aufmerksamkeitsvariablen über die gesamte Stunde ab. Es handelt sich dabei um statische, vorab definierte Fragen, die mit „Ja“ oder „Nein“ beantwortet werden können. Liegt der Durchschnitt der Aufmerksamkeitsvariablen über dem Schwellenwert der Frage, antwortet die jeweilige Schülerin durch Nicken oder Kopfschütteln.

Der erwähnte zweite Bildschirm zeigt die vier auswählbaren Szenarien und die Namen der SuS an, des Weiteren wird die Aufmerksamkeitsvariable für SuS grafisch, in bestimmten Szenarien auch in Prozentwerten dargestellt. Außerdem finden sich auf dem Tablet-Bildschirm Tipps und Hinweise zur Steuerung der Simulation. Die Fragen für das Quiz am Ende der Stunde sind dort ausformuliert angegeben. Sie müssen entweder exakt so vorgelesen werden oder ein Druck auf einen Button reicht, damit die Frage gestellt wird (Zidik, 2015).

Auf Nachfrage zum Projekt via E-Mail meldete sich leider niemand. Einige Projektwebseiten der Penn State U sind mittlerweile nur noch über die Waybackmaschine zu erreichen. Da nie wissenschaftliche Publikationen erschienen sind und auch Zac Zidik nach 2016 keine weiteren Informationen zu First Class liefert, darf mit einiger Sicherheit davon ausgegangen werden, dass der Simulator nicht mehr genutzt wird.

4.9 Inner-City Simulation Laboratory / Teaching Problems Laboratory

Interaktionsmodellierung:	Typ 6 (Human-in-the-loop)
Immersionsform:	(Hyper-)Text-based Simulation / Interactive Video-based Simulation (IVBT)
Granularität:	Task simulation / Teacher simulation
Noch in Benutzung:	Nein (Fischler, 2006, S.55)
Vorgänger:	Mr. Lands Sixth Grade (indirekt, als Inspiration) (Cruickshank, 1988, S.137)
Nachfolger:	The Tigerlake Project (Fischler, 2006, S.58)

Bei dem Inner-City Simulation Laboratory und dem Teaching Problems Laboratory handelt es sich um zwei verschiedene nicht-computergestützte Szenarien, die dieselbe technische Umsetzung (d.h. denselben Simulator) verwenden. Beide Simulationen wurden von Donald R. Cruickshank am College at Brockport der State University of New York (z.T. in Zusammenarbeit mit Frank W. Broadbent) entwickelt (Cruickshank, 1988, S.137; Cruickshank, 1971, S.5).

Das Teacher Problems Laboratory wurde 1967 fertiggestellt, Cruickshank (1971, S.5) nennt uneindeutig dieses Jahr als Erscheinungsjahr. An anderer Stelle gibt er als Datum für das Inner-City Simulation Laboratory das Jahr 1969 an, in welchem auch eine entsprechende Monographie von ihm erschien (Cruickshank, 1988, S.141). Burns (1973, S.21) verweist in seiner Dissertationsschrift auf einen früheren Artikel von Cruickshank und Leonard (1967), allerdings beschreiben die Autoren hier nur die Entwicklung des Instruments zur Generierung der *critical incidents*. Die Nicht-Existenz einer fertigen Simulation wird aus folgendem Satz klar: „In addition, simulated experiences, which admit of more careful design and control than direct experiences, need to be developed“ (Cruickshank & Leonard, 1967, S.3).

In Cruickshanks Inner-City Simulation Laboratory wurde Simulationspartizipantinnen durch Mediennutzung eine kritische Klassenraumsituation vorgestellt. Sie übernehmen die Rolle von Pat Taylor, der Lehrkraft einer fünften bis sechsten Klasse einer innerstädtischen (Problem-)Schule (Fischler, 2006, S.58; Tansey, 1970, S.287; Cruickshank, 1988, S.137). Im Falle des Teaching Problems Laboratory ist die fiktive Schule hingegen ländlich geprägt. Zielgruppe für beide Simulationen waren sowohl Lehramtsstudierende, als auch neu eingestellte Lehrkräfte (Cruickshank, 1971, S.13). Es war beabsichtigt

einen Simulator zu entwerfen, der ohne große technische Anforderungen eine große Verbreitung finden kann (Cruickshank, 1988, S.138).

Die Identifizierung der *critical incidents* fand durch eine Erhebung unter Beteiligung von 163 neuen Lehrkräften für das Teaching Problems Laboratory und von 287 Lehrkräften für das Inner-City Simulation Laboratory statt (Burns, 1973, S.21; Cruickshank & Leonard, 1967, S.1; Cruickshank, 1971, S.11). Für das Inner-City Simulation Laboratory wurden 1400 *critical incidents* zusammengetragen und zu 184 unterscheidbaren Situationen destilliert (Cruickshank & Leonard, 1967, S.1).

Die Situationen wurden visuell mittels Kinescope Aufzeichnungen projiziert; eine simultan ablaufende Langspielplatte diente der Bereitstellung von Klang. Später wurden die Videos auf 16mm Film mit Sound aktualisiert. Wie bei Kersh's Mr. Land's Sixth Grade imitiert die Kameraperspektive den Blick der fiktiven Lehrkraft (Cruickshank, 1988, S.137; Tansey, 1970, S.287). Die Klasse besteht aus 32, primär afroamerikanischen, SuS. Andere Situationen wurden durch Rollenspieleinführungen oder durch Text illustriert. Wenn SuS durch Partizipantinnen geschauspielert werden sollten, erhielten diese sogenannte Charakterkarten. Diese Karten beschrieben, wie sich ihr Charakter zu verhalten habe. Die Karten wirken aus inhaltlicher Perspektive heute recht befremdlich. Zudem erhielten die Partizipantinnen Soziogramme und Arbeitsproben der simulierten SuS, ein nicht näher beschriebenes ‚Fakultätshandbuch‘ und ein ‚audiovisuelles Manual‘ (Cruickshank, 1988, S.137). Die Simulationspartizipantinnen verbrachten 15 Minuten in eigenständiger Stillarbeit zu verschiedenen Fragen und diskutierten dann ihre Ergebnisse zunächst in Kleingruppen, danach im Plenum. Die Kleingruppen waren dazu angehalten, Handlungsalternativen durch Rollenspiele zu explorieren (Cruickshank, 1988, S.138). Die Rollenspielpassagen der Simulation erinnern heutzutage mehr an *microteaching-peerteachings*, die Fragen, die zusätzlich noch gestellt worden sind, an Assessment (Cruickshank, 1971, S.12).

Inhaltlich geht es darum, die Simulationspartizipantinnen auf die „häufigsten und schwierigsten“ *classroom management* Probleme vorzubereiten (Cruickshank, 1988, S.137). Forschungsparadigmatisch steht der Simulator dem Behaviorismus nahe (Cruickshank, 1969, S.26), ist aber weniger rigide im Verständnis von Lernen wie die operante Konditionierung in Kersh's Mr. Land's Sixth Grade.

Das Inner-City Simulation Laboratory wurde unter anderem von Schulen in Montreal, Baltimore, Cleveland, Akron und Youngstown benutzt (Cruickshank, 1971, S.26). Das Teaching Problems Laboratory wurde (wahrscheinlich ohne Videokomponenten) kommerziell vertrieben und verkaufte sich bis 1971 ca. 300 mal; Tansey (1970, S.288) arbeitete mit dieser Simulation in England.

4.10 IVT (Interactive Virtual Training) / IVT-T (Interactive Virtual Training for Teachers))

Interaktionsmodellierung:	Typ 5 (Komplexe algorithmische Steuerung mit eingeschränktem Entscheidungsbaum) / Typ 1 (Deterministischer Entscheidungsbaum)
Immersionsform:	3D graphical simulation
Granularität:	Task simulation
Noch in Benutzung:	Ja
Vorgänger:	–
Nachfolger:	–

Der IVT (Delamarre et al., 2017) oder auch IVT-T (Delamarre, Lunn et al., 2019) Simulator wurde ab 2015 in einem interdisziplinären Forschungsprojekt zwischen Psychologinnen, Informatikerinnen und Systemingenieurinnen entwickelt, das durch das National Center of Educational Research (NCER) (Teil des Institute of Education Sciences (IES), was wiederum im U.S. Department of Education eingegliedert ist) gefördert wurde. Die Projektverantwortliche Elisa Shernoff der Rudgers University hatte die Leitung des psychologisch-erziehungswissenschaftlichen Verbundes inne, während der informatische Verbund unter der Leitung von Christine Lisetti an der Florida International University angesiedelt war. Zusätzlich unterstützten auch andere Institutionen wie das Centre Européen de Réalité Virtuelle und das LAB-STICC (Laboratoire des sciences et technologies de l'information, de la communication et de la connaissance) aus Frankreich das Projekt (Affective Social Computing Laboratory, 2016).

IVT benutzt in der graphischen Umsetzung die WebGL Komponente der Game Engine Unity, ist also somit in jedem Browser nutzbar und erfordert geringe technische Voraussetzungen (Delamarre et al., 2017, S.401). Leider kann auf die Demos des Simulators nicht mehr zugegriffen werden, da die verlinkten Seiten offline sind (<http://users.cis.fiu.edu/~gruiz/IVT/>). Da aber noch im Juni 2019 zum Projekt publiziert worden ist, kann davon ausgegangen werden, dass nach wie vor am Simulator gearbeitet wird und der Zugriff anderweitig erfolgt. Es existieren zwei separat modellierte Klassenräume mit jeweils fünfzehn individuell modellierten SuS der jeweils ersten und sechsten Klasse (Delamarre et al., 2017, S.400f. Delamarre, Lunn et al., 2019, S.185). Die simulierten Klassenräume sollen urbane amerikanische Brennpunktschulen abbilden und sind nach real existierenden Räumen modelliert. Die mo-

dellierten SuS sind analog dazu an die demographische Zusammensetzung in diesen Schulen angelehnt. Die SuS wurden vollständig via *motion-capturing* animiert. Alle verbalen Äußerungen können mittels vorab aufgenommener Audiodateien ausgelöst werden (Delamarre et al., 2017, S.402). Aussagen der Lehrkraft werden via Textfenster dargestellt, um eine bessere Identifikation der Partizipantinnen mit der von ihnen verkörperten Lehrkraft zu ermöglichen (Delamarre, Lunn et al., 2019, S.186). Alle SuS tragen in der Simulation Schuluniformen. Simulationspartizipantinnen betrachten den Raum immer aus der *first-person* Perspektive.

Thematisch behandelt die Simulation den Umgang mit Störungsverhalten (Delamarre et al., 2017, S.400). Es existieren insgesamt zwölf sogenannte Vignetten (alternative Bezeichnungen wären Szenarien oder Level), in denen die Simulationspartizipantinnen maximal 50 Pfade beschreiten können (Delamarre, Lunn et al., 2019, S.187). Als Pfad verstehen die Autorinnen eine dargestellte Szene, d.h. in der Regel werden nicht alle 50 Pfade durchschritten. In jeder Szene finden ca. 20 Ereignisse statt. Während eines Ereignisses kann die Situation entweder gleich bleiben oder SuS initiieren eine Handlung / verbale Äußerung. Nach der Szene können die Simulationspartizipantinnen aus einer von minimal 4 und maximal 9 möglichen Handlungsalternativen die für sie sinnvollste auswählen (Delamarre, Lunn et al., 2019, S.186). Der Simulator enthält drei verschiedene Schwierigkeitsstufen: Auf der niedrigsten Stufe kann man sich für die Entscheidungen so lange Zeit lassen wie man will, in der mittleren Stufe gibt es eine Zeitbegrenzung für die Auswahl (20–30 Sekunden) und in der höchsten Stufe gibt es zusätzlich zur Zeitbeschränkung noch Hintergrundgeräusche. Jede Vignette enthält drei mögliche Enden, egal welcher Pfad gewählt wurde. Während eine Vignette läuft, kann sie nicht pausiert werden (Delamarre, Lunn et al., 2019, S.186). Die SuS-Agenten können zwei generelle Stati besitzen, entweder sind sie disruptiv oder nicht-disruptiv. Die Auswahl, ob eine Schülerin disruptiv ist, erfolgt zufallsgesteuert (Delamarre, Lunn et al., 2019, S.186). Es ist allerdings unklar, ob nicht bestimmte disruptive Verhalten fest im Skript der Situation verankert wurden. Sind die SuS-Agenten disruptiv, so können sie entweder in einen *off-task* Zustand wechseln, wie z.B. aus dem Fenster schauen, oder aggressives Störverhalten zeigen. Dieses Störverhalten kann sich verbal oder physisch manifestieren. Zeigen sie nicht-disruptives Verhalten, steuert eine FSM (*finite state machine*) verschiedene Zustände von sozial erwünschtem Verhalten.

Die Prototypen des Simulators wurden von sechs Lehrkräften mit langer Unterrichtserfahrung in Brennpunktschulen evaluiert und der Simulator daraufhin angepasst (Delamarre et al., 2017, S.402).

Während die Simulation läuft, werden Simulationspartizipantinnen auf

Wunsch via Web-Cam aufgezeichnet. Während der Feedbackkomponente der Simulation diskutieren die Partizipantinnen mit einer Expertin und können die Web-Cam-Aufzeichnungen zur Reflexionshilfe benutzen (Delamarre et al., 2017, S.401). Zielgruppe der Simulation sind Lehrkräfte mit weniger als drei Jahren Berufserfahrung (Delamarre et al., 2017, S.400).

Es ist geplant den Simulator mit einer Komponente zu erweitern, die auch Emotionen algorithmisch steuert. Die Entwicklung am Simulationsmodell dieser AIMER (Appraisal Interpersonal Model of Emotion Regulation) genannten Komponente ist zum Zeitpunkt des Verfassens dieser Arbeit noch nicht abgeschlossen (Delamarre, Buche & Lisetti, 2019, S.183).

4.11 Kognito Interactive Simulations

Interaktionsmodellierung:	Typ 1 (Deterministischer Entscheidungsbaum) / Typ 3 (Komplexe algorithmische Steuerung mit Entscheidungsbaum)
Immersionsform:	3D graphical simulation
Granularität:	Task simulation
Noch in Benutzung:	Ja
Vorgänger:	–
Nachfolger:	–

Bei den Simulationen von Kognito Interactive handelt es sich um 29 einzelne, kommerzielle Simulationen bzw. Szenarien mit der Zielgruppe von Lehrkräften und Lehramtsstudierenden. Kognito Interactive, ansässig in New York, produziert außerdem noch weitere Simulationen für drei andere Märkte, nämlich für die Hochschulbildung, das Gesundheitswesen und NGOs. Kognito verkauft die Einzellizenzen für \$30–60 USD pro Szenario (45–120 Minuten Länge) und Lizenzbündel für Bildungseinrichtungen. Auf Anfrage fertigt Kognito auch neue Simulationsszenarien innerhalb ihres Frameworks an.

Inhaltlich behandeln die Simulationen für alle Märkte ähnliche Themen: Umgang mit Medikamentenmissbrauch und Drogen, Suizidprävention und andere Themen mit Bezug auf psychische Störungen oder Konfliktprävention. Das Framework der Simulation sieht immer ein Einzelgespräch zwischen zwei Gesprächspartnerinnen vor, wird also nicht zur Simulation von Unterricht genutzt. Explizit werden in einem Übersichtsartikel zu Simulationen von Bradley und Kendall (2014) die Szenarios „At-Risk for High School Educators“, „At-Risk for Middle School Educators“ (Themen: Depressionen und

Suizidprävention bei SuS, psychische Störungen) und „Stand Up, Speak Up!“ (Thema: LGBTQ) erwähnt. Die Charaktere sind dreidimensional, sehr detailliert, durchgehend animiert und werden aus der Third-Person Perspektive dargestellt. Die ästhetische Umsetzung der Figuren und Hintergrundumgebung könnte man als ‚realistischen Comiclook‘ bezeichnen, der sich an die Gestaltung der Figuren des Computerspiels ‚Die Sims‘ anlehnt. Die Simulationsteilnehmerinnen können aus einer Reihe vorgefertigter Fragen und Aussagen auswählen, wie das Gespräch weitergehen soll. Sowohl die Aussagen des Charakters, den die Simulationsteilnehmerinnen übernehmen, als auch die Aussagen der Gesprächspartnerin wurden von Synchronsprecherinnen eingelesen und werden nach Auswahl einer Handlungsalternative abgespielt. Obwohl die Fragen häufig eine Variable wie z.B. das sogenannte „Trust-Meter“ beeinflussen, sind die algorithmischen Interaktionen jedoch so gering, dass man in der Interaktionsmodellierung eher von einem Typ 1 sprechen müsste, u.a. weil jede Handlungsalternative stets dieselbe Reaktion hervorruft.

Obwohl die Studien über beeindruckende Populationszahlen verfügen, z.B. 18.896 Lehrkräfte, Lehramtsstudierende und Schulpersonal für die „At-Risk“-Simulationen, so muss doch kritisch angemerkt werden, dass sämtliche Studien von Mitarbeiterinnen von Kognito, insbesondere von dessen Co-Gründer Dr. Glenn Albright verfasst wurden und bis auf zwei Ausnahmen nicht über Journals publiziert worden sind, sondern nur über die eigene Webseite (Albright et al., 2013b; Albright et al., 2013a; Albright et al., 2016, diese Studie wurde vom mHealth Journal veröffentlicht: Bradley & Kendall, 2019, diese Studie wurde vom Journal of Technology in Behavioral Science veröffentlicht). Kognito weist darauf hin, dass die selbst publizierten Studien in dem NREPP (National Registry of Evidence-Based Programs and Practices) gelistet sind, einem Register, was von der SAMHSA (Substance Abuse and Mental Health Services Administration) bereitgestellt wurde, die wiederum dem U.S. Department of Health and Human Services untersteht. Mit dem Beginn des Jahres 2018 wurde das NREPP durch das SAMHSA eingestellt, was kontrovers diskutiert wurde (<https://pgdf.org/samhsas-registry-of-evidence-based-programs-nrepp-suspended/>).

4.12 Lesson Sketch / SimTeach / ThEMaT I-III

Interaktionsmodellierung:	Typ 3 (Komplexe algorithmische Steuerung mit Entscheidungsbaum)
Immersionsform:	2D graphical simulation (z.T. animiert)
Granularität:	Task simulation
Noch in Benutzung:	Unklar
Vorgänger:	–
Nachfolger:	–

Die zwei Simulationen SimTeach und ThEMaT I-III nutzen dieselbe Engine bzw. denselben Simulator, welcher LessonSketch heißt. Sie unterscheiden sich in den dargestellten Unterrichtsszenarien. Alle Szenarien haben fachdidaktisch-mathematischen Bezug und modellieren eine Unterrichtsaufgabe – nicht den gesamten Unterricht – in detaillierter Granularität. Der Simulator zeigt im einfachen Comic-Stil gehaltene Bilder, zum Teil mit, zum Teil ohne Animationen, welche die Unterrichtssituation illustrieren (Chieu & Herbst, 2011, S.107). Die Figuren sind stark abstrahiert in blauer Hautfarbe abgebildet.

Die Simulationsteilnehmerinnen können sich für eine von vier vordefinierten Handlungsalternativen entscheiden. Wie die Simulation fortgesetzt wird, entscheidet nicht nur die Auswahl der Handlungsalternative, sondern auch, welche Handlungsalternativen zuvor von den Teilnehmerinnen ausgewählt worden sind – mathematisch implementiert über bayer'sche Netzwerke (Chieu & Herbst, 2011, S.111f.). Die Fortsetzung der Simulation enthält nicht nur eine Beschreibung des Unterrichtsgeschehens, das auf die Handlungsalternativen folgt, sondern auch Seiten mit Hinweisen, Links oder Lernressourcen. Welche Szene in der Simulation folgt, ist abhängig davon, wo der Algorithmus noch Trainingsbedarf sieht. Dieses Simulationskonzept ist stark beeinflusst vom dem medizinischen Simulator Technology Enhanced Learning Environment for Orthopedic Surgery (TELEOS) (Chieu & Herbst, 2011, S.107). Ziel der Simulation ist es durch Komplexitätsreduktion, Reflexion und automatisiertem Feedback Lernen anhand des cKc Modells zu ermöglichen (Chieu & Herbst, 2011, S.107, S.113). Es wurde ein Editor entwickelt, der es Fachdidaktikerinnen ermöglicht, eigene Lernszenarien zu erstellen, einmal „Lesson Sketching Tool“ (Chieu & Herbst, 2011, S.107), einmal „Depict“ (Dimmel et al., 2016, S.1) genannt. Diese Tools enthalten unter anderem graphische Schablonen für Gegenstände und Figuren. Da der Simulator technisch via Flash

betrieben wurde, was von allen gängigen Browsern zum Zeitpunkt dieser Arbeit nicht mehr unterstützt wird, ist davon auszugehen, dass der Simulator nicht mehr aktiv verwendet wird. Er wird aber noch in neuerer Forschungsliteratur unter dem Namen LessonSketch erwähnt (Sweeney et al., 2018, S.678; Weston et al., 2018, S.153).

4.13 Mr. Land's Sixth Grade

Interaktionsmodellierung:	Typ 1 Deterministischer Entscheidungsbaum / Typ 6 Human-in-the-loop / Typ 8 Reale Handlungsdurchführung
Immersionsform:	Interactive Video-Based Simulation (IVBT)
Granularität:	Task Simulation / Teacher simulation ⁶
Noch in Benutzung:	Nein
Vorgänger:	–
Nachfolger:	Low-Cost Instructional Simulation

Mr. Land's Sixth Grade ist eine durch Videos und andere Medien unterstützte, nicht-digitale Simulation des Oregon College of Education unter Leitung von Bert Y. Kersh, deren Entwicklung 1961 (Twelker, 1967, S.197) begann und die in der Mitte des Jahres 1962 fertiggestellt wurde (Kersh, 1963, S.28, 42). Ähnlich wie in der bereits vorgestellten Begriffsdefinition distinguieren Kersh (1963, S.i, S.4f.) und Twelker (1967, S.198) die Begriffe ‚Simulation‘ und ‚Simulator‘. Der Simulator – im Sinne der physischen Forschungseinrichtung mit Projektoren, Leinwänden und Stühlen – wird als Classroom Simulator bezeichnet (Cruickshank, 1971, S.8; Kersh, 1962, S.109; Kersh, 1963, S.32). Ob der generischen Simulatorbezeichnung ‚Classroom Simulator‘ und der häufigeren Verwendung der Simulationsbezeichnung wird in dieser Arbeit der Begriff ‚Mr. Land's Sixth Grade‘ gewählt.

Mr. Land's Sixth Grade wird bei einigen Autorinnen als die erste oder eine der ersten Simulationen in der Lehrkräftebildung bezeichnet (Cruickshank, 1988, S.134; Tansey, 1970, S.283; Brown, 1999, S.313). Nach eingehender Analyse ist dies zu bestätigen. Kersh (1963, S.2) verweist auf einige Vorarbeiten wie die ab 1959 durchgeführte DCS-Studie (Development of Criteria for Success), die die Umsetzung von Mr. Land's Sixth Grade beeinflusst hat. Es sind zwar methodische Ähnlichkeiten zu erkennen (Videoverwendung, Rollenspielelemente), aber die in der DCS-Studie verwendete Simulation war für Schulleiterinnen konzipiert. Diese Simulation fokussierte nicht auf die Schulung von Unterrichtshandeln (Cunningham, 1959, S.4f.).

Für Mr. Land's Sixth Grade wurden 60 Klassenraumprobleme plus jeweils 2–3 Fortführungssequenzen (insgesamt ca. 120–180 Videoausschnitte) mit einer Länge von ca. 30 Sekunden (Kersh, 1963, S.25) im Rahmen einer *staged*⁷ Videoaufzeichnung gefilmt (Kersh, 1963, S.5; Cruickshank, 1971, S.8). Die Filme wurden aus der Perspektive der fiktiven Lehrkraft Mr. Land aufgezeichnet, der selbst nicht zu sehen ist und dessen Rolle die Simulationspartizipantinnen einnehmen (Kersh, 1963, S.27). Die zweiundzwanzig Kinder, die die SuS in der Simulation spielen, wurden von einer realen Lehrkraft gecastet und spielen ihre eigenen Persönlichkeiten, jedoch unter einem Pseudonym. Sie übten auch Einfluss auf den Regisseur der Filme aus, damit dieser das Skript abänderte, so dass sie es als realistischer empfanden (Kersh, 1963, S.22, 5).

Die 60 Klassenraumprobleme wurden in drei Teile à 20 Sequenzen geteilt, die jeweils Ausschnitte aus einem simulierten Schultag zeigten (Kersh, 1963, S.38, 23, 25). In der Hauptstudie von Kersh (1963, S.23) wurden diese Videos N=40 Lehramtsstudierenden, welche die Zielgruppe des Simulators bildeten, sukzessive präsentiert. Es wurden zudem andere schriftliche Materialien bereitgestellt wie (1) standardisierte Testdaten der fiktiven SuS, (2) Krankendaten der SuS, (3) eine Anzahl von ‚Lehrkraft-Anekdoten‘ zu den einzelnen SuS, sowie (4) eine Beschreibung der Schule.

Die Studie bestand aus drei Phasen: Dem Pretest, dem Treatment und dem Posttest (Kersh, 1963, S.25). In jeder Phase wurden im rotierenden Verfahren 20 Klassenraumsequenzen verwendet. Die Lehramtsstudierenden wurden in vier Gruppen eingeteilt, welche sich durch die verwendete Simulationsfidelität im Treatment unterschieden. Gruppe 1 lernte in der Treatmentphase mittels Bewegtbildern in Lebensgröße. Diese wurden von drei 16mm Filmprojektoren auf eine Leinwand projiziert. Gruppe 2 arbeitete auch mit Bewegtbildern, allerdings mit kleineren Projektionen, Gruppe 3 mit statischen Bildern in Lebensgröße. Die Ereignisse während der Simulation wurden in dieser Gruppe durch die Untersuchungsleiterin vorgelesen. Gruppe 4 arbeitete mit statischen Bildern in kleinerer Projektion (Kersh, 1963, S.8). Die Pre- und Posttests wurden bei allen Gruppen mittels lebensgroßer Bewegtbilder durchgeführt (Kersh, 1963, S.9).

Im Simulator kann sich immer nur eine Partizipantin befinden (Cruickshank, 1988, S.136). Bevor eine Problemsequenz startet, ist die Partizipantin dazu angehaltenen die ggf. vorliegenden schriftlichen Materialien zu lesen. Danach wird das Video abgespielt. Anschließend spielt die Partizipantin in

⁷In sogenannten *staged* Videoaufzeichnungen werden SuS von schauspielenden Kindern verkörpert. Der Begriff steht im Gegensatz zu authentischen Unterrichtsaufzeichnungen, bei welchen reale SuS in realem Unterricht mitwirken.

einem rollenspielähnlichen Setting ihre präferierte Handlung in Richtung der SuS auf der Leinwand aus. Ein sogenannter *experimenter* stellt nun fest, welche der 2–3 Folgevideos das Geschehen angemessen fortsetzen würden (Kersh, 1963, S.5). Kriterien und Standards dafür wurden im Vorfeld von einem aus drei Personen bestehenden Expertinnengremium aus erfahrenen Lehrkräften festgelegt (Kersh, 1963, S.6, 22, 28, 36). Eine weitere Person unterstützt den *experimenter* in der technischen Bedienung der Projektoren (Kersh, 1963, S.32). Eines der Folgevideos zeigt ein SuS-Verhalten, dass nach einer Lehrkrafthandlung folgt, die das Expertinnengremium als korrekt befunden hat. In anderen Worten: Es gibt eine richtige Folgesequenz, aber mehrere Lösungen, um diese zu erreichen (Cruickshank, 1988, S.134; Kersh, 1963, S.6f.). Erreichten die Simulationspartizipantinnen durch ihr Handeln nicht die korrekte Folgesequenz, so wurde die Situation so lange wiederholt, bis dies eintraf. Es war geplant bei mehr als zehn Wiederholungen die Simulationssequenz abubrechen, dies wurde in der Studie jedoch nie erreicht (Kersh, 1963, S.38). Der *experimenter* war dazu angehalten keine direkten Handlungsempfehlungen auszusprechen, unterstützte durch Nachfragen aber die Suche der Simulationspartizipantinnen nach geeigneten Handlungsalternativen (Kersh, 1963, S.40).

Um die 20 Problemsequenzen zu absolvieren, benötigten die Simulationspartizipantinnen ca. eine Stunde Realzeit (Kersh, 1963, S.38). Nach der Simulation reflektierten alle Partizipantinnen und der *experimenter* ihre Simulationserfahrungen (Kersh, 1963, S.6). Der Post-Test fand eine Woche nach dem Treatment statt (Kersh, 1963, S.44). Während Pre- und Posttest gab es keine Intervention des *experimenters* und keine Wiederholungen der Sequenzen.

Inhaltlich beziehen sich die vorgestellten Problemsequenzen schwerpunktmäßig auf *classroom management* bezogene Themen wie Diagnose und Lösungsansätze für Unaufmerksamkeit, Verwirrung, Ablenkung und Müdigkeit bei SuS (Kersh, 1963, S.4). Das Forschungsparadigma von Kersh ist – auch nach eigener Aussage – fest im Behaviorismus und operanten Konditionieren verwurzelt (Cruickshank, 1988, S.138; Kersh, 1963, S.14).

Mr. Land's Sixth Grade hatte in den 60er und 70er Jahren eine gewisse Rezeption gefunden und wurde anschließend von Twelker (1967) in eigenen Studien benutzt. Auch einige modernere Studien verweisen auf Mr. Land's Sixth Grade (z.B. G. R. Girod et al., 2007, S.209).

4.14 SimAula 1.0

Interaktionsmodellierung:	Typ 3 (Komplexe algorithmische Steuerung mit Entscheidungsbaum)
Immersionsform:	3D graphical simulation
Granularität:	Lesson simulation
Noch in Benutzung:	Nein, Projektende 31. Oktober 2012, Downloads zur Demo-Version des Prototyps funktionieren nicht mehr
Vorgänger:	–
Nachfolger:	–

SimAula (auch SimAULA) ist eine Simulation, die im Rahmen eines länderübergreifenden Verbundprojekts des europäischen Programms ‚Education, Audiovisual and Culture Executive Agency Lifelong Learning: Comenius, ICT and Languages‘ (2010–2012) erstellt worden ist. Beteiligte Länder an dem Projekt waren Spanien, Italien, Bulgarien, Griechenland und Großbritannien.

SimAula wurde in der Game-Engine Unity programmiert und orientiert sich in seiner blockhaften Ästhetik an Spielen wie Minecraft. SuS-Figuren sind algorithmisch gesteuert, geben aber kein visuelles Feedback zu Mimik oder Arbeitsvorgängen. Die SuS-Avatare stehen für die Dauer der Simulation statisch hinter ihren Tischen. Auch die Darstellung des Klassenraums als offene Fläche auf einer Wiese ist ungewöhnlich.

Die Modellierung des SuS-Verhaltens wurde nicht selbst durch die beteiligten Wissenschaftlerinnen entwickelt, sondern an Indra abgegeben, eine externe Technologie- und Beratungsfirma, die das Projekt mitkoordinierte (Indra, 2010). Wie genau das Verhalten modelliert wurde, legt Indra leider nicht offen.

Die Partizipantinnen übernehmen die Rolle einer Biologielehrkraft und behandeln im Prototyp das Thema ‚Die Zelle‘. Im Rahmen des Projektes waren zusätzliche Fächer und Unterrichtsthemen geplant, aber es ist unklar, inwieweit diese Planungen umgesetzt worden sind. Am Anfang wählen die Simulationspartizipantinnen eines von drei möglichen didaktischen Konzepten: *experiment-based learning* (EL), *problem-based learning* (PBL) oder *collaborative learning* (CL). Anschließend wählt man eine von mehreren vorgegebenen Lernaktivitäten in Textform aus, die man für den am besten geeigneten Ansatz hält, um das didaktische Konzept umzusetzen. Zusätzlich kann man bestimmen, ob SuS bestimmte Lernmaterialien (wie Mikroskope, Power-Point-

Präsentationen, Laptops, etc.) nutzen sollen oder nicht. Mittels eines Sliders wird festgelegt, wie lange eine Aktivität durchgeführt werden soll. Insgesamt werden pro Simulationsdurchgang drei Lernaktivitäten ausgewählt, während derer einzelne SuS Fragen stellen und Aussagen machen, auf die mittels eines Single-Choice-Dialogs reagiert werden kann (Suev, 2013; Clarke et al., 2016).

Am Ende erhalten die Simulationspartizipantinnen einen summarischen Punktwert für die Beteiligung der SuS und einen Text, der beschreibt, welche Handlungen mit dem didaktischen Konzept kompatibel sind. Eine grafische oder tabellarische Auswertung ist nicht vorhanden.

SimAula wurde wenig in der Forschungsliteratur rezipiert (Chuang & Shih, 2014, S.343).

4.15 simClass / simClass II

Interaktionsmodellierung:	Typ 3 (Komplexe algorithmische Steuerung mit Entscheidungsbaum)
Immersionsform:	2D graphical simulation (,Flash SimClass‘), (Hyper-)Text-based simulation (,Web SimClass‘, ,SimClass II‘)
Granularität:	Lesson simulation
Noch in Benutzung:	Nein, Webseite offline
Vorgänger:	–
Nachfolger:	–

SimClass ist eine südkoreanische Unterrichtssimulation, die von DonGuk Cheong und SunHwa Kim unter Leitung von Dr. Youngkyun Baek von der Korea National University of Education entwickelt wurde (Gibson, 2009a, S.672; Cheong & Kim, 2009). Problematisch ist, dass drei unterschiedliche Simulationen mit dem Begriff SimClass bezeichnet werden:

- (a) Die erste Version, im folgenden als ,Flash SimClass‘ bezeichnet, ist webbasiert und wurde mit Macromedia FlashMX und PHP entwickelt, wobei z.T. Animationen vorhanden sind (Cheong & Kim, 2009, S.67). Inhaltlicher Fokus dieser Version ist *classroom management*, insbesondere Umgang mit Störungsverhalten (Cheong & Kim, 2009, S.64). Die Simulation zeigt vier SuS, die 48 verschiedene Zustände von (Stör-)Verhalten zeigen. Gedanken und Aussagen der SuS werden mittels Sprechblasen dargestellt (Cheong & Kim, 2009, S.68f.). Die Algorithmen zur Interaktionsmodellierung wurden zunächst über Excel prototypisiert

und anschließend in die Simulation überführt, sind aber vergleichsweise einfach gehalten (Cheong & Kim, 2009, S.67). So gibt es nur drei Variablen, die das SuS-Verhalten beeinflussen, nämlich „Selbstbewusstsein“, „Selbstwirksamkeit“ und „Selbstkontrolle“, die die Ausprägungen „sehr niedrig“, „niedrig“, „normal“ und „hoch“ haben können (Cheong & Kim, 2009, S.68f.).

Auch in diesem Simulator findet man die Trennung zwischen situationsbezogenen und personenbezogenen Variablen (siehe [Abschnitt 5.3](#), S. 155). Die personenbezogenen Variablen können über ein virtuelles Schulaktensystem eingesehen und sich über einen simulierten Austausch mit virtuellen Lehrkräftekolleginnen erschlossen werden.

- (b) Die zweite Version, im folgenden als ‚Web SimClass‘ bezeichnet, basiert auf PHP und MySQL. In dieser Version werden sechs SuS modelliert, allerdings ohne jede Form von Animation (Kim & Cheong, 2009, S.302). Die Interaktionsmodellierung ist komplexer: Alle SuS verfügen über sieben Arten von Intelligenzen (linguistische, logisch-mathematische, spatiale, körperlich-kineästhetische, musische, intrapersonelle und interpersonelle), die anschließend zu einer Gesamtgröße verrechnet werden. Außerdem besitzen alle SuS die fünf OCEAN/Big-Five-Persönlichkeitsmerkmale (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism), welche zu einem Motivationslevel verrechnet werden. Beide Variablen sind nach dem Verständnis der Zone der proximalen Entwicklung von Vygotsky modelliert (Vygotskij, 1987b; Vygotskij, 1987a; Kim & Cheong, 2009, S.293), reichen von 0 bis 100 und bestimmen den Erfolg bzw. das Verhalten der SuS (Kim & Cheong, 2009, S.298). Die Simulationsteilnehmerinnen können diese Werte über die Handlungsalternativen „SuS motivieren“, „Aufmerksamkeit der SuS einfordern“ oder „Lerngelegenheiten allokatieren“ beeinflussen (Kim & Cheong, 2009, S.299). SuS-Verhalten wird in fünf Ausprägungen modelliert: „gelangweilt“, „herausgefordert“, „ängstlich/unruhig“, sowie positiv oder negativ auf motivationales Feedback reagierend (Vgl. auch Baek, 2009, S.42).
- (c) Eine dritte Version, die als ‚SimClass II‘ bezeichnet wird, ist ebenfalls webbasiert und beschäftigt sich mit der Schulung von motivierendem Lehrkraftverhalten. Diese Version wurde sowohl für die Schulung von Lehramtsstudierenden, als auch für Lehrkräfte konzipiert (Cheong & Kim, 2009, S.233). Die Interaktionsmodellierung berücksichtigt ausschließlich motivationale Faktoren und basiert auf dem sogenannten ARCS-Modell von John M. Keller (J. Keller & Kopp, 1987; J. M. Kel-

ler, 1983). Die Variablen „Attention“, „Relevance“, „Confidence“ und „Satisfaction“ bestimmen das Verhalten der simulierten SuS. Genau so wie in ‚Flash Sim Class‘ können die Simulationspartizipantinnen SuS-Akten einsehen und Informationen von virtuellen Lehrkräftekolleginnen einholen, um Hinweise zur Ausprägung der SuS-Variablen zu bekommen.

SimClass wird vereinzelt in der anglo-amerikanischen Literatur erwähnt (Gibson, 2009a, S.672; Christensen et al., 2011, S.215; Baek, 2009, S.42), hat allerdings keine Weiterentwicklung oder Nutzung außerhalb der Studien von Cheong, Kim & Baek erhalten, insbesondere nachdem sich Gonguk Cheong 2010 rollenspielgestützten Unterrichtssimulationen in Second Life zugewendet hat (Cheong, 2010). Die offizielle Seite <http://www.simclass.co.kr> ist heutzutage nicht mehr verfügbar und wurde auch nicht in der Wayback-Machine archiviert.

4.16 simSchool 1.0

Interaktionsmodellierung:	Typ 4 (Komplexe algorithmische Steuerung ohne Entscheidungsbaum)
Immersionsform:	2D graphical simulation
Granularität:	Lesson simulation
Noch in Benutzung:	Nein, aufgegangen in simSchool 2.0, Nutzung bis 2018
Vorgänger:	–
Nachfolger:	simSchool 2.0

4.16.1 Einführung und Entwicklungsgeschichte

simSchool ist neben TeachLivE eine der wichtigsten und am weitesten verbreiteten kommerziellen, digitalen Simulationen in der Lehrkräftebildung. Auch wenn es an einigen Stellen als *serious game* (Gibson, 2014a, S.148; McPherson et al., 2011, S.93f. Tyler-Wood et al., 2015, S.17) bezeichnet wird, findet sich in der Literatur häufiger eine Identifikation als Simulation bzw. die Begriffe werden synonym verstanden. simSchool ist insbesondere für diese Arbeit von großer Bedeutung, da das benutzte Konzeptions- und Simulationsmodell – wenn auch an vielen Stellen kritisierbar – elaboriert ist und bis zu einem gewissen Grad transparent wissenschaftlich vorgestellt wird.

Die Idee für simSchool entstand beim Zusammentreffen von David Gibson, Melanie Zibit und B. Halverson mit M. Searson an der Society for Information Technology and Teacher Education 13th International Conference in Nashville im Jahre 2002 (Gibson, 2012a, S.10). Aus diesen Gesprächen entstand schnell ein Proposal, das im Rahmen des U.S. Department of Education PT3 Programms akzeptiert wurde und den Forscherinnen eine Initialfinanzierung von 1.7 Millionen Dollar ermöglichte. Die Projektstruktur war ein Verbund zwischen der Western Oregon University und den Vermont Instituten (Gibson, 2012a, S.10). Der erste Prototyp, entwickelt von David Gibson, der später die Firma CurveShift Inc. gründete (Hettler et al., 2008, S.2), wurde 2003 fertiggestellt. Weitere Studien in den Jahren 2007, 2009 und 2011 wurden genehmigt und führten neben neuen wissenschaftlichen Erkenntnissen auch zu einer iterativen Verbesserung des Simulators (Gibson, 2012a, S.10). Im Rahmen dieser Anschlussforschung wurde auch ein Kooperationsvertrag mit Gerald Knezek und Rhonda R. Christensen von der University North Texas (UNT) geschlossen, wobei das ‚national center of research for simSchool‘ an der UNT eingerichtet wurde (Vgl. auch Gibson et al., 2013, S.1320). Der erste Einsatz von simSchool in der Lehrkräftebildung fand 2007 an der UNT, gefördert durch den Fund for the Improvement of Postsecondary Education (FIPSE), statt. Die Studien ab dem Jahr 2011 wurden vom EDUCAUSE Next Generation Learning Challenges (NGLC) Programm der Bill & Melinda Gates Foundation finanziert (Hopper, 2018, S.108). Es wurde eine Partnerschaft mit Pragmatic Solutions eingegangen, welche eine Software zur Useranalyse von simSchool Sessions bereitstellte (Gibson & Jakl, 2013, S.1f.). Die Nutzerinnengruppe explodierte in dieser Zeit von 600 auf 10,000 Nutzerinnen (Monsour, 2012; Gibson et al., 2013, S.1316) innerhalb von neun Monaten. Laut Gibson (2012a, S.1) benutzten 100 Colleges und Universitäten in 164 Ländern simSchool, mit vergleichbaren Werten im Jahr 2013 (Gibson et al., 2013, S.1316). Eine Nutzungskarte zeigt, dass sich der überwiegende Anteil an Institutionen zwar in Nordamerika befindet, simSchool jedoch auf allen Kontinenten Verbreitung gefunden hat (Gibson et al., 2013, S.1316). Im Jahre 2014 berichtete Gibson (2014b) über 11,936 simSchool Nutzerinnen in 115 Institutionen und 146 Ländern. Das Nutzungsdaten aus dem Jahr 2018 zeigen nur noch ein graduelles Wachstum mit 13,000 Nutzerinnen und Anwendung in 156 Ländern (Hopper, 2018, S.79).

simSchool ist eine rein webbasierte Anwendung (Badiee & Kaufman, 2015; Hettler et al., 2008, S.2).

4.16.2 Grafischer Aufbau und UI

simSchool ist eine zweidimensionale graphische Simulation in einer simplen Comic-Ästhetik. Die SuS werden über einen festen, nicht änderbaren Blickwinkel aus der Perspektive der Lehrkraft dargestellt (Zibit & Gibson, 2005, S.9). Es gibt fünf visuell unterschiedliche SuS-Schablonen, die durch Änderungen an der Haut- und Haarfarbe, Haarart und Kleidungsfarbe weiter differenziert werden können (Deale & Pastore, 2014, S.211). Die Auswahl dieser Parameter erfolgt zufallsgesteuert, wobei die Simulationspartizipantinnen auswählen können, ob sie in einer ländlichen, suburbanen oder urbanen Schule unterrichten möchten. Daraufhin wird die Zusammensetzung der SuS nach demographischen Faktoren angepasst (Zibit & Gibson, 2005, S.2; Deale & Pastore, 2014, S.211). Innerhalb der Simulation wechseln die SuS je nach Situation zwischen verschiedenen Körperposen und Verhaltensweisen für SuS (z.B. träumen, schlafen, Musik hören, schreiben), diese sind aber nicht animiert, sondern bestehen aus Standbildern. Nach der Verteilung einer Aufgabe wird ein Tachometer und ein Thermometer über den Köpfen der SuS angezeigt. Das Thermometer zeigt an, wie geeignet der Task für die Schülerin ist, das Tachometer zeigt an, wie schnell die Schülerin lernt (Hettler et al., 2008, S.4; Deale & Pastore, 2014, S.206).

Wählt man eine Schülerin durch Anklicken aus, werden drei Werte am unteren Bildschirmrand angezeigt, nämlich ‚Power‘, ‚Happiness‘ und ‚Academic‘.

simSchool simuliert SuS der Jahrgangsstufe 7–12 (Zibit & Gibson, 2005, S.1; Badiie & Kaufman, 2015, S.3). Die Schülerinnenanzahl in simSchool ist variabel und kann zwischen einem und zwanzig Schülerinnen liegen (Deale & Pastore, 2014, S.211; Hopper, 2014, S.24 gibt achtzehn an, was sich aber nicht mit den Screenshots von simSchool im selben Artikel deckt, die zwanzig SuS anzeigen). Die Anzahl der SuS ist von den Partizipantinnen vor Simulationsbeginn festlegbar.

Die Simulation in simSchool bzw. das Ablaufen der simulierten Zeit, kann jederzeit angehalten und wieder fortgesetzt werden.

4.16.3 Interaktionsmodellierung

Im Gegensatz zu den meisten anderen Simulationen, aber ähnlich wie VCS von Simon Skrødal, existiert kein festgelegtes Szenario mit Entscheidungsbaum und damit kein expliziter fachlicher Inhalt der simulierten Stunde. Simulationspartizipantinnen können beispielsweise SuS auffordern, ihre Hausaufgaben vorzutragen, aber die konkrete inhaltlich-sprachliche Ausgestaltung wird nicht modelliert. Damit ist simSchool generell fachunabhängig. Es gab die Entwicklung einzelner Szenarios mit spezifischem Fachbezug, z.B. ein

Erdkunde-Szenario mit dem Thema ‚Leben in einem Teich‘ (Christensen et al., 2011, S.213). Diese Szenarien, auch als Module bezeichnet, konnten auch von NutzerInnen erstellt werden. Gibson et al. (2013, S.1319) geben die Zahl von öffentlich zugänglichen Modulen mit 128 an. Wie spezifisch der fachliche Inhalt modelliert wurde, ist leider intransparent. Es ist davon auszugehen, dass es sich dabei um eine Anpassung der Fragebausteine, sowie spezifizierte Zusammensetzungen einer Klasse und nicht die Implementierung von geskripteten Ereignissen (und damit Entscheidungsbäumen) handelte. Scheinbar widersprüchlich dazu ist, dass Gibson auf einem Vortrag im Jahr 2014 (Gibson, 2014b) angibt, dass solche bereichsspezifischen Fragen noch nicht existieren. Dies könnte sich auf die Basisversion von simSchool (ohne Module) beziehen, bleibt aber unklar. Deale und Pastore (2014, S.213) erwähnen, dass es für NutzerInnen möglich war eigene Tasks mit dem ‚SimCreator Feature‘ zu erstellen.

Die simulierten SuS in simSchool besitzen eine reaktive algorithmische Steuerung. Dies bedeutet, dass SuS nicht von sich aus Handlungen initiieren oder ihre MitschülerInnen beeinflussen können, sondern dass ihr Verhalten ausschließlich durch ihr Persönlichkeitsprofil und die Handlungen der Lehrkraft determiniert werden (Badiie & Kaufman, 2015, S.4; Hopper et al., 2013, S.2872; Christensen et al., 2011, S.205; Gibson & Jakl, 2013, S.2; Gibson, 2014a, S.151; Tyler-Wood et al., 2015, S.19). Auch wenn die Konstruktion von simSchool insgesamt eher dem kognitivistischen Lehr-/Lernparadigma zuzuordnen ist, zeigt sich in der interaktionalen Simulationsmodell-Ausgestaltung ein behavioristisch inspiriertes Reiz-Reaktionsschema. Mit Programmieraufwand hätte das Simulationsmodell so gestaltet werden können, dass SuS von sich aus Handlungen initiieren. Damit hätte der Eindruck größerer *agency* erzeugt werden können. Gibson selbst gesteht das ein, wenn er davon spricht, dass die modellierten SuS zum Teil „pavlovian creatures“ und „popperian creatures“ sind (Gibson, 2009a, S.687f.). Ob eine identische Schülerin auf einen identischen Stimulus der Lehrkraft immer identisch reagiert, oder ob Heuristiken verwendet werden (wie Christensen et al., 2011, S.204 anzudeuten scheint), um leicht abweichende Ergebnisse zu erzielen, geht aus der Literatur nicht vollständig hervor. Der Algorithmus von simSchool aktualisiert den Zustand der SuS alle zehn Sekunden (Gibson & Jakl, 2013, S.4; Hopper et al., 2013, S.2872; Gibson, 2014b).

simSchools Algorithmus ist ein sogenannter Bergsteigeralgorithmus (Hopper et al., 2013, S.2871f. Deale & Pastore, 2014, S.210; Gibson, 2009a, S.680) was ein einfaches heuristisches Optimierungsverfahren bezeichnet, das beendet wird, sobald ein lokales oder im besten Falle globales Maximum gefunden ist. Alle simulierten SuS versuchen ein Equilibrium basierend auf physischen, emotionalen und kognitiven Persönlichkeitsvariablen und Handlungen der

Lehrkraft zu erreichen.

4.16.4 Theoretische Grundlage / Konzeptionsmodell

Die theoretische Grundlage zur Modellierung des Lernens von SuS-Agenten bietet das von David Gibson erarbeitete COVE-Modell, welches wiederum auf vier theoretischen Ebenen basiert und ein komplexes Konstrukt ist. Die theoretische Basis für COVE speist sich aus zwei Frameworks und einer Theorie:

- (a) Dem „How people learn“ Framework des National Research Councils (Bransford et al., 2000; Gibson, 2009b, S.120).
- (b) Einer nicht unproblematischen forschungsparadigmatischen Vermischung von Behaviorismus, Kognitivismus und Konstruktivismus (von Gibson, 2009b BCC Theorien genannt), mit Nennung aber ohne explizite Bezugnahme auf das ADDIE-Modell⁸ (Dick et al., 2004; Gibson, 2009b, S.121), das ASSURE-Modell⁹ (Heinich et al., 1992; Gibson, 2009b, S.121) und das ARCS-Modell¹⁰ (J. Keller & Kopp, 1987; J. M. Keller, 1983; Gibson, 2009b, S.121).
- (c) eine modifizierte Form der Cattell-Horn-Carroll (CHC) *theory of intelligence* (Gibson, 2009b, S.123; Gibson, 2011, S.2).

Gibson ordnet den einzelnen C-, O- und V-Ebenen des COVE Modells (Erklärung des Akronyms folgt später im Text) die 16 generellen Fähigkeiten nach der CHC Theorie zu, wobei einige dieser Zuordnungen diskussionswürdig sind. Es ist beispielsweise nicht wirklich ersichtlich, warum die Fähigkeit der ‚fluiden Intelligenz/Reasoning‘ der emotionalen ‚O-Ebene‘ und nicht der kognitiven ‚C-Ebene‘ zugeordnet wurde, die bei Gibson nur auf Speicher- und Abrufvorgänge reduziert ist (Gibson, 2009b, S.123). Die Fähigkeiten nach der CHC Theorie und den fünf OCEAN-Persönlichkeitsmerkmalen sind als bipolares Kontinuum konzipiert (Gibson, 2009b, S.125).

Im folgenden sind die einzelnen Ebenen, deren Beschreibungen, Akronyme und die Zuordnungen zur CHC Theorie angegeben:

⁸Das Akronym steht für Analysis, Design, Development, Implementation, and Evaluation

⁹Das Akronym steht für (1) Analyze learner characteristics, (2) State objectives, (3) Select, modify, or design materials, (4) utilize materials, (5) require learner response und (6) Evaluation

¹⁰Das Akronym steht für Attention, Relevance, Confidence, Satisfaction

(a) ‚C-Ebene‘, ausgehend vom Begriff *cognitive* (Deale & Pastore, 2014, S.210 leiten die ‚C-Ebene‘ von Cattle-Horn-Carrol ab). Sie beschreibt nach Gibson den Vorgang des Speicherns und Abrufens von Wissen. Die C-Ebene ist folgenden Fähigkeiten nach der CHC Theory of Intelligence zugeordnet (Gibson, 2009b, S.124; Gibson, 2014a, S.154ff.):

- (1) kristallisierte Intelligenz
- (2) langfristige Gedächtnisspeicher und Aufruf derselben
- (3) generelles domänenspezifisches Wissen
- (4) quantitatives Wissen
- (5) Lesen/Schreiben
- (6) psychomotorische Fähigkeiten

(b) ‚O-Ebene‘, ausgehend vom OCEAN- und OCC-Modell(Ortony et al., 1988; Gibson, 2009b, S.127). Diese Ebene modelliert – auch dies ist kritisierbar – gleichzeitig zwei Vorgänge. Zum einen beschreibt sie das emotionale SuS-Profil, beziehend auf das bekannte Big Five Modell der Persönlichkeitspsychologie, auch OCEAN-Modell (McCrae & Costa Jr., 1996; Howard & Howard, 2000) genannt. ‚O‘ steht hierbei für *openness*, ‚C‘ für *conscientiousness*, ‚E‘ für *extraversion*, ‚A‘ für *agreeableness* und ‚N‘ für *neuroticism* (Gibson, 2009b, S.123ff. Gibson, 2011, S.3f. Christensen et al., 2011, S.204; Gibson, 2014a, S.150). Für eine detaillierte Beschreibung aller OCEAN-Faktoren und Zusammenstellung empirischer Evidenz sei an Hopper (2014, S.28ff.) verwiesen. Außerdem beschreibt die ‚O-Ebene‘ das Verarbeiten von Informationen.

Das OCC-Modell, nach der Monographie ‚The cognitive structure of emotions‘ von Andrew Ortony, Gerald Clore and Alan Collins (Ortony et al., 1988) beschreibt den Einfluss von Emotionen auf Verhalten (Gibson, 2009b, S.129).

Die O-Ebene ist folgenden Fähigkeiten nach der CHC *theory of intelligence* zugeordnet (Gibson, 2009b, S.124; Gibson, 2014a, S.154ff.):

- (1) Fluide Intelligenz/Reasoning
- (2) Kognitive Verarbeitungsgeschwindigkeit
- (3) Entscheidungs- oder Reaktionszeit
- (4) Psychomotorische Geschwindigkeit
- (5) Kurzzeitgedächtnis

(c) ‚V-Ebene‘, ausgehend vom Begriff *visual-auditory-kinesthetic perception* / an anderer Stelle viszerale Gehirn(ebene) (Gibson, 2011, S.2). Diese Ebene modelliert die prä-emotionale Wahrnehmungsebene, also die sensorischen und motorischen neuronalen Netzwerke (Gibson, 2011, S.2). Die V-Ebene ist folgenden Fähigkeiten nach der CHC *theory of intelligence* zugeordnet (Gibson, 2009b, S.124; Gibson, 2014a, S.154ff.):

- (1) Visuell/Räumliche Fähigkeiten
- (2) Auditive Verarbeitung
- (3) Kinästhetische Fähigkeiten
- (4) Olfaktorische Fähigkeiten
- (5) Taktile Fähigkeiten

(d) ‚E-Ebene‘ (für Environment). Der ‚E-Ebene‘ sind keine Fähigkeiten nach der CHC *theory of intelligence* zugeordnet (Gibson, 2009b, S.124), sondern sie wird vielmehr aus einer anderen Theorie abgeleitet, nämlich der *interpersonal circumplex theory* (Plutchik, 1997; Hopper, 2018, S.83; Gibson, 2009b, S.143; Tyler-Wood et al., 2015, S.19). Diese besteht aus 16 emotionalen Faktoren einer zweidimensionalen Skala der Konstrukte ‚*power*‘ und ‚*affiliation*‘. Sie bestimmt, wie ein Kommunikationsreiz emotional aufgeladen werden kann. Die Konstrukte ‚*power*‘ und ‚*affiliation*‘ befinden sich auf einem bipolaren Kontinuum, nämlich (1) dominant zu (2) unterwürfig, (3) sicher zu (4) unsicher, (5) exhibitionistisch zu (6) gehemmt, (7) umgänglich zu (8) reserviert, (9) freundlich zu (10) feindlich, (11) warm zu (12) kalt, (13) vertrauend zu (14) misstrauisch, (15) ehrerbietig zu (16) kompetitiv (Zibit & Gibson, 2005, S.12; Gibson, 2009b, S.143f.). Nach der *interpersonal circumplex theory* führen Interaktionen in der *power*-Dimension zu einer gegenteiligen Reaktion (Reziprozität), während Interaktionen auf der *affiliation*-Dimension zu einer gleichartigen Reaktion führen (Korrespondenz) (Gibson, 2009b, S.143; Plutchik, 1997). So führt beispielsweise ein dominanter Interaktionsreiz zu einer unterwürfigen Reaktion, während ein freundlicher Interaktionsreiz auch zu einer freundlichen Reaktion führt. Obwohl die Literatur keine expliziten Angaben zur Implementation im Simulationsmodell macht, ist anzunehmen, dass Reziprozität und Korrespondenz in ähnlicher Weise in Lehrkraft-SuS-Interaktionen (in simSchool ‚Kommentare‘ genannt) verarbeitet werden.

Die E-Ebene steht außerdem nach Gibson (2014a, S.154) für sämtliche Nutzerinneninteraktionen mit den SuS-Agenten.

Dass SuS Aufgaben nur dann effektiv lösen können, wenn sie sich im kognitiv individuell passenden Bereich befinden, wird theoretisch aus der Zone der proximalen Entwicklung (Vygotskij, 1987b; Vygotskij, 1987a) hergeleitet (Gibson, 2011, S.7; Gibson, 2014a, S.150f. Gibson et al., 2013, S.1317).

4.16.5 Algorithmische Umsetzung – Simulationsmodell

Zunächst ist festzustellen, dass COVE ein Konzeptionsmodell und kein Simulationsmodell ist. Die wirkliche Umsetzung der algorithmischen Strukturen im Simulationsmodell von simSchool weicht davon ab. Gibson begründet dies folgendermaßen:

COVE assumes that a modeler can swap in and out as many factors as needed for a specific simulation purpose, interacting with a core of processes taking place in the layers (Gibson, 2009b, S.125).

Außerdem weist Gibson darauf hin, dass mittelstarke bis starke Korrelationen zwischen den CHC Fähigkeiten festgestellt wurden, was für eine Simplifizierung durch Vereinheitlichung dieser Faktoren spreche (Gibson, 2009b, S.125). So bestätigen Christensen et al. (2011, S.204): „academic components are currently represented by a single variable representing overall academic performance“ (Vgl. auch Knezek et al., 2015, S.150).

Im Simulationsmodell von simSchool sind nach einigen Quellen allen SuS neun Variablen zugewiesen, nämlich die fünf Faktoren des OCEAN-Modells, die visuelle, auditive und kinästhetische Wahrnehmung (VAK), sowie eine Variable für die akademische Leistungsfähigkeit (Hettler et al., 2008, S.23f. Christensen et al., 2011, S.204; Hopper, 2018, S.82; McPherson et al., 2011, S.98). Gibson spricht hingegen von zehn Variablen, wobei er zusätzlich noch *language capability* nennt (Gibson, 2014a, S.151; Gibson et al., 2013, S.1317). In anderen Studien sowie in den Screenshots und Videos zu simSchool finden sich Hinweise, dass eine *language capability* implementiert worden ist (beispielsweise Hopper, 2014, S.24), wie genau diese Variable mit den anderen Variablen im Simulationsmodell interagiert, bleibt leider unklar. Gibson und Jakl (2013) sprechen auch von zehn den SuS zugewiesenen Variablen, allerdings handelt es sich um zum Teil unterschiedliche Variablen: Genannt werden die fünf OCEAN-Faktoren, die aus diesen Faktoren aggregierten Variablen ‚*power*‘ und ‚*affiliation*‘, eine Variable zur akademischen Leistungsfähigkeit, eine als ‚*circumplex*‘ bezeichnete Variable und schließlich eine Variable zur Bestimmung der Körperposition der SuS am Tisch (Gibson & Jakl, 2013, S.4). Die Variable ‚*power*‘ wird berechnet aus dem Durchschnitt von *extraversion*, *conscientiousness* und *openness*, die ‚*affiliation*‘-Variable (auch

als *happiness* bezeichnet) basiert auf dem Durchschnittswert von *agreeableness* und *neuroticism* (Gibson & Jakl, 2013, S.4). Die *circumplex*-Variable kennzeichnet den emotionalen Status und basiert auf den sechzehn Faktoren des *Interpersonal Circumplex Models*. Bis auf den ungeklärten Status der *language capability* Variable lassen sich diese Unterschiede folgendermaßen erklären: (1) Gibson und Jakl beschreiben das vereinfachte Simulationsmodell ohne VAK-Variablen, (2) aggregierte Variablen wie ‚*power*‘, ‚*affiliation*‘ und ‚*circumplex*‘ werden in vielen Quellen nicht als Variablen klassifiziert, (3) die Variable zur Speicherung der Pose hat einen eher datenbankbezogenen Charakter und wurde deswegen bei Gibson und Jakl nicht miteinbezogen. Aufgrund der Komplexität der Rekonstruktion des Simulationsmodells kann man künftigen Beschreibungen von digitalen Simulationen nur empfehlen wirklich alle Variablen, die in der Datenbank vorkommen, zu erwähnen und eine strikte Trennung von Simulations- und Konzeptionsmodell einzuhalten.

Die VAK-Variablen sind direkt aus der CHC *theory of intelligence* übernommen, wobei auf die Modellierung der olfaktorischen und taktilen Fähigkeiten verständlicherweise verzichtet wird, da diese keinen großen Anteil an der Modellierung einer Klassenraumumgebung haben (Gibson, 2009b, S.128).

Alle OCEAN-Faktoren besitzen einen Wert von -1 (nicht vorhanden/extrem schwache Ausprägung) bis 1 (extrem starke Ausprägung) (Hettler et al., 2008, S.24; Gibson, 2009b, S.126; Hopper, 2018, S.82). Die Variablen werden in einem Intervall von 0.1 festgelegt, was bedeutet, dass eine Variable 21 verschiedene Ausprägungen haben kann (Gibson, 2009b, S.126; In Hettler et al., 2008, wird von 20 verschiedenen Ausprägungen berichtet). Diese Variablen sind für die Simulationspartizipantinnen nicht einsehbar (Christensen et al., 2011, S.207; Hopper, 2018, S.82, 90), aber sie können sich aus der Lektüre einer in Textform präsentierten SuS-Akte erschlossen werden. Die Partizipantinnen erreichen diese Funktionalität von simSchool über einen Mausklick auf den Laptop, der am Lehrerpult positioniert ist (Deale & Pastore, 2014, S.206). Außerdem werden die Veränderungen der aggregierten Konstrukte ‚*power*‘, ‚*happiness*‘ und ‚*academic*‘ bei der Auswahl einer Schülerin graphisch dargestellt. Für jede Variable der SuS gibt es fünf Beschreibungen, jeweils in den Clustern von -1, -0.5, 0, +0.5 und +1 (Christensen et al., 2011, S.207).

Verwirrenderweise berücksichtigt Gibson (2009a, S.126) in seiner Rechnung, wie viele unterschiedliche SuS damit denkbar sind, nur die OCEAN- und nicht die VAK-Variablen¹¹. So errechnet er einen Wert von $21^5 = \text{ca. } 4$ Millionen unterschiedlichen SuS und $5^5 = 3125$ verschiedenen SuS-Beschreibungen. An anderer Stelle berücksichtigen er und seine Kolleginnen die Va-

¹¹In Gibson (2014a, S.156) spricht er hingegen nicht von unterschiedlichen SuS sondern von ‚OCEAN-States‘, was die treffendere und weniger verwirrende Formulierung ist.

riable für akademische Leistungsfähigkeit, ignorieren aber die VAK-Variablen und setzen für die Anzahl der verschiedenen Ausprägungen einen Wert von 20 an, was zum Ergebnis von $20^6 = \text{ca. } 64$ Millionen verschiedener SuS und 15.625 verschiedenen Beschreibungen führt (Hettler et al., 2008, S.6). Berücksichtigt man alle 9 Variablen, kommt man zu dem astronomischen Wert von $21^9 = 749$ Milliarden verschiedener SuS und $5^9 = 1.9$ Millionen verschiedenen Beschreibungen. Christensen et al. (2011, S.7) und Gibson (2014a, S.151) geben ähnliche Werte an, wobei beide mit 20 Ausprägungsstufen rechnen und Gibson 10 Variablen angibt; bei der Anzahl der Beschreibungen sind sich Christensen et al. (2011, S.8), Deale und Pastore (2014, S.210) und Hopper (2018, S.90) mit einem angegebenen Wert von $5^9 = 1.9$ Millionen einig.

Problematisch an allen Berechnungen ist aber, dass an anderer Stelle (Deale & Pastore, 2014, S.210; Gibson, 2009b, S.129) angegeben wird, dass die VAK-Variablen nicht von -1 bis 1 gespeichert werden, sondern einen Bereich von 0 bis 1, mit jeweils 10 Zwischenstufen, abdecken. Berücksichtigt man dies errechnet sich ein Wert von $21^6 + 3^{10} = \text{ca. } 85.8$ Millionen verschiedener SuS. Letztendlich ist die Zahl der unterschiedlichen SuS zwar nicht von großer Bedeutung, da in einer Intervallskalierung auch theoretisch unendlich viele individuelle SuS erzeugt werden könnten, aber die genaue Betrachtung der Angaben in der Literatur ist deswegen sinnvoll, weil die unterschiedlichen Angaben so verwirrend sind, dass es schwer ist die Anzahl und Ausprägungsstufen der im Simulationsmodell verwendeten Variablen nachzuvollziehen.

Die unterschiedlichen Angaben werden durch einen Screenshot von simSchool in Hettler et al. (2008, S.2) fast vollständig erklärt: Aus diesem geht hervor, dass simSchool in dieser Version über zwei Simulationsmodelle verfügte, wobei eines davon die VAK-Variablen ausklammerte („*Simplified simulation*“ und „*Include visual, auditory, and kinesthetic variables*“). Offen hingegen sind die widersprüchlichen Angaben zu den Ausprägungsstufen und die Frage, welche Rolle der zehnte von Gibson (2014a, S.151) genannte Faktor *language capability* im Simulationsmodell spielt. Dieser Faktor wird schlicht nicht thematisiert, obwohl ihn Gibson (2014a, S.152) in die Kalkulation verschiedener SuS-Agenten aufnimmt.

In simSchool gibt es vorgefertigte SuS, deren Schülerinnenprofile (VAK-OCEAN-A) auf Aufzeichnungen praktizierender Lehrkräfte basieren (Badiee & Kaufman, 2015, S.3; Kruse & Gibson, 2011). Daneben gibt es die Möglichkeit SuS als Simulationsnutzerin selbstständig zu erstellen (Hopper, 2018, S.90).

Die Simulationspartizipantinnen können während des Simulationsverlaufs zwei grundsätzliche Handlungen durchführen: Zum einen können Aufgaben gestellt werden („assign tasks“), zum anderen kann verbal mit einzelnen Schülerinnen kommuniziert werden („comment“) (Hettler et al., 2008, S.27; Deale

& Pastore, 2014, S.205).

4.16.6 Aufgaben (Tasks)

Aufgaben können sowohl an eine einzelne Schülerin, als auch an die gesamte Klasse gestellt werden (Hettler et al., 2008, S.4). Werden Aufgaben gestellt, so enthält jede Aufgabe exakt die neun Variablen, die zur Modellierung der SuS verwendet werden, nämlich die sensorischen VAK-Variablen, die OCEAN-Variablen und eine Variable zur akademischen Leistung (Hettler et al., 2008, S.25). Akademische Leistung ist in diesem Einsatzszenario besser als Aufgabenschwierigkeit zu interpretieren. Anschließend vergleicht der Algorithmus von simSchool die neun Variablen aller SuS mit denen der Aufgabe (Gibson, 2014a, S.152). Übersteigt eine Variable der Aufgabe eine Variable einer Schülerin, so hat die Schülerin Probleme mit dieser Komponente (siehe Abb. 7, S. 116). Die genaue Verarbeitung der Variablen bleibt hierbei unklar. Insbesondere wird nicht thematisiert, wie sich das SuS-Verhalten (z.B. Pose, Beteiligung) durch Abgleich von Aufgabe und SuS-Variablen ändert. Gibson (2014a, S.155) spricht davon, dass die Verarbeitung zumindest im Konzeptionsmodell gestaffelt passiert, was durchaus einleuchtend ist: Erst nimmt die Schülerin die Aufgabe sensorisch wahr (VAK), dann finden die präkognitiven emotionalen Verarbeitungen statt (OCEAN) und abschließend folgt die bewusste kognitive Verarbeitung (A).

Die akademische Variable und die OCEAN-Variablen der SuS ändern sich während der Simulation konstant. Dies ist unter anderem den präsentierten Screenshots zur Feedbackkomponente zu entnehmen (Gibson & Jakl, 2013, S.2). In Hettler et al. (2008, S.25) wird erwähnt, dass simSchool in der Basisversion insgesamt 20 vorgefertigte Aufgaben enthält, wie z.B. „do a brief presentation from memory“. Obwohl in der selben Forschungsarbeit erwähnt wird, dass ein Task „durch dieselben VAK-OCEAN-A Variablen charakterisiert wird [Übers. des Autors]“ (Hettler et al., 2008, S.25), zeigt ein weiterer Graph andere (abgekürzte) Variablen (Hettler et al., 2008, S.6, S.25) (Abschnitt 4.16.6, S. 117). Neben den bekannten *extraversion*-, *agreeableness*- und *academic*-Variablen sind dort weitere Variablen zu finden: *persistence*, *emotion[al stability]* und *intellect[ual openness]*. Wie aus einem Screenshot zur Erstellung von individuellen SuS-Avataren in Hettler et al. (2008, S.40) (ebenfalls in Christensen et al., 2011, S.208) ersichtlich wird, wurden drei Faktoren des OCEAN-Modells umbenannt bzw. geändert. Die Konstrukte *persistence* und *emotional stability* sind nach dem Verständnis der simSchool-Autorinnen annähernd äquivalent zu *conscientiousness* und *neuroticism* zu verstehen (Christensen et al., 2011, S.208). Auch *intellectual openness* enthält einen Teilbereich des Konstrukts *openness*. Es wird nicht transparent

gemacht, warum hier eine Änderung der Benennung erfolgte.

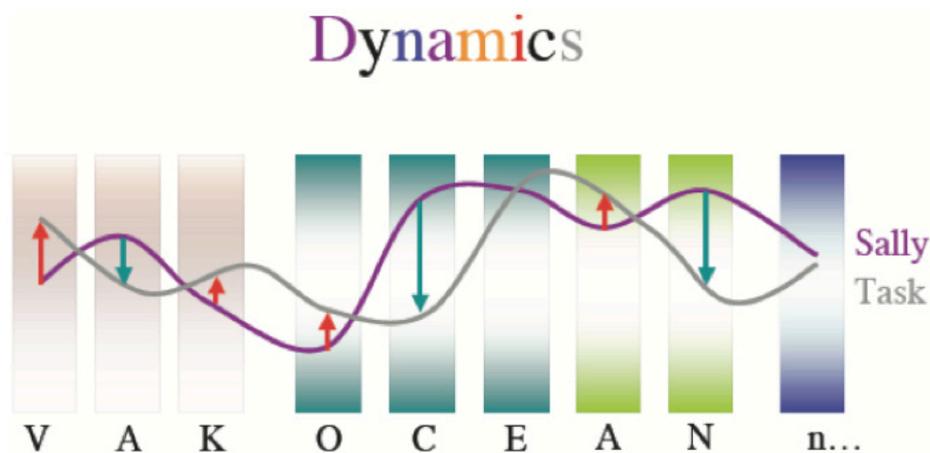


Abbildung 7: Aufgaben-SuS-Wechselwirkung in simSchool, Quelle: Hettler et al. (2008, S.25)

Nachdem Hopper (2018, S.83) und Gibson (2014a, S.160) in neueren Arbeiten aber wieder die originalen OCEAN-Faktoren (dargestellt als Kontinuum-Tabelle / Feedbackscreenshot) benutzen, während im Jahr 2013 (Gibson & Jakl, 2013, S.2) die adaptierten Variablen verwendet wurden, ist zu vermuten, dass um das Jahr 2014 eine Umbenennung zu den originalen OCEAN-Faktoren erfolgte. Es ist zu vermuten, dass die Autorinnen von einer Synonymie der alten und neuen Bezeichnungen ausgehen, da Knezek et al. (2015, S.150) wieder von den adaptierten Variablen berichten.

Abgesehen von den vorgefertigten Tasks haben die Simulationsnutzerinnen die Möglichkeit eigene Tasks zu implementieren. Im Jahre 2012 gab es insgesamt 110 solcher user-generierten Tasks, welche von 62 Nutzerinnen erstellt worden sind (Monsour, 2012).

Aufgaben sind in vier Kategorien geordnet, nämlich (1) *recall*, (2) *skill/concept*, (3) *thinking* und (4) *extended thinking*, was eine Adaption bzw. Vereinfachung der Bloom'schen Aufgabentaxonomie (Bloom et al., 1976) ist (Hopper, 2014, S.113). Dieses Simulationsmodell bildet allerdings nur die kognitive Komponente von Aufgaben ab. Wissensdimensionen, wie sie Appel (2015, S.74ff.) verwendet, werden nicht modelliert.

Die Konzeption dieses Teils des Simulationsmodells ist nebulös und in den offengelegten Teilen problematisch.

Zunächst sind die OCEAN-Faktoren unmissverständlich *traits*, d.h. personenbezogen. Ausprägungsänderungen sind nur sehr partiell möglich und

wenn, dann in einem sehr langen Zeitraum. Aus den Feedbackscreenshots sowie aus den extrahierten Datenbanken zur Useranalyse (Gibson & Jakl, 2013, S.4) geht klar hervor, dass sich die OCEAN-Werte der SuS während einer Unterrichtsstunde innerhalb weniger Minuten massiv ändern (Christensen et al., 2011, S.206; Zibit & Gibson, 2005, S.11; Gibson, 2009a, S.138; Hopper, 2018, S.97, 100, 103), was sich eigentlich nicht mit der Konzeption des OCEAN-Modells vereinbaren lässt. In einem Forschungsartikel thematisiert Gibson dies sogar:

The OCEAN model is typically used as a ‚trait‘ theory, but here it is used as scope for the dimensionality of a large but finite number of ‚states‘ of preconscious processes (e.g. affects and emotions if made conscious), so additional explanation is needed (Gibson, 2014a, S.150).

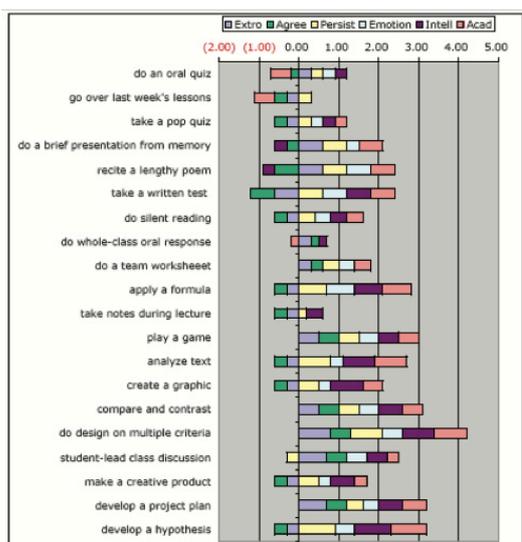


Abbildung 8: OCEAN-Änderungen nach Aufgaben in simSchool, Quelle: Hettler et al. (2008, S.25)

Gibsons Erklärung soll hier kurz skizziert werden. Angenommen, eine an sich offene Person (‚O‘ als *trait* im OCEAN-Modell) wird gebeten die Nationalhymne vor einem großen Publikum zu singen. Dies erfordert psychologische und mentale Ressourcen, welche ein Ungleichgewicht zwischen Anforderung der Aufgabe und dem persönlichen *trait* erzeugen. Nachdem die Aufgabe abgeschlossen ist, näherte sich der *openness*-Faktor wieder dem Equilibrium an. Ohne dies explizit zu thematisieren sind hier Anknüpfungspunkte zur Ego-

Depletion Theory (Baumeister et al., 1998) zu erkennen, welche nach neuesten Meta-Studien kritisiert wurde (Sripada et al., 2014).

Zudem ist es nicht leicht zu begründen, dass einer Aufgabe inhärent alle Komponenten des OCEAN-Modells innewohnen. Es wäre noch halbwegs plausibel zu begründen, dass die Präsentation eines Textes vor der Klasse für eine introvertierte Person schwerer als für eine extravertierte Person durchzuführen ist. Demzufolge könnte man dieser Aufgabe einen höheren Extraversionsanteil zusprechen. Aber wie kann eine Aufgabe auf Neurotizismus wirken? Welchen Extraversionsanteil hat eine einfache Additionsaufgabe? Ist der Aufgabe „apply a formula“ wirklich eine negative Verträglichkeit inhärent (Wie auf [Abschnitt 4.16.6](#), S. 117 zu erkennen ist), weil man während dieser Aufgabe nicht mit anderen Personen zusammenarbeitet? Geht man davon aus, dass jeder Aufgabe auch die VAK-Variablen zugewiesen werden (siehe [Abb. 7](#), S. 116), so ist unklar, was die Intervalle dieser Skala inhaltlich aussagen. Während es ersichtlich ist, dass eine blinde Schülerin (V Wert von 0) Probleme damit haben wird einen Tafelanschrieb ohne simultanes Vorlesen niederzuschreiben, so stellt sich die Frage, welchen V-Wert man dieser Aufgabe nun zuweisen soll. Von 0.1 bis 1.0 erscheint fast alles möglich, so dass dieser V-Wert sehr willkürlich erscheint.

4.16.7 Kommentare

Kommentare werden, da sie für die Autorinnen von simSchool die externen Faktoren im Sinne des COVE-Modells darstellen, theoretisch aus der *interpersonal circumplex theory* (Plutchik, 1997) abgeleitet. Jeder Kommentar wird mit einer der sechzehn emotionalen Komponenten versehen, indem diese von den Simulationsnutzerinnen aus einer zweidimensionalen, kreisförmigen Userschnittstelle ausgewählt werden. Die zwei Dimensionen sind ‚power‘ (von dominant zu unterwürfig und ‚affiliation‘ (von feindlich zu freundlich) (Hopper, 2018, S.92f.). Da an anderer Stelle beschrieben wurde, dass sich diese Variablen in simSchool aus der Aggregation von OCEAN-Faktoren ergeben, ist zu vermuten, dass Kommentare Einfluss auf diese Variablen nehmen. Dies wird aber in der Literatur nicht expliziert.

Interessanterweise wird ein großer Anteil an Kommunikation im Klassenraum, nämlich das Stellen von Aufgaben, nicht mit einer solchen Komponente versehen. Außerdem stellt sich bei diesem Verfahren die Frage, ob der Simulator dann noch zielführend als Assessment-Tool eingesetzt werden kann. Es ist wohl kaum anzunehmen, dass eine Simulationspartizipantin – wenn es darum geht ihre Leistung zu messen – freiwillig ihren Kommentar mit einer ‚feindlichen‘, ‚unterwürfigen‘ oder ‚exhibitionistischen‘ Komponente versehen würde. Im Rahmen eines Experimentes ist dies sicher möglich, aber nicht im

Sinne eines Assessments. Eine Möglichkeit dieses Problem zu umgehen wäre es, Kommentare in narrativer Form bereitzustellen, die zwar eine emotionale Komponente beinhalten, diese aber nicht für die Simulationspartizipantinnen einsehbar sind. Dies ist allerdings mit großem Konstruktions- und Programmierungsaufwand verbunden und sicher nicht leicht zu implementieren.

Kommentare sind in zwei Kategorien geordnet, zum einen Verhaltenskommentare (*behavioral comments*), zum anderen akademische Kommentare (*academic comments*) (Hopper, 2018, S.93; Christensen et al., 2011, S.207). Jede dieser Kategorien enthält drei Unterkategorien, nämlich Fragen, Beobachtungen und Feststellungen (*assertions*).

simSchool verfügt über 32 vordefinierte Kommentare, die acht unterschiedlichen Farben zugeordnet sind, die sich auf das zweidimensionale *interpersonal circumplex model* beziehen (Hopper, 2018, S.93f.). Damit werden acht Zwischenstufen, die im Modell noch existieren, in der Umsetzung des Simulationsmodells nicht berücksichtigt. Die Kommentare sind kontextunspezifisch und kurz, z.B. „You just earned two points towards the pizza party“, „Go to the office“ oder „You need to recall those details“ (Hopper, 2018, S.93).

4.16.8 Anwendung, Zielgruppe und wahrgenommene Fidelität

Zielgruppe für simSchool sind sowohl Lehramtsstudierende, als auch angehende Lehrkräfte, wobei in den Studien schwerpunktmäßig Lehramtsstudierende betrachtet werden (Gibson, 2014a, S.149; Hopper et al., 2013; Collum et al., 2019, S.735; Deale & Pastore, 2014, S.204). Ziel der Simulation ist es zum einen Teil Tools zum Performance-Assessment bereitzustellen, wobei ‚*knowledge-in-action*‘ gemessen wird (Gibson, 2013; Gibson, 2014a; Gibson et al., 2013, S.1316). Zum anderen sollen handlungsnahen Kompetenzen (auch als Know-How, Heuristik, prozedurales Wissen, stilles Wissen bezeichnet) bzw. erfolgreiche Lehrstrategien geschult werden (Gibson, 2012a, S.1f, 4; Hopper, 2018, S.79; Gibson, 2009a, S.135; Tyler-Wood et al., 2015, S.17; Hopper, 2014, S.27). Außerdem sollen Lehrkräfte einen besseren Blick für die Bedürfnisse der SuS bekommen und ihre Instruktionen diesen Bedürfnissen anpassen (Gibson & Halverson, 2004, S.2).

In der Studie von Christensen et al. (2011, S.213) mit N=157 Teilnehmerinnen, welche genau wie diese Arbeit die wahrgenommene Fidelität des Simulationsmodells untersuchte, wurde die Passgenauigkeit zwischen modellierten und realen SuS als „moderately well“ bezeichnet.

72.8% der N=22 Teilnehmerinnen der Untersuchungen von Badiie und Kaufman (2015, S.5) stuften die simulierten Konversationen zwischen SuS und Lehrkräften als „unrealistisch“ oder „sehr unrealistisch“ ein. Zudem wur-

den mangelnde, unangemessene oder unrealistische Handlungsalternativen beklagt, sowie die graphische Fidelität der Simulation kritisiert (Badiie & Kaufman, 2015, S.7). Weiterhin wünschten sich die Teilnehmerinnen eine größere Adaptivität der Simulation bezüglich der Simulationsnutzerinnen (Badiie & Kaufman, 2015, S.5f.). Einschränkend sei erwähnt, dass sich die Kritik primär auf die Interaktionsmöglichkeiten bezieht; immerhin 50% der Partizipantinnen beurteilten die Charakteristika der virtuellen SuS als „realistisch“ oder „sehr realistisch“. 68.1% der Teilnehmerinnen beurteilten die Herausforderungen, mit denen sie im Simulator konfrontiert wurden, als „realistisch“ oder „sehr realistisch“ (Badiie & Kaufman, 2015, S.6). Insgesamt in eine ähnliche Richtung verweist die Studie von Stavroulia et al. (2015a, S.211), in der 58.3% der N=63 Studierenden der School of Pedagogical and Technological Education in einem Fragebogen angaben, dass sie simSchool für keinen sehr realistischen Klassenraum hielten. Allerdings bezeichnete kein Studierender simSchool als „überhaupt nicht realistisch“. Letztlich gaben 75% der Studierenden an, dass simSchool zur Schulung von Lehrkraftinteraktionen und *classroom management* benutzt werden kann; 83.3% wünschten sich die Aufnahme der Software ins universitäre Curriculum.

Die Dauer der Simulation ist je nach Studie unterschiedlich, bei Deale und Pastore (2014, S.209) wurden drei Trainingssessions à 10 Minuten von N=13 Pädagogikstudierenden durchgeführt, bei Hopper et al. (2013) verbrachten die N=58 Lehramtsstudierenden nur 7.5 Minuten im Simulator. In der Dissertationsschrift von Hopper (2014, S.39) (Auch in Hopper, 2018, S.80 beschrieben) werden 2x3 Stunden Sessions mit N=37 *learning technology* Studierenden erwähnt, wobei mehrere simSchool Szenarien genutzt wurden, für die jeweils 30 Minuten vorgesehen waren. Die dreistündigen Sessions beinhalten auch Anteile, die nicht die Nutzung des Simulators erfordern, wie Einführung und Reflexion. Gibson (2014a, S.151) nennt Simulationszeiten von 10 bis 30 Minuten. In der Studie von McPherson et al. (2011, S.98) verbrachten N=151 Sonderpädagogikstudierende 7.5 Minuten in der Simulation. Knezek et al. (2015, S.152) erwähnen zwei Experimentalstudien. In der ersten Studie nutzen N=31 Lehramtsstudierende den Simulator acht Stunden, verteilt über drei Trainingssessions. In der zweiten Studie wurden N=36 Lehramtsstudierende für 2x3 Stunden in der Simulation geschult. Ebenso wie in Hoppers Studien bestanden diese Sessions aus verschiedenen Szenarien und zeigen nicht nur reine Simulatornutzung an. Alle Studien verwendeten auch Kontrollgruppen.

Die simulierte Zeit der fiktiven Unterrichtsstunde vergeht in simSchool um den Faktor 6 schneller als in der Realzeit (Hopper et al., 2013, S.2872; Gibson, 2014b; Hettler et al., 2008, S.4; Gibson, 2012b).

Neben der Nutzung des digitalen Simulators gehört auch immer ein BAR

und AAR mit Feedbackrunde zur Simulation. Als BAR wird das *before-action review* bezeichnet, als AAR das *after-action review*. Diese Komponenten dienen den Simulationspartizipantinnen dazu, sich auf die Simulation vorzubereiten (BAR), sowie gelernte Fähigkeiten zu festigen und zu reflektieren (AAR) (Gibson, 2009a, S.694). Grundlage der Besprechung bietet dabei die automatisierte Feedbackkomponente von simSchool, welche den Verlauf von SuS-Variablen und die Interaktionen der Simulationspartizipantinnen über die Zeit visualisiert.

4.17 simSchool 2.0

Interaktionsmodellierung:	Typ 4 (Komplexe algorithmische Steuerung ohne Entscheidungsbaum)
Immersionsform:	2D graphical simulation
Granularität:	Lesson simulation
Noch in Benutzung:	Ja
Vorgänger:	simSchool 1.0
Nachfolger:	–

simSchool 2.0 wurde im Herbst 2018 (Collum et al., 2019, S.735f.) auf der Webseite <http://simschool.org> veröffentlicht, Informationsvideos zu den gemachten Änderungen folgten im März 2019 (simSchool Educator Training, 2019a). Der Simulator ersetzt das ältere simSchool 1.0, welches seit der Veröffentlichung nicht mehr verfügbar ist. Über simSchool 2.0 gibt es zum Zeitpunkt des Verfassens dieser Arbeit leider kaum Informationen. Der Artikel von Collum et al. (2019) erwähnt die Entwicklung von simSchool 2.0, fasst aber nur Ergebnisse aus simSchool 1.0 zusammen und erläutert nicht das geänderte Simulationsmodell. Ein Artikel von Demoulin et al. (2012) trägt zwar „SimSchool 2.0“ [sic!] im Titel, beschreibt aber lediglich ein Proposal zur Entwicklung einer neuen Simulation in der Lehrkräftebildung, wobei sich die Autorinnen auf simSchool 1.0 beziehen: „Our design is based on an existing product (<http://www.simschool.org/>) and it is in this sense that we titled our work SimSchool 2.0, which is an update, or rather an improvement consequent commodity.“ (Demoulin et al., 2012, S.15). Diese Simulation wurde als *Second Life*-Anwendung konzipiert und schlug als Interaktionsmodellierung *virtual puppetry* vor. Daher kann ausgeschlossen werden, dass sich der Artikel von Demoulin et al. auf das hier beschriebene simSchool 2.0 bezieht. Die wichtigsten Informationsquellen sind mangels Alternativen die Schulungsvideos für simSchool 2.0 und die Nutzung der Versuchsversion. Diese Videos

waren im Schreibprozess dieser Arbeit im Jahre 2019 noch verfügbar, sind aber mittlerweile offline. Da es sich um Videoinhalte handelt wurden die Informationen leider nicht archiviert.

Es ist möglich eine 14-tätige Versuchsversion von simSchool 2.0 zu beziehen, und es gibt einige von simSchool produzierte Videos, die das Programm aus verschiedenen Perspektiven beleuchten. Der Grafikstil von simSchool 2.0 ist nach wie vor im zweidimensionalen Comic-Stil gehalten, allerdings werden diesmal Vektorgrafiken eingesetzt. Diese Vektorgrafiken ermöglichen simple, recht statisch wirkende Animationen. In simSchool 2.0 sind nun auch Grundschülerinnen modellierbar und grafisch ausgestaltet (simSchool Educator Training, 2019b). Eine spezifische Version mit dem Namen SimRwanda wird präsentiert, in welcher ruandische Schulen simuliert werden (simSchool Educator Training, 2019d).

Ebenso wie im Vorgänger ist das Verhalten der SuS rein reaktiv. Interagieren die Simulationspartizipantinnen nicht mit den SuS, führt dies zu keiner Verhaltensänderung. Während sich das Simulationsmodell von SimSchool 1.0 zu SimSchool 2.0 geändert hat, sind die theoretischen Grundlagen bzw. das Konzeptionsmodell identisch geblieben (Collum et al., 2019, S.734).

In simSchool 2.0 ist es wie in der Vorgängerversion möglich Aufgaben an SuS zu verteilen und Kommentare zu geben. Darüber hinaus können nun auch sogenannte ‚Strategien‘ durchgeführt werden.

Es gibt zwei Kategorien von Strategien: Zum einen zehn *classroom management*-bezogene, wie (a) Begrüßung der SuS , (b) Geben non-verbaler Signale oder (c) Ermahnung zur Aufmerksamkeit. Zum anderen existieren vier versorgungsbezogene Strategien (*accommodation*), wie das Austeilen von Kopfhörern oder die Benutzung eines Timers, um die Arbeitszeit einer Aufgabe für alle SuS sichtbar zu machen.

Die Kommentare sind in drei Kategorien geordnet und damit ähnlich aufgebaut wie in simSchool 1.0: Es gibt (1) Fragen, (2) Beobachtungen und (3) Anweisungen (*directive*). Jede Kategorie enthält drei vordefinierte, wiederum recht abstrakte Aussagen wie „Can you tell me more?“ oder „You are working hard“. Außerdem können in einem Textfeld eigene Kommentare eingegeben werden. Die emotionale Ausprägung der Aussage kann von den Simulationspartizipantinnen über das zweidimensionale *interpersonal circumplex model* (Plutchik, 1997) in Form eines Kreises festgelegt werden.

Instruktionen werden nun nach der Taxonomie von Anderson et al. (2000) in sechs kognitiven Dimensionen eingeordnet, nämlich (1) *remember*, (2) *understand*, (3) *apply*, (4) *analyze*, (5) *evaluate* und (6) *create*. Die eigene Adaptierung der Bloom’schen Taxonomie aus dem vorherigen simSchool 1.0 wird damit fallengelassen (Collum et al., 2019, S.734). In jeder Kategorie der kognitiven Dimension befinden sich konkrete, ausformulierte Aufgaben-

aufträge, welche die Erstellerin der Simulation zuvor einer kognitiven Dimension zugewiesen hat. Außerdem wurde jeder Aufgabe eine akademische Schwierigkeit und eine ‚soziale Voraussetzung‘ zugewiesen, welche die Simulationspartizipantinnen über zwei Slider sehen, aber nicht verändern können. Ebenso unveränderbar sind benötigte Arbeitsmaterialien und die erwartete Blickrichtung der SuS.

Sowohl Instruktionen, Strategien als auch Kommentare können an eine spezifische Schülerin, an eine Gruppe oder an die gesamte Klasse gestellt werden.

Während der Simulation kann nun direkt über den Progress-Button beobachtet werden, welche SuS hochperformant sind, welche im Mittelfeld liegen und welche mit der gestellten Aufgabe Probleme haben.

Der Profile-Button erfüllt dieselbe Funktion wie der Lehrkraftlaptop im vorherigen simSchool 1.0: Durch einen Klick werden die Schülerinnenprofile angezeigt. Aus diesen Profilen wird ersichtlich, dass sich das Simulationsmodell geändert hat. Zunächst werden in der aktuellsten Version allen SuS 32 *traits* zugewiesen, die in 6 Kategorien geordnet sind. Jeder *trait* ist ein bipolares Kontinuum von -1 bis 1 (Collum et al., 2019, S.734), welches für die Nutzerinnen als Slider dargestellt wird. Die 32 *traits* sind in [Tabelle 3](#) (S. 123) aufgeführt.

Tabelle 3: *Traits* in simSchool 2.0

Executive Functions	Comprehension-Knowledge	Academic Ability
Short Term Memory	Academic Language Skill	General Academic
Long Term Memory	Conversational Fluency	Quantitative Knowledge (Math)
Fluid Reasoning	Discrete Language Skill	Reading and Writing (ELA)
Processing Speed		Science
Decision-Reaction Time		
Cognitive Processing	Social and Emotional	Physical
Psychomotor Ability	Curiosity	Sight
Psychomotor Speed	Excitable	Hearing
Visual Processing	Self-Discipline	Mobility
Auditory Processing	Motivation	General Health
Kinesthetic Processing	Seeks Achievement	
Tactile (Haptic) Processing	Outgoing	
	Compliant	
	Sympathetic	
	Vulnerable to Stress	
	Confident	

Unklar ist, wie diese *traits* miteinander interagieren, um das Verhalten der SuS zu steuern. Wie in simSchool 1.0 ist der Begriff ‚*traits*‘ an sich falsch gewählt, da sich alle Faktoren, bis auf die physischen, im Verlauf der Schulstun-

de ändern (simSchool Educator Training, 2019c). In der aktuellsten Demo-Version von simSchool finden sich die erwähnten 32 *traits*, im Schulungsvideo sind allerdings nur 12 *traits* zu sehen.

Drei der *traits* sind kognitiver Natur, fünf bilden die originalen OCEAN-Faktoren ab. Die letzten vier Faktoren sind physischer Art, nämlich die VAK Variablen plus eine ‚*general health*‘-Variable (simSchool Educator Training, 2019c).

Es ist geplant, die Nutzerinnengruppe von simSchool 2.0 auf Schulleiterinnen, Beraterinnen, pädagogisches Fach- und Unterstützungspersonal zu erweitern (Collum et al., 2019, S.737).

4.18 Smifta / SimInClass

Interaktionsmodellierung:	Typ 4 (Komplexe algorithmische Steuerung ohne Entscheidungsbaum)
Immersionsform:	3D graphical simulation
Granularität:	Lesson simulation
Noch in Benutzung:	Ja
Vorgänger:	–
Nachfolger:	–

Bei Smifta (in der englischsprachigen Version SimInClass) handelt es sich um eine 2016 veröffentlichte türkische dreidimensionale Unterrichtssimulation, die tendenziell dem Bereich der *serious games* zuzuordnen ist. Diese Zuordnung ist neben dem generellen Aufbau des Programms auch dadurch zu belegen, dass Smifta mehrere Preise für das „Best Serious Game“ (z.B. den Kristal Pıksel Indie Prize 2017; Finalist der Serious Games Showcase & Challenge 2018) gewonnen hat.

Smifta wurde im Rahmen des von der türkischen Anstalt für Wissenschaftliche und Technologische Forschung (TÜBİTAK) geförderten Projektes „FATİH“ in Kooperation mit der Technischen Universität des Nahen Ostens in Ankara (ODTÜ), der Siirt Universität und der Firma Simsoft Bilgisayar Teknolojileri Ltd. Sti. entwickelt (Yıldırım, 2016; o.A., 2019). Die Leitung der akademischen Struktur hatte Prof. Dr. Veysi İşler inne, der bis 2016 Präsident des MODSIMMER (Modeling and Simulation Research and Application Center) war und im Fachbereich der technischen Informatik/Computergrafik arbeitet. Die Firma Simsoft war primär für die technisch/graphische Umsetzung zuständig und hat ihren Hauptsitz in Ankara.

Ziel der Simulation ist es, korrektes *classroom management* bei Störungsverhalten zu schulen. Zudem soll die Integration von Technik in den Unterricht eingeübt werden (Serious Games Showcase and Challenge, 2018). Zielgruppe für Simfta sind sowohl Lehramtsstudierende als auch Lehrkräfte (Serious Games Showcase and Challenge, 2018). Laut der Simfta Webseite (o.A., 2019) gibt es 920 Nutzerinnen der Simulation und 34 – ausschließlich türkische – kooperierende Universitäten.

Simfta benutzt als Engine Unity und unterstützt Windows und Android (Serious Games Showcase and Challenge, 2018). Man merkt der Simulation ihre Herkunft aus der technischen Informatik/Computergrafik an – die grafische Ausgestaltung wirkt sehr professionell, sämtliche Figureninteraktionen sind animiert und die Simulationspartizipantinnen können sich frei durch den Raum bewegen und in jede Richtung blicken. Während die normale Simulatorbenutzung über Desktop-PCs geschieht – was auch auf vielen Videos und in Promomaterial zu beobachten ist – findet sich auf der Webseite von Simfta (o.A., 2019) der Hinweis, dass z.T. auch CAVE-Umgebungen (Cave Automatic Virtual Environment) zur Simulation verwendet werden. Es ist nicht klar, ob dies Zukunftspläne sind oder ob diese VR-Simulationen bereits durchgeführt wurden.

Die Ausgestaltung des Simulationsmodells basiert auf dem BDI-Modell (*belief, desire, intention*) (Köknar, 2015). SuS werden durch die neunstufige Skala der fünf OCEAN-Faktoren (siehe simSchool 1.0, Abschnitt 4.16, S. 105) beschrieben, besitzen eine Biographie in Textform und verfügen über drei von 0 bis 100 reichende situative Variablen für Konzentration, Vergnügen und Wissen. Diese Variablen werden durch folgende Simulationshandlungen beeinflusst: (1) Die Ausgestaltung des Stundenplans, (2) das Verhalten der Simulationspartizipantinnen auf Störungsverhalten der SuS, (3) der Umgang mit Technologie im Klassenraum, (4) das Aufrechterhalten des Blickkontakts zwischen Lehrkraft und SuS (Serious Games Showcase and Challenge, 2018).

Im Gegensatz zu simSchool sind die OCEAN-Faktoren tatsächlich *traits* und verändern sich nicht während der Simulation. Nur die situativen Variablen ändern sich. Die SuS-Variablen können sich die Partizipantinnen, ähnlich wie in simSchool, über virtuelle Schülerinnenakten erschließen.

Beim Start des Simulators kann eines von sechzehn Level ausgewählt werden, wobei Level zum Teil auch gesperrt sein können. In diesem Fall müssen sie nacheinander abgearbeitet werden müssen. Die Level sind fest in den Simulator encodiert und können von externen Personen nicht erweitert werden. Nachdem ein Level ausgewählt wurde, startet die Simulation des Unterrichtsgeschehens noch nicht direkt, sondern die Partizipantinnen entwickeln zunächst einen Stundenplan aus verschiedenen vorgefertigten Methoden-Bausteinen, wie z.B. ‚Hausaufgabenkontrolle‘ oder ‚Einführung durch lebenswelt-

lichen Bezug'. Ein vorgefertigtes Textfeld beschreibt, wie sich die Autorinnen des Simulators eine verbale Umsetzung der Unterrichtskomponente vorstellen (Simsoft Computer Technologies, 2018). Die in einer anderen Simulations-Version (SGSChallenge, 2018) ausgeführten Änderungen dieses Bildschirms sind rein kosmetischer Natur. Die auswählbaren Aktivitäten beziehen sich auf das jeweils ausgewählte Level. Jede Aktivität ist zeitlich begrenzt; die Simulationspartizipantinnen können die Zeit mittels Slider festlegen. Eine Stunde in Simfta ist immer in drei Teile aufgeteilt, die alle mit Aktivitäten gefüllt werden müssen: (1) Einführung, (2) Lehr-/Lernprozess und (3) Messung und Evaluation. Während dieser Vorplanung der Stunde kann eine von vier Sitzordnungen ausgewählt und die Anzahl der SuS bestimmt werden (Simsoft Computer Technologies, 2018).

Startet die Simulation des Unterrichtsgeschehens, so ist am unteren Bildschirmbereich eine Zeitachse mit den ausgewählten Aktivitäten der Unterrichtsplanung zu sehen, welche durch Ausgrauen anzeigt, wie viel Zeit vergangen ist und welche Aktivität gerade stattfindet. Während der Aktivitäten wird eine Sprechblase der Lehrkraft angezeigt. Dort findet sich die bereits aus der Stundenplanung bekannte Verbalisierung der Lehrkraft-Interaktionen (Sinifta, 2018).

Es gibt drei Buttons, mit denen man technische Geräte wie Tablets bedienen kann. Drückt man auf den Tablet-Button, so kann man unter anderem über die grafische Darstellung eines solchen Tablets Bilder an die SuS senden (Sinifta, 2018).

Wenn SuS Störungen initiieren, wird dies sowohl durch Animation, als auch durch ein kleines Icon über dem Kopf des SuS-Avatars signalisiert. Die Simulationspartizipantinnen können mit verschiedenen Maßnahmen darauf reagieren, z.B. „Frage stellen“, „Die Eltern kontaktieren“, „Privatgespräch suchen“ oder „Aus der Klasse schicken“.

Die Simulation ist jederzeit anhaltbar, vor- und zurückspulbar und die Simulationsgeschwindigkeit steuerbar. Den Simulationspartizipantinnen werden in der oberen linken Ecke des Simulatorfensters aggregierte Durchschnittswerte der situativen Variablen präsentiert. Am Ende der Simulation erfolgt ein Feedback in Form von (a) einer Bewertung, ausgehend von den Durchschnittswerten der situativen Variablen und (b) einem Graphen, welcher die Entwicklung der situativen Variablen aller SuS über die Zeit visualisiert (Simsoft Computer Technologies, 2018).

Simfta wurde bislang noch nicht international rezipiert.

4.19 STAR Simulator / TeachME / TeachLivE (TLE) / SimLab / Mursion

Interaktionsmodellierung:	Typ 7 (digital/virtual puppetry) / Typ 6 (human-in-the-loop)
Immersionsform:	3D graphical simulation
Granularität:	Lesson simulation / Task simulation
Noch in Benutzung:	Ja
Vorgänger:	siehe Text
Nachfolger:	siehe Text

Die Begriffe STAR Simulator / TeachME / TeachLivE (TLE) / SimLab und Mursion bezeichnen alle unterschiedliche Versionen desselben Simulators. Zunächst als STAR Simulator (Dieker et al., 2007) konzipiert, wurde der Begriff um das Jahr 2008 in TeachME (Teaching in a Mixed Reality Environment) umbenannt (Dieker et al., 2008; weitere Erwähnung dieser Bezeichnung in Andreasen & Haciomeroglu, 2009). Die Simulation wurde in Kooperation mit der Habermann Educational Foundation entwickelt, während die Programmierarbeiten von Simiosys, LLC durchgeführt wurden (Dieker et al., 2008). Dr. Charles Hughes war für den Großteil der technischen Umsetzung verantwortlich (Dieker et al., 2015; Nagendran et al., 2013).

Die Bezeichnung TeachME wird noch 2011 verwendet (Aronin, 2011), aber im Jahr 2012–2013 erfolgt eine Umbenennung in TLE TeachLivE (Hayes et al., 2013). Dies geht mit der Etablierung einer nationalen TeachLivE Konferenz (2013–2016) in den USA einher, deren Proceedings alle diesen Begriff nutzen. In neueren Veröffentlichungen ab dem Jahr 2017 wird die Simulation erneut umbenannt, was aus Dawson und Lignugaris/Kraft (2017) hervorgeht: „TLE was the name of the technology at the time this study was conducted. It is now called Mursion“ (Dawson & Lignugaris/Kraft, 2017, S.267). Die Bezeichnung TeachLivE wird ebenfalls im Titel der Publikation geführt. Mursion ist ein Startup-Unternehmen, das das Framework für die Simulation vertreibt. Es existieren spezifische Anpassungen, die unter anderen Namen wie „SimLab“ an verschiedenen Universitäten eingesetzt werden (Ledger & Fischetti, 2020; Carmody, 2019), wobei die Entwicklung von SimLab an der Murdoch University in Australien realisiert wurde (o.A., 2023).

Da die meisten Publikationen sich auf die TeachLivE-Version beziehen, wird im Folgenden dieser Begriff verwendet, um die Funktionsweise der Simulation zu illustrieren. Bei Besonderheiten in der Funktionsweise werden die jeweiligen Versionen genannt. Da TeachLivE einen virtual puppetry/human-

in-the-loop Ansatz gewählt hat, ist die Simulation trotz der hohen Zahl an Publikationen nur von bedingtem Interesse für diese Arbeit. Eine umfassende Betrachtung des Literaturfundus' soll deswegen an dieser Stelle nicht stattfinden. Der Grund dafür ist, dass in der gewählten Interaktionsmodellierung das Simulationsmodell nur implizit ist und daher nicht zur Analyse herangezogen werden kann. TeachLivE wird als *mixed reality* Simulation bezeichnet. Simulationspartizipantinnen interagieren mit SuS über einen großen Monitor (Dawson & Lignugaris/Kraft, 2017, S.29; Aronin, 2011), wobei sie mit den SuS kommunizieren als stünden sie vor einer realen Klasse. Der Raum, in dem sich der Monitor befindet, ist z.T. mit Requisiten wie Whiteboards, Schultischen und -stühlen ausgestattet (Dieker, Rodriguez et al., 2014; Chini et al., 2016). Ein Kinect-Sensor bestimmt die Position der Partizipantinnen im Raum und passt den Zoomgrad und den Winkel der Darstellung an (Dawson & Lignugaris/Kraft, 2017; Barmaki & Hughes, 2015). Die SuS werden als dreidimensionale Figuren von einer sogenannten Interaktantin gesteuert. Dies wurde in einigen Anwendungen über *motion capturing* realisiert (Barmaki & Hughes, 2015), in anderen, insbesondere in der neusten Version Mursion, ist die Steuerung über Gamepads realisiert (Dies ist auf den abgebildeten Fotos in Carmody, 2019, zu erkennen; o.A., 2023) und konstituiert damit nicht mehr *virtual puppetry*, sondern vielmehr eine human-in-the-loop Interaktionsmodellierung. In beiden Anwendungen übernimmt die Interaktantin die Rolle von SuS und interagiert verbal über ein Mikrofon mit den Simulationspartizipantinnen (Dieker et al., 2008, S.7; Dawson & Lignugaris/Kraft, 2017). Den Interaktantinnen werden Skripts aushändigt, die spezifizieren, wie die SuS zu spielen sind. Die Skripts basieren unter anderem auf Beschreibungen und detaillierten fiktiven Biographien von 12–14 jährigen middle school SuS (Dieker et al., 2015; Dieker et al., 2008), wobei später andere SuS und Personengruppen erschlossen wurden (Chini et al., 2016, nennt high school SuS; Kelley & Wenzel, 2018, nennt Eltern). Die herangezogene theoretische Grundlage zur Erstellung dieser Skripts variiert nach der konkreten Anwendung (Dieker et al., 2015, nennen *student archetypes* nach William Long als Ausgangsbasis; Dieker et al., 2008, nennen die Beschreibung von typischen SuS nach der American Academy of Child and Adolescent Psychiatry). TeachLivE simuliert vier bis fünf SuS, wobei die Interaktantin alle SuS simultan steuert (Dieker et al., 2008, S.11; Dieker et al., 2015; Barmaki & Hughes, 2015; Dieker, Rodriguez et al., 2014; Andreasen & Haciomeroglu, 2009; Hayes, 2015; Hayes et al., 2013). Nicht-Unterrichtssituationen wie Elternabende (mit Eltern-Agenten) wurden ebenfalls in der Simulation umgesetzt (Kelley & Wenzel, 2018; Dieker, Rodriguez et al., 2014). Es wird berichtet, dass Pläne existierten, die Zahl der Interaktantinnen zu erhöhen (Dieker, Rodriguez et al., 2014), aber in der vorliegenden Literatur sind keine Hinweise darauf zu finden, dass

dies umgesetzt wurde. Bei der Steuerung der SuS können automatische Animationen eingesetzt werden (Chini et al., 2016, S.4). Die Verwendung der automatischen Animationen wird über Slider realisiert, welche den Grad der Unruhe im Klassenraum abbilden. Dieser Wirkmechanismus wird in der Literatur manchmal auch als KI bezeichnet, wobei intransparent ist, wie das Simulationsmodell konkret aufgebaut ist (Dieker et al., 2015). Die Simulationsdauer beträgt 5 min (Dieker et al., 2008; Aronin, 2011; Chini et al., 2016), 2x7min (Barmaki & Hughes, 2015), „weniger als 10 min pro Session im Labor“ (Dieker, Rodriguez et al., 2014) oder 10–20 min (Andreasen & Haciomeroglu, 2009). Die Feedback-Komponenten unterscheiden sich nach konkreter Anwendung und beinhalten immer ein verpflichtendes AAR (after action review), das z.T. mit automatisiertem Feedback angereichert ist (Dieker et al., 2015; Barmaki & Hughes, 2015). Bei diesem automatisierten Feedback handelt es sich um Proximität der Lehrkraft zu den SuS und Reaktionszeiten von Lehrkraft und SuS (Barmaki & Hughes, 2015).

Thematisch ist die Simulation offen, es finden sich in der Literatur jedoch viele Anwendungen aus der Sonderpädagogik. Oft wird der Umgang mit Autismus behandelt (Dawson & Lignugaris/Kraft, 2017; Dieker et al., 2015; Garland et al., 2015; Garland et al., 2012), aber auch andere Wissensgebiete wie die generelle Vermittlung handlungsnaher Kompetenzen (oft *classroom management*) werden geschult (Dawson, 2016; Dieker et al., 2015; Andreasen & Haciomeroglu, 2009). TeachLivE nutzten im Jahr 2015 geschätzt 10.000 Partizipantinnen (Dieker et al., 2015), Barmaki und Hughes (2015) nennen die Zahl von 12.000 Partizipantinnen an 55 Universitäten.

4.20 The Tigerlake Project

Interaktionsmodellierung:	Typ 6 (Human-in-the-loop)
Immersionsform:	(Hyper-)Text-based simulation
Granularität:	Teacher simulation
Noch in Benutzung:	Nein
Vorgänger:	Inner-City Simulation Laboratory (Fischler, 2006, S.58)
Nachfolger:	SimTeach (Fischler, 2006, S.58)

Das Tigerlake Projekt ist ein rein textbasierter Simulator, der 1996 von Robert B. Fischler an der Indiana University Bloomington entwickelt wurde. Er wurde vor dem Hintergrund der konstruktivistischen *learner-centered pedagogy* entwickelt (Fischler, 2006, S.59) und ist von den Möglichkeiten der Frühphase des Internets geprägt. Es gibt keine direkten Interaktionsmöglichkeiten, der Simulator besteht primär aus statischen Webseiten, mit der Ausnahme eines sogenannten ‚Lehrkräfterraums‘, der ein Webforum zum Austausch zwischen Simulationsnutzerinnen bezeichnet. In den statischen Webseiten werden in Form von Textvignetten Schulsituationen beschrieben, wozu auch Nicht-Unterrichtssituationen wie die Bewerbung an einer Schule zählen¹². Anschließend wurden offene Fragen zu diesen Situationen gestellt, die sich auf das Lehrkraft Handeln beziehen. Die Fragen wurden von Lehramtsstudierenden beantwortet und via E-Mail oder postalisch an eine Kursleiterin übertragen, die dann individuelles Feedback dazu gab (Fischler, 2006, S.61f.). Die Situationen enthielten zwölf SuS, modelliert wurden die Jahrgangsstufen 1, 4, 8 und 11. Didaktisch fokussierte die Simulation laut Fischler (2006, S.61) auf „domänenspezifisches, prozedurales und metakognitives Wissen,.. Nach Egbert et al. (2000) wurde der Simulator im Rahmen der Spracherziehung eingesetzt.

¹²Einige der Situationsbeschreibungen sind auf der archivierten Webseite <http://web.archive.org/web/19991001201925/http://www.indiana.edu:80/%7Etiger/> einzusehen.

4.21 SimTeacher

Interaktionsmodellierung:	Typ 1 (Deterministischer Entscheidungsbaum)
Immersionsform:	(Hyper-)Text-based Simulation
Granularität:	Teacher simulation
Noch in Benutzung:	Nein, Benutzungsende ca. Ende 2005 (Abschaltung der Webseite)
Vorgänger:	Tigerlake Projekt
Nachfolger:	–

SimTeacher ist der Nachfolger des Tigerlake Projects und wurde von Robert B. Fischler entwickelt. Technisch umgesetzt wurde das Projekt als Webanwendung via PHP und MySQL (Fischler, 2006, S.75; Fischler, 2007). Für den Simulator selbst wurden nur Texte und Bilder genutzt, keine Animationen, Videos oder Sound (Fischler, 2006, S.150). Zielgruppe für die Simulation waren Lehramtsstudierende und Lehrkräfte im Vorbereitungsdienst (Fischler, 2006, S.77). Eine wichtige Neuerung gegenüber dem Tigerlake Project ist die Möglichkeit für InstruktorInnen selbst Simulationsszenarien zu entwickeln (Fischler, 2006, S.76); diese sind nicht unveränderlich mit dem Quelltext des Programms verknüpft (Fischler, 2007, S.44).

Die Interaktionsmodellierung ist ein deterministischer Entscheidungsbaum: Die SimulationspartizipantInnen wählen eine Handlungsalternative, begründen diese und werden dann auf eine weitere, im Vorfeld festgelegte, Webseite mit dem Ergebnis der Handlung weitergeleitet (Fischler, 2006, S.80; Fischler, 2007, S.45). SimulationspartizipantInnen können darauf hinweisen, dass sie die Fortsetzung der Simulation für unplausibel halten (Fischler, 2007, S.45). Die Zeit, die für einen Durchlauf der Simulation benötigt wurde, differiert nach den gewählten Szenarios, ist im Durchschnitt aber ca. 10 Minuten lang (Fischler, 2006, S.117).

Unabhängig von der Simulationskomponente gehörte zum Lernkonzept von SimTeacher auch die Arbeit mit einem Portfolio (Fischler, 2006, S.88f.). Im Sommer 2005 waren mit 53 Lehrkräften die meisten NutzerInnen in der Historie der Simulation auf der Plattform registriert. Es existierten zu diesem Zeitpunkt 49 Kurse mit 20 Simulationsszenarien (Fischler, 2007, S.46).

SimTeacher wurde in der Literatur kaum rezipiert; eine Ausnahme bildet Collum (2015, S.45f).

4.22 Transition to Teaching Classroom Management

Interaktionsmodellierung:	Typ 4 (Komplexe algorithmische Steuerung ohne Entscheidungsbaum)
Immersionsform:	3D graphical simulation
Granularität:	Lesson simulation
Noch in Benutzung:	Nein
Vorgänger:	Verbal Abuse Simulation (Verknüpfung)
Nachfolger:	Verbal Abuse Simulation (Verknüpfung)

Bei Transition to Teaching Classroom Management handelt es sich um einen Simulator, der im Rahmen des Projekts Transition to Teaching, gefördert durch das U.S. Department of Education, entwickelt worden ist. Während des Projekts arbeitete das Digital Media Program der Drexel University mit der Drexel School of Education und anderen Abteilungen der Universität im sogenannten RePlay Lab zusammen, um den Simulator zu entwickeln (Diefenbach et al., 2019, S.1f.). Ein zweiter Simulator mit anderem Fokus wurde parallel entwickelt und wird unter ‚Verbal Abuse Simulation‘ (siehe [Abschnitt 4.24](#), S. 135) näher beschrieben.

Zielgruppe sind Quereinsteigerinnen in den Lehramtsberuf (Diefenbach et al., 2019, S.2). Thematisch behandelt die Simulation Mobbing, sowie verbale und physische Belastung von Lehrkräften. Lehrkräfte sollten diesen Simulator und die „Verbal Abuse Simulation“ auf einer angedachten Webseite benutzen können (Diefenbach et al., 2019, S.2). Da sich kein Hinweis dazu findet, dass diese Idee je umgesetzt wurde, ist damit zu rechnen, dass der Simulator nicht mehr in Benutzung ist.

Technisch wurde der Simulator über die Unity Game Engine realisiert, der grafische Stil wirkt heutzutage etwas veraltet, bemüht sich aber um eine naturalistische Darstellung. In einer vermutlich früheren Version haben die SuS-Agenten noch ein comicartiges Aussehen (Drexel University, 2010b). SuS werden in dem Simulator durch Algorithmen gesteuert und sind animiert (Drexel University, 2010a). Transition to Teaching Classroom Management ist eine der wenigen Simulationen, in welcher sich SuS auch gegenseitig beeinflussen und der Zustand nicht nur von den Aktionen der Simulationspartizipantinnen abhängig ist. Ähnlich wie in Simfta legen die Nutzerinnen vor Beginn der Stunde einen Stundenplan fest, wobei folgende Phasen zur Auswahl stehen: (1) Einführung, (2) thematischer Vortrag, (3) angeleitetes Üben, (4) eigenständiges Üben, (5) Abschluss/Zusammenfassung. Jede dieser Phasen

ist nochmal in verwendete Methoden unterteilt, beispielsweise Lehrkraftvortrag, Vorstellung eines Problems an der Tafel oder ein Quiz. In dem der Simulation vorgelagerten Optionsmenü ist festlegbar, um welchen Faktor die simulierte Zeit schneller abläuft als die reale, wie lange die einzelnen Unterrichtsphasen dauern sollen und welche von vier Sitzarrangements verwendet werden sollen (Drexel University, 2010a). Startet die Unterrichtssimulation zeigt ein Fortschrittsbalken an, in welcher Phase der Stunde und welcher Methode sich die Simulationspartizipantinnen aktuell befinden. Jeder Wechsel der Methode ändert die internen Variablen der SuS-Agenten und die Interaktionsmöglichkeiten der Simulationspartizipantinnen. Es gibt zwei generelle Möglichkeiten mit den SuS zu interagieren: Zum einen *class commands*, zum anderen *student commands*. Letzteres modelliert die Mikroebene der Interaktionen mit einzelnen SuS, ersteres die Makrointeraktion mit der gesamten Klasse. *Class commands* sind beispielsweise (1) loben, (2) Fragen stellen, (3) Aufmerksamkeit fordern oder (4) ermahnen. *Student commands* beinhalten dieselben Interaktionsmöglichkeiten wie *class commands*, verfügen zusätzlich aber über die Handlungsalternativen (5) Aufforderung, den Platz zu wechseln, (6) Nachsitzen lassen, (7) aus dem Klassenraum schicken und (8) zur Schulleitung schicken. Während der Simulation ist eine freie Bewegung im Klassenraum möglich (Drexel University, 2010a).

Da die Autorinnen davon sprechen, dass alle Interaktionen eine positive oder negative Verstärkung hervorrufen, (Diefenbach et al., 2019, S.3) darf man ein lernparadigmatisches Fundament im Behaviorismus annehmen, obwohl dies nicht explizit gemacht wird. Allen SuS sind archetypische Persönlichkeitsprofile zugewiesen wie „Klassenclown“, „aggressiv“ oder „akademisch“. Anhand eines Screenshots ist erkennbar (siehe [Abschnitt 8.3](#), S. 256), dass drei durch Balken dargestellte SuS-Variablen existieren. Was diese darstellen, was sie bestimmen und wie das Simulationsmodell konkret funktioniert, erläutern die Autorinnen leider nicht.

Der Simulator beinhaltet ein zu- und abschaltbares Tutorialsystem, das bei der Auswahl der Interaktionen hilft.

Nachdem die Unterrichtssimulation beendet ist, findet eine automatische Bewertung der Interaktionen statt. Es ist möglich eine Aufzeichnung der Simulation abzuspielen und an Entscheidungspunkten einzugreifen, um eine bessere Bewertung zu erhalten (Diefenbach et al., 2019, S.3).

Laut den Autorinnen war Transition to Teaching Classroom Management im Jahre 2010 noch in der Entwicklung (Diefenbach et al., 2019, S.3).

4.23 VCS – Virtual Classroom Simulator

Interaktionsmodellierung:	Typ 4 (Komplexe algorithmische Steuerung ohne Entscheidungsbaum)
Immersionsform:	2D graphical simulation (ohne Animation)
Granularität:	Task simulation
Noch in Benutzung:	Nein, Deregistrierung der EduPeX Homepage im Jahr 2013
Vorgänger:	–
Nachfolger:	–

Bei VCS handelt es sich um die Simulation von SuS-Reaktionen auf verschiedene Aufgabenstellungen. Der Simulator wurde von Simon Skrødal im Rahmen seiner Dissertation an der School of Education der University of Adelaide (Australien) entwickelt und programmiert, mit späteren Programmierhilfen von Paul Calder. Ein erster Prototyp – der sich allerdings erheblich von der finalen Version von VCS unterschied – wurde erstmals 2002 eingesetzt (Skrødal, 2010).

Es steht für die Simulation nicht im Vordergrund eine Schulstunde immersiv darzustellen. Stattdessen ist sie vielmehr als Experiment angelegt, bei der Simulationspartizipantinnen beobachten können, wie die simulierten SuS auf Änderungen von Aufgabenstellungen (siehe [Abschnitt 8.3](#), S. 256) reagieren. Eine Aufgabe wird in VCS durch zwei Variablen, ‚Schwierigkeit‘ und ‚Neuheitswert‘, definiert, welche die Simulationsnutzerinnen durch zwei Slider manipulieren können. Die simulierten SuS werden durch folgende Variablen modelliert: *mood*, *distractibility*, *self-efficacy*, *knowledge* und *metacognition* (Skrødal, 2010, S.5). Diese Variablen ändern sich mit den Eingaben der Simulationsnutzerinnen, jedoch nicht deterministisch (Skrødal, 2010, S.5). In den Worten von Skrødal: „It is highly unlikely that two identical runs would ever produce the exact same output“ (Skrødal, 2010, S.7). Der aktuelle *mood*-Status der SuS wird über 2D-Screenshots von vorgerenderten 3D-Kopf-Modellen dargestellt. Es gibt insgesamt sieben *mood*-Werte: *happy*, *content*, *normal*, *tired*, *sad* und *upset*. Nachdem eine Aufgabe gestellt wurde und sich der *mood*-Wert geändert hat, erhalten Nutzerinnen visuelles Feedback durch eine Änderung der Kopf-Grafik (ohne Animation).

Neben der Interaktionsmöglichkeit SuS Fragen zu stellen, können die Simulationsnutzerinnen einzelne SuS zu mehr Aufmerksamkeit ermahnen (Skrødal, 2010, S.62, 128). Eine Uhr zeigt die simulierte Zeit an und kann pausiert oder gestoppt werden (Skrødal, 2010, S.79).

Am Ende erhalten die Simulationsnutzerinnen von VCS Feedback in Form von Graphen, die anzeigen, wie sich das Verhalten der SuS über die Simulationsdauer geändert hat (Skrødal, 2010, S.6, 80).

VCS fand in der Forschungsrezeption keine große Beachtung. Das Programm selbst oder der Quellcode sind nicht mehr öffentlich verfügbar. Dennoch ist es wichtig VCS zu betrachten, da in Skrødals Dissertation die Entwicklung des Simulationsmodells explizit thematisiert wird.

4.24 Verbal Abuse Simulation

Interaktionsmodellierung:	Typ 1 (Deterministischer Entscheidungsbaum)
Immersionsform:	3D graphical simulation
Granularität:	Task simulation
Noch in Benutzung:	Nein
Vorgänger:	Transition to Teaching Classroom Management (Verknüpfung)
Nachfolger:	Transition to Teaching Classroom Management (Verknüpfung)

Die Verbal Abuse Simulation wurde im Zusammenhang mit dem Projekt Transition to Teaching (siehe [Abschnitt 4.22](#), S. 132) in einem Joint Venture zwischen der Drexel University und dem externen Dienstleister Vcom3D Inc. entwickelt. Der externe Dienstleister war für Animation und Charaktermodellierung zuständig, alles andere wurde vom RePlay Lab der Drexel University erstellt.

Im Gegensatz zu Transition to Teaching Classroom Management handelt es sich bei der Interaktionsmodellierung dieses Simulators um einen deterministischen Entscheidungsbaum, weswegen hierfür auch Audioaufzeichnungen der verbalen Äußerungen der SuS angefertigt werden konnten. Thematisch geht es darum, eine aggressive Störungssituation nicht weiter eskalieren zu lassen, sondern die Situation durch korrekte Handlungs- und Wortwahl zu beruhigen. Die benutzten Szenarien sind nach Erfahrungen von Lehrkräften und Schulleitungen modelliert. Leider beschreiben die Autorinnen diesen Simulator nur kurz ohne weitergehende Details zu nennen (Diefenbach et al., 2019, S.4).

4.25 Virtual School Class (VSC)

Interaktionsmodellierung:	Typ 2 (Randomisierter Entscheidungsbaum) / Typ 3 (Entscheidungsbaum mit komplexer algorithmischer Steuerung)
Immersionsform:	(Hyper-)Text-based simulation / Interactive video-based simulation (IVBT)
Granularität:	Task simulation
Noch in Benutzung:	Ja
Vorgänger:	Der Simulierte Klassenraum (SKR) / Simulated Classroom
Nachfolger:	U.a. Virtual Simulated Classroom (V.C.R.)

Virtual School Class (VSC) ist eine Weiterentwicklung des simulierten Klassenraums (SKR) und teilt die meisten Eigenschaften mit diesem Simulator.

Ein wichtiger Unterschied ist die Immersionsform, welche ein Hybrid aus einer (Hyper)-Text-based simulation und einer interaktiven videobasierten Simulation ist. Während Meldungen als statische Fotografien von zwei Zuständen (sich meldend, sich nicht meldend) dargestellt werden, triggert ein Klick auf eine sich meldende Person ein Video, das eine richtige oder falsche Antwort zeigt (Fiedler, 2012, S.5f.). Die dargestellten zwölf Personen sind alle im Alter von Studierenden und der Hintergrund zeigt einen Universitätshörsaal (Fiedler, 2012, S.25, 5f.). Warum diese Darstellungsform gewählt wurde und in welchem inhaltlichen Kontext diese spezielle Form des Simulators genutzt worden ist, geht leider aus der Literatur nicht hervor.

4.26 Virtual Simulated Classroom (V.C.R.)

Interaktionsmodellierung:	Typ 2 (Randomisierter Entscheidungsbaum)
Immersionsform:	(Hyper-)Text-based simulation / Interactive video-based simulation (IVBT)
Granularität:	Task simulation
Noch in Benutzung:	Ja
Vorgänger:	Der Simulierte Klassenraum (SKR) / Simulated Classroom
Nachfolger:	–

Die Simulation Virtual Simulated Classroom (V.C.R.) wurde im Rahmen des BMBF-geförderten „ProSim“ Projektes (01.03.2017-29.02.2020) der Universität Kiel unter der Leitung von Prof. Dr. Ute Harms und Prof. Dr. Jens Möller entwickelt und verbessert (<https://www.ipn.uni-kiel.de/de/forschung/projekte/prosim>). Es ist anzumerken, dass der ähnliche Begriff ‚Virtual Class Room (VCR)‘ ein anderes, nicht simulationsbezogenes, Projekt bezeichnet.

In der neuesten Fassung der Software scheint der simulierte Handlungsspielraum erweitert worden zu sein. Es wurden Anstrengungen unternommen, die Software auch als Trainingsinstrument in universitären Lehrveranstaltungen von angehenden Lehrkräften einzusetzen (Fischer et al., 2017). Zum Zeitpunkt des Verfassens dieser Arbeit existieren noch keine weitergehenden Publikationen, die den Simulator oder seine Nutzung näher beschreiben.

4.27 VR-Klassenraum

Interaktionsmodellierung:	Typ 6 Human-in-the-loop
Immersionsform:	VR-simulation
Granularität:	Task simulation
Noch in Benutzung:	Ja (in Entwicklung)
Vorgänger:	–
Nachfolger:	–

Der VR-Klassenraum ist ein neuer, zur Zeit noch in Entwicklung befindlicher, deutscher Simulator, der in Potsdam von Axel Wiepke (Komplexe Multimediale Anwendungstechniken) und Prisca Paulicke (Institut für Innovation und Technik, Digitalisierung der Bildung) entwickelt wird. Der Simulator

nutzt die HTC Vive als HMD und die Game Engine Unity zur Darstellung einer hochimmersiven Umgebung (Wiepke & Paulicke, 2018, S.4).

Thematisch beschäftigt sich der Simulator mit dem Umgang von Störungsverhalten und *classroom management* (Wiepke & Paulicke, 2018, S.1). Zielgruppe sind Lehramtsstudierende. Die Simulationspartizipantinnen sind dazu angehalten in der virtuellen Umgebung einen kurzen thematisch-didaktischen Vortrag zu halten (z.B. einen Unterrichtseinstieg im Fach Geschichte) (Wiepke & Paulicke, 2018, S.2). Während dieser Zeit steuert ein Coach das Verhalten der virtuellen SuS-Agenten, wobei diese mitarbeiten oder Unterrichtsstörungen initiieren können (Wiepke & Paulicke, 2018, S.2f., 5). Die möglichen Zustände der SuS basieren auf Kounins sieben SuS-Profilen: (1) Eindeutige und vollständige Mitarbeit, (2) wahrscheinliche Mitarbeit, (3) eindeutig keine Mitarbeit, (4) Unruhe, (5) Lustlosigkeit, (6) arbeitsbezogenes Fehlverhalten, (7) nicht-arbeitsbezogenes Fehlverhalten (Wiepke & Paulicke, 2018, S.3; Kounin, 1976). Die Simulationspartizipantinnen können sich während der Simulation frei im Raum bewegen, was für die VR-Umgebung einen Platz von mindestens 3x3 Metern erfordert. Zudem können virtuelle Bewegungen im Raum via Teleportation zu realisiert werden; eine mittlerweile gängige Methode in der VR-Entwicklung um Bewegungen zu ermöglichen. Mit dieser Technik kann bei empfindlichen Personen recht effektiv Übelkeit und Schwindelgefühl (Simulatorkrankheit) vermieden werden. Während der Simulation beobachten andere Teilnehmerinnen des Seminars die VR-Interaktionen über eine Leinwand. Abschließend werden die Erfahrungen und Beobachtungen gemeinsam im Plenum ausgewertet (Wiepke & Paulicke, 2018, S.5).

Gewünscht ist, dass der Umgang mit diesen Unterrichtsstörungen zum Abbau von Stereotypisierungen und dem Aufbau von situativen Handlungskompetenzen wie präventives Monitoring und adaptiven Interventionen führt. Es wird angestrebt noch andere Szenarien, wie die Ausgestaltung eines konkreten Lehr-/Lernszenarios, in den Simulator zu implementieren (Wiepke & Paulicke, 2018, S.3). Bisher wurde der Simulator von zwei Expertinnen aus den Bereichen Bildungswissenschaft und Schule, sowie von fünf Lehramtsstudierenden evaluiert. Dabei gaben alle Partizipantinnen ein hohes Präsenzepfinden zu Protokoll (Wiepke & Paulicke, 2018, S.5).

4.28 VR-Tool von Bailenson et al. (ohne Bezeichnung)

Interaktionsmodellierung:	Keine / Unklar
Immersionsform:	Virtual Reality (VR) Simulation
Granularität:	Task simulation
Noch in Benutzung:	Unklar
Vorgänger:	–
Nachfolger:	–

Ebenso wie beim VR-Tool von Stavroulia et al. handelt es sich bei diesem VR-Tool von Bailenson et al. (2008) nicht im engeren Sinne um einen Simulator. Die Applikation wird von den Autorinnen als (*immersive*) *virtual environment* bezeichnet. Entwickelt wurde das Tool von der Stanford University. Es nutzte ebenfalls HMD, genauer gesagt ein nVisor SX (2x1280x1024) oder eine Virtual Research HMD (2x640x480) zur Darstellung der VR-Inhalte (Bailenson et al., 2008, S.103).

Ziel des Einsatzes dieses VR-Tools war es, herauszufinden, ob (1) eine augmentierte Sicht in einem virtuellen Klassenraum einer Lehrkraft hilft, ihre Aufmerksamkeit gleichmäßiger auf alle SuS zu verteilen, (2) es beim Hineinversetzen in eine SuS-Rolle einen Vorteil bietet, direkt vor der Lehrkraft zu sitzen (3) das Lernen mit aufmerksamen oder nicht-aufmerksamen Co-Lernern in einer virtuellen Umgebung die Lernperformanz beeinflusst (Bailenson et al., 2008, S.111f.).

Im ersten Experiment versetzten sich die N=40 studentischen Simulationsteilnehmerinnen in die Rolle einer Lehrkraft, die einen achtminütigen Vortrag vor neun virtuellen SuS-Agenten hielt. Die SuS-Agenten spielten während der Simulation vorab aufgezeichnete Kopfbewegungen ab (Bailenson et al., 2008, S.112f., 115). Bei einer Gruppe verringerte sich die Opazität der 3D-Figuren der SuS, wenn die Simulationsteilnehmerinnen diese länger nicht mit einem Blick fixiert hatten, bei der anderen Gruppe geschah dies nicht (Bailenson et al., 2008, S.113).

4.29 VR-Tool von Yaacov J. Katz (ohne Bezeichnung)

Interaktionsmodellierung:	Unklar
Immersionsform:	Virtual Reality (VR) simulation
Granularität:	Lesson simulation
Noch in Benutzung:	Nein
Vorgänger:	–
Nachfolger:	–

Diese israelische VR-Simulation, die 1999 von Yaacov J. Katz entwickelt wurde, ist zur Schulung von Kindergartenlehrkräften entwickelt worden. Die Simulation basiert auf der Entwicklungsumgebung SuperScope 5.5 und wurde unter MS-DOS entwickelt. Die Simulationspartizipantinnen benutzten schon damals Head-Mounted Displays und einen Datenhandschuh (Katz, 1999, S.154). Damit dürfte diese Simulation eine der frühesten, wenn nicht die früheste, VR-Simulation in der Lehrkräftebildung sein. Es gibt noch einen zweiten Grund, warum diese Simulation erwähnenswert ist: Die Prämisse war nämlich nicht, wie in den meisten anderen Simulationen, eine Lehrkraft zu verkörpern, sondern im Gegenteil die Umgebung des Kindergartens aus der Perspektive eines Kindes wahrzunehmen. Supervisorinnen bewerteten die N= 37 Partizipantinnen anhand von 40 Items zum Thema ‚Verständnis von Emotionen, Bedürfnissen und Wahrnehmungen von Kindern‘. Die Simulations-Treatment-Gruppe schnitt hierbei besser ab als die Kontrollgruppe von N=35 Partizipantinnen, welche einen thematisch äquivalenten Workshop besuchten.

4.30 VR-Tool von Stavroulia et al. (ohne Bezeichnung)

Interaktionsmodellierung:	Keine / Unklar
Immersionsform:	Virtual Reality (VR) simulation
Granularität:	Task simulation
Noch in Benutzung:	Ja
Vorgänger:	–
Nachfolger:	–

Das namenlose griechische VR-Tool, dass um 2015 im Arbeitskreis vom Kalliopi Evangelia Stavroulia mit Unterstützung des in der Schweiz ansässigen interdisziplinären MIRALab, unter Leitung von Nadia Magnenat-Thalmann,

entwickelt wurde (Stavroulia, 2015), wird von den Forscherinnen korrekterweise nicht als Simulator, sondern als *virtual (reality) environment* bezeichnet. Die Applikation wurde im Rahmen des NOTRE Projekts entwickelt, welches durch das Horizon 2020 Programm der EU gefördert wurde (Stavroulia et al., 2018, S.282).

Zunächst wurde von den Forscherinnen simSchool verwendet (Stavroulia et al., 2015a), bis 2015 mit der Entwicklung eines eigenen Tools begonnen wurde (Stavroulia, 2015). Das VR-Tool wurde konzipiert, um es Partizipantinnen in bestimmten Szenarien zu ermöglichen die Perspektive verschiedener Personen einzunehmen (Stavroulia et al., 2019, S.196). Während eine vorgefertigte Szene abläuft, können Partizipantinnen lediglich die Szene sitzend beobachten, jedoch nicht mit der Umgebung interagieren (Stavroulia et al., 2018, S.281). Die Partizipantinnen müssen nach Angabe der Forscherinnen sitzen, da sie an empfindliche EEG-Geräte angeschlossen sind, während die Szenen laufen. Technische Umsetzung ist die Nutzung einer HTC Vive VR-Brille (Stavroulia et al., 2018, S.277); programmiert wurde die Applikation in der Game Engine Unity (Stavroulia, 2015, S.56; Stavroulia et al., 2017, S.3566). Zielgruppe sind Lehramtsstudierende in den ersten Semestern. Das Tool wurde auch in der Lehrkräfteweiterbildung eingesetzt (Stavroulia et al., 2017, S.3562f.).

Thematisch behandeln die Szenarien (a) Bullying/Rassistische Anfeindungen (Stavroulia et al., 2018, S.276; Stavroulia & Lanitis, 2018), (b) Hyperaktivität, (c) aggressives Störverhalten (Stavroulia, 2015, S.55; Stavroulia et al., 2018, S.3562) und (d) Umgang mit Drogenmissbrauch bei SuS (Stavroulia et al., 2019, S.202). Die simulierten SuS und die modellierten Situationen basieren auf Erfahrungen von Lehrkräften, die tatsächlich so stattgefunden haben. Die dargestellten SuS gehören der Sekundarstufe an (Stavroulia, 2015, S.55).

In einem geschilderten Szenario zum Drogenmissbrauch versetzen sich Partizipantinnen beispielweise (1) in eine Schülerin, die einen anderen Schüler beim Drogenkonsum beobachtet hat, (2) in die Lehrkraft, die hinzugerufen wurde oder (3) in den konsumierenden Schüler selbst, inklusive graphisch dargestellter Wahrnehmungsverzerrungen (Stavroulia et al., 2019, S.202). Davon wird sich erhofft, die Lehrkräfte zu mehr Empathie und angemessenem Umgang mit solchen Problemen anzuleiten. Die Studien analysieren zum einen emotionale Reaktionen der Partizipantinnen während der VR-Erfahrung (Stavroulia et al., 2015b; Stavroulia, 2016; Stavroulia et al., 2019), zum anderen erforschen sie den Zusammenhang zwischen *presence* und grafischer Ausgestaltung/Simulationsfidelität der Repräsentation (Stavroulia et al., 2018).

4.31 Andere Simulationen

Es existieren weitere Simulationen, die in der Lehrkräftebildung eingesetzt wurden. Diese sind an dieser Stelle nur kurz aufgeführt und nicht genauer analysiert, da sie keinen großen Mehrwert bezüglich einer Simulationsmodell-Entwicklung zeigen. Viele der Simulationen sind veraltet und werden nicht mehr verwendet:

- (a) Classroom Teacher Training 3D simulation (Kaufman & Ireland, [2016](#), S.264)
- (b) Connect.ed / Connect[ED] (Bradley & Kendall, [2014](#), S.8; Artist Made ProductionsPlus, [2011](#))
- (c) GAM-WATA (Chuang & Shih, [2014](#), S.344)
- (d) Sim: Classroom (Praetorius et al., [2016](#))
- (e) TD-Simulation (Strang, [1996](#), S.616; Strang et al., [1996](#), S.620; Sullivan et al., [1998](#))
- (f) Teacher Talk Game (Bradley & Kendall, [2014](#), S.7)
- (g) Teaching Simulator (Indonesien) (Fatimah et al., [2015](#); Fatimah et al., [2017](#))
- (h) The Curry Teacher Simulation (Strang, [1997](#))
- (i) The Teaching Simulator (Fatimah et al., [2015](#))
- (j) The Virtual School (o.A., [2002](#); Eason, [2002](#); Fischler, [2004](#))

5 Entwicklung des interaktionalen Simulationsmodells

5.1 Planung der Simulation Clasivir

Ziel dieses Kapitels ist es, zunächst mögliche Anwendungsgebiete der geplanten Simulation Clasivir (Classroom Simulator in Virtual Reality) zu entwickeln. Im Gegensatz zum Simulationsmodell, das im Rahmen dieser Arbeit entwickelt wurde, ist die Implementierung der konkreten Simulation noch nicht umgesetzt. Daher rekurriert dieses Unterkapitel auf einen geplanten Einsatz; es wird gezeigt, welche Umsetzungsideen die Ausgestaltung des Simulationsmodell beeinflusst haben. Ein Simulationsmodell muss stets an den Anforderungen eines geplanten didaktischen Einsatzes ausgerichtet sein. Um die Bedingungen für einen gelungenen Einsatz zu analysieren, empfiehlt sich die Verwendung des DOIT-Modells (=Didaktik, Organisation, Individuum, Technologie) von Horz und Ulrich (2013) (siehe Abb. 9, S. 143).

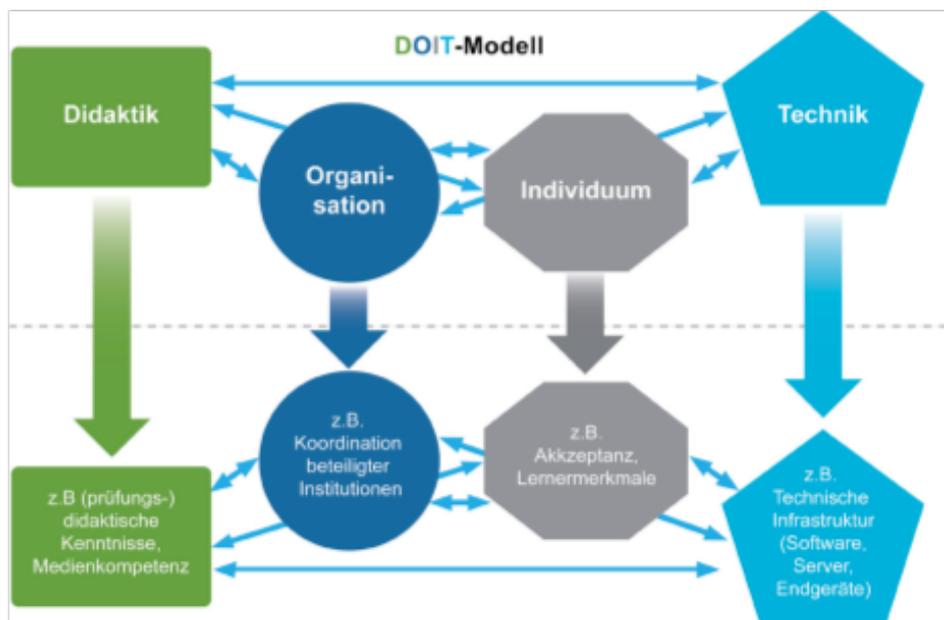


Abbildung 9: Das DOIT-Modell, Quelle: Horz und Ulrich (2013)

Das DOIT-Modell wurde zur Analyse innovativer digitalisierter Lehr-/Lernprozesse entwickelt, weswegen es einen wertvollen Beitrag zur Planung der VR-Simulation Clasivir liefert. Die erste Komponente des DOIT-Modells ist die Didaktik. Darunter werden sowohl Voraussetzungen der Medienkompetenz zur Benutzung von digitalen Angeboten, als auch deren didaktischen

Zielsetzungen verstanden. Unter Organisation werden alle organisationalen Verflechtungen beteiligter Institutionen aufgeschlüsselt, etwa Rechenzentren, Hochschuldidaktik und Prüfungsämter. Auf der Individualebene muss sich die digitale Lehr-/Lerninnovation um bestmögliche Akzeptanz bemühen und individuelle Bedürfnisse und Anforderungen berücksichtigen. Letztlich ist die verwendete Technik zu analysieren, die bedarfsgerecht und aktuell sein muss, sich jedoch immer den didaktischen Zielsetzungen unterzuordnen hat (Horz & Schulze-Vorberg, 2017).

Nach der Vorstellung der Simulation sollen theoretische Grundannahmen und Grenzen des Simulationsmodells transparent gemacht werden. Die Zielgruppe der Simulation *Clasivir* sind Lehramtsstudierende und potentielle Studieninteressierte. Diese Entscheidung berücksichtigt, dass vor allem Novizinnen von einer eingeschränkten Fidelität des Simulationsmodells profitieren, was sich aus den theoretischen Überlegungen zur Simulationsfidelität ergibt (siehe [Abschnitt 2.5.12](#), S. 38). Eine Simulationsumgebung wie *TeachLivE* oder *microteachings* wären demzufolge eher für Fortgeschrittene oder Expertinnen geeignet. Die erarbeiteten didaktischen Ziele legen eine Zielgruppe aus den frühen Fachsemestern nahe. Insbesondere die intendierte Theorie-Praxis-Verzahnung ist relevant für eine Zielgruppe, die über keine oder nur eingeschränkte eigene Lehrerfahrungen verfügt. Die Nutzung der Simulation durch Studieninteressierte erfolgt mit dem Ziel des (Self-)Assessments.

5.1.1 DOIT-Modell: Didaktik

Nach dem DOIT-Modell sollen zunächst die didaktischen Zielsetzungen einer digitalen Lerngelegenheit expliziert werden. Primäre didaktische Ziele der Simulationsnutzung sind:

- (a) Umgang mit Störungsverhalten, insbesondere aggressivem Störungsverhalten und Vermittlung handlungsnaher *classroom-management* Kompetenzen.

Rationale für dieses Ziel ist, dass physische und verbale aggressive Störungsverhalten in realen Hospitationen selten erfahrbar sind. Simulationen haben als Werkzeug den Vorteil, eine ethische, kontrollierte Umgebung (*controlled-exposure*) schaffen zu können, um Kompetenzen in Umgang mit solchen Störungsverhalten zu schulen. Weiterhin wünschen sich viele Lehramtsstudierende, dass aggressives Störungsverhalten thematisiert wird, was die Akzeptanz der Simulation auf Individualebene steigern könnte. Die meisten der analysierten Simulationen in der Lehrkräftebildung beschäftigen sich ebenfalls mit der Schulung von

classroom-management. Dies hat auch damit zu tun, dass derartige Verhalten technisch einfacher zu modellieren ist, als beispielsweise die Etablierung von Regeln und Ritualen.

(b) Schulung von (Aufgaben-)Adaptivität.

Die Simulationspartizipantinnen sollen die Möglichkeit haben, den SuS verschiedene Aufgabentypen, differenziert nach kognitiven Dimensionen, Wissensdimensionen und Aufgabenschwierigkeit zu stellen, auf welche diese individuell unterschiedlich reagieren. Das Simulationsmodell ist so zu gestalten, dass Aufgaben Einzelpersonen, Gruppen oder der ganzen Klasse zugeordnet werden können. Es soll ein Bewusstsein dafür geschaffen werden, dass SuS unterschiedliche Aufgaben benötigen. Bereits früh in ihrer professionellen Karriere würden Lehramtsstudierende durch Simulationsnutzung das Bewusstsein für die Notwendigkeit adaptiver Aufgabenstellung entwickeln. So kann eine möglichst optimale Passung der gestellten Aufgaben an die Anforderungen der SuS eingeübt und im besten Falle später in den Berufsalltag integriert werden. Weiterhin soll vermieden werden, dass angehende Lehrkräfte primär Fragen auf niedrigem kognitivem Niveau stellen (Lipowsky et al., 2007, S.127). Nach Sacher (1995, S.98) befinden sich 80% der gestellten Lehrkräftefragen auf der kognitiven Dimension *knowledge* und *comprehension* (aus der Bloom'schen Taxonomie entlehnt, entspricht *remember* und *understand* in der Anderson'schen Taxonomie).

Zuletzt ist ein didaktisches Ziel die Einhaltung von optimalen Wartezeiten nach der Fragestellung bis zum Aufrufen der SuS (nach Sacher, 1995, S.102; Lipowsky et al., 2007, S.127, ca. 3-5 Sekunden; Borich, 2014).

(c) (Self-)Assessment

Die Erfahrung der Simulationsnutzung soll den Partizipantinnen schon frühzeitig ein Bild davon vermitteln, in welchen Situationen sich eine Lehrkraft bewegt und welchen Herausforderungen sie gegenübersteht. Im Rahmen eines präcurricularen (Self-)Assessments können am Lehramtsstudium Interessierte schon früh herausfinden, ob das Studium für sie eine Zukunftsperspektive sein könnte. Der simulierte Rollenwechsel dient auch dazu durch die eigene Schülerinnenperspektive verzerrte Konzeptionen von schulischem Unterricht aufzubrechen.

(d) Umgang mit Diversität und Heterogenität

Bereits in der Simulation sollen die Partizipantinnen lernen, mit den Herausforderungen einer immer heterogener werdenden Schülerinnen-schaft umzugehen. Es ist erforderlich ein realistisches Klassengebilde zu konstruieren, welches auch eine authentische, kulturell, ethnisch und kognitiv diverse Zusammensetzung abbildet, mit den Gruppen von SuS, die die Lehrkräfte in Zukunft auch erwarten werden. Insbesondere ist mit einer Simulation messbar, ob ein bestimmter Teil der SuS in Aufrufen, Taskadressierung und ggf. Blickkontakt bevorzugt wird. Dies wäre in anschließenden Feedbackgesprächen zu thematisieren.

(e) Fachdidaktische Zielsetzungen

Da Clasivir nach einem modularen Baukastensystem (siehe [Abschnitt 5.2](#), S. 153) aufgebaut werden soll, sind einzelne Szenarien/Level so gestaltet, dass verschiedene fachdidaktische Zielsetzungen verfolgt werden können. Es ist beispielsweise denkbar ein Szenario zu konstruieren, in dem es darum geht, Fachsprache möglichst exakt zu verwenden und falsche Begriffsnutzung didaktisch sinnvoll und konstruktiv zu korrigieren.

Sekundäre Ziele sind:

(a) Schulung von *digital literacy* / *ICT (Information and Communication Technology) literacy*

Wenn Partizipantinnen Clasivir benutzen, erfahren sie aus erster Hand die Möglichkeiten der VR-Technologie und des Lernens mit Simulationen. Da VR-Geräte bislang noch nicht flächendeckend verbreitet sind, ermöglicht dies einen eigenen Zugang der künftigen Lehrkräfte auf eine potentielle digitale Zukunftstechnologie. Es besteht die Hoffnung, dass die Partizipantinnen dazu angeregt werden, diese oder ähnliche Technologien in die Klassenzimmer der Zukunft zu tragen und didaktisch sinnvolle Einsatzszenarien zu entwickeln. Nur durch Vertrautheit mit digitalen Technologien, im Sinne von *ICT literacy* (International ICT Literacy Panel, [2002](#); Horz & Schulze-Vorberg, [2017](#)) kann ein mündiger Umgang mit den digitalen Werkzeugen ermöglicht werden.

(b) Positive Änderungen von Selbstwirksamkeitserwartungen oder Lehrkraftüberzeugungen (*belief systems*)

Viele Literaturquellen berichten von positiven motivationalen Effekten durch Simulationsnutzung (Badiee & Kaufman, [2015](#); Ferry et al., [2005](#), S.90; Alessi & Trollip, [2001](#), S.229, 245; Cheong & Kim, [2009](#); Höntzsch et al., [2011](#); Y.-C. Yeh, [2004](#); IsaBella et al., [2006](#)). Schon in der ersten

Version von Clasivir 1.0 (siehe [Abschnitt 8.3](#), S. 238), die im Verlauf dieser Arbeit entstand, waren diese motivationalen Effekte beobachtbar (Glocker et al., 2023). Da es Hinweise darauf gibt, dass positive Selbstwirksamkeitserwartungen und Überzeugungen von Lehrkräften mit deren Performanz zusammenhängen (Gundel et al., 2019, S.244), kann auch dieses didaktische Ziel erheblichen Einfluss auf die Verbesserung der Unterrichtsqualität haben.

(c) Messung diagnostischer Kompetenzen

Ähnlich wie im simulierten Klassenraum (SKR) (siehe [Abschnitt 4.6](#), S. 85) kann auch Clasivir benutzt werden, um diagnostische Kompetenzen zu messen. Dabei kann es sich um die Herausarbeitung eines impliziten Bias der Lehrkraft gegenüber bestimmten SuS handeln, was bereits schon beim Umgang mit Heterogenität erwähnt wurde. Ebenfalls ist vorstellbar Anteile an korrekten Schülerantworten im Vergleich zu den Schätzungen der Partizipantinnen auszuwerten.

Es lohnt sich die Grenzen der didaktischen Simulationsziele herauszuarbeiten. Man sollte höchst skeptisch gegenüber Simulationen sein, die vorgeben, sämtliche didaktischen Ziele realisieren zu können; in der Regel wird dann kein Ziel adäquat umgesetzt. Mit Clasivir lassen sich folgende didaktische Zielsetzungen nicht oder nur mit aufwendigen Anpassungen realisieren:

(a) Schulung von Gesprächsführungskompetenzen, außerhalb des Systems der IRE (Initiation-Replay-Evaluation)-Kaskaden (Mehan, 1979).

Die technische Umsetzung des Interaktionsmodells bedingt IRE-Kaskaden. Mehan (1979) legt nahe, dass IRE-Kaskaden ein häufig genutztes Kommunikationsmittel im Unterrichtsgeschehen sind: „The three-part sequential pattern [gemeint ist das initiation-replay-evaluation pattern, Anm. d. Autors] predominates in teacher-initiated interaction in the nine lessons in the corpus“ (Mehan, 1979, S.54). Zudem sind 81.1% aller IRE-Kaskaden Lehrkraft-initiiert (Mehan, 1979, S.80), was sich in der Konstruktion des Simulationsmodells von Clasivir widerspiegelt. Dort können SuS zwar eigenständige Handlungen durchführen, werden aber zu einem Großteil von den Aktionen der Lehrkraft beeinflusst.

(b) Regeln und Ritualetablierung, z.T. Störungsprävention

Im Rahmen der in Clasivir verwendeten Agentenmodellierung mittels *behavior trees* ist es nicht möglich selbstlernende Agenten zu erschaffen, die in einer sinnvollen Art und Weise neu geschaffene Regeln und Rituale verinnerlichen. Dasselbe gilt für Störungsprävention. Es mag auf

den ersten Blick paradox erscheinen, dass Umgang mit Störungsverhalten ein zentraler Teil des didaktischen Ziels ist, jedoch ein wichtiger Aspekt, nämlich die Störungsprävention nicht modelliert werden kann. Ähnlich wie Regeln und Rituale, erfordert auch eine Implementation von Störungsprävention selbstlernende Agenten, die mit den heutigen Technologien nicht umzusetzen sind. Es handelt sich demzufolge um eine technische Limitation. Will man den Umgang mit Störungsverhalten holistisch vermitteln, ist neben der Simulationsnutzung, die den akuten Umgang mit Störungsverhalten thematisiert, noch eine weitere Lernkomponente (z.B. Seminar) zum Thema Störungsprävention erforderlich.

(c) Selbstreguliertes Lernen (Boekaerts, 1999) / Lernbegleitung

Beim selbstregulierten Lernen haben die SuS eine weitaus wichtigere Rolle für den Ablauf einer Stunde als in einem Setting der direkten Instruktion (Kinder & Carnine, 1991). Ähnlich wie bereits unter Regeln und Ritualetablierung erläutert, ist es mit derzeitigen Technologien nicht möglich, plausibles Agentenverhalten dieser Art in einer Interaktionsmodellierung des Typs 5 zu generieren.

(d) Tiefgreifender Umgang mit Lernstörungen (z.B. ADHS, Autismus) oder visuellen, auditiven oder kineästhetischen Lernbeeinträchtigungen

Es wurde die Entscheidung getroffen im bisherigen Zustand des Simulationsmodells auf die Modellierung von Lernstörungen oder Lernbeeinträchtigungen zu verzichten. Begründung dafür ist, dass andere Interaktionsmodellierungen (insbesondere Typ 7, siehe [Abschnitt 4.19](#), 127) besser geeignet sind, um Interaktionen zu modellieren, bei denen es auf eine hohe feingranulare Ausgestaltung der SuS ankommt. In simSchool werden SuS VAK-Werte (visuelle, auditive und kinästhetische Fertigkeiten bzw. Beeinträchtigungen) zugewiesen, aber es scheint für die Partizipantinnen äußerst schwierig diese SuS zu erkennen und Kompetenzen im Umgang mit diesen Störungen/Beeinträchtigungen zu erwerben. Die Interaktionen können, bedingt durch die Interaktionsmodellierung, nur auf einer oberflächlichen Ebene ablaufen, weswegen bei Clavir auf eine entsprechende Modellierung verzichtet wurde.

(e) DaZ/DaF-Kompetenzen(Deutsch als Zweitsprache/Deutsch als Fremdsprache)

Aktuell berücksichtigt das Clavir-Interaktionsmodell keine Variable zur Sprachkenntnis, d.h. alle SuS verstehen den Unterricht – bei gleichen kognitiven Kapazitäten – in gleich guter Form. Dies ist zwar nicht

vollständig realistisch, aber ähnlich wie bei der Betrachtung der Lernstörungen, würde eine zusätzliche Variable zur ‚Sprachkenntnis‘ entweder die benötigten Interaktionskaskaden potenzieren oder aber nur eine sehr oberflächliche Modellierung zulassen. In künftigen Versionen von Clasivir könnte eine solche Variable ergänzt und eine Modellierung mit einem gewissen Aufwand ermöglicht werden. Für den Prototyp sprengte dies allerdings die zur Verfügung stehenden Ressourcen.

Zudem sei erneut bekräftigt, dass es nicht Ziel von Clasivir ist reale Praxiserfahrungen zu verdrängen oder zu ersetzen. Stattdessen sollen Synergieeffekte geschaffen werden. Beide Lerngelegenheiten haben das Potential sich gegenseitig zu bereichern.

5.1.2 DOIT-Modell: Organisation

Die Organisations-Komponente des DOIT-Modells beschäftigt sich mit den beteiligten Institutionen und der Koordination der digitalen Lerngelegenheit. Zur Realisierung von Clasivir ist es erforderlich, Programmier- und Grafikarbeiten durchzuführen, welche nicht allein durch universitäre Ressourcen zu realisieren sind. Es war daher angedacht, dies an einen Drittanbieter auszulagern, was in Form von Clasivir 1.0 auch bereits geschehen ist. Die technische Infrastruktur (HMD-Brillen, eventuell DOME-Umgebungen) wird hingegen von der Universität erworben und verwaltet. Zunächst wurde das Simulationsmodell in (*Blended Learning*-)Seminaren evaluiert und im Rahmen dieser Arbeit validiert. Außerdem wurde in diesen Seminaren die erste Version einer VR-Simulation (Clasvir 1.0) getestet. Bei weiteren positiven Forschungsergebnissen wäre eine Verstetigung durch curriculare Einbindung wünschenswert.

Da die VR-Brillen und ggf. DOME-Umgebungen von der Universität verwaltet werden, muss eine Nutzungsordnung für die Räumlichkeiten und Geräte entworfen werden. Da es geplant ist, aufgrund der potentiellen curricularen Verankerung später viele Studierende mit einer Simulationserfahrung zu versorgen, kann die Zeit für eine Simulationssession nicht beliebig lang gewählt werden. Ausgehend von den didaktischen Zielen und den Erfahrungswerten aus anderen Simulationen ist eine kurzfristige VR-Nutzung von 15 Minuten Simulationsnutzung, plus 10 Minuten Anleitungs- und Eingewöhnungszeit pro Simulationspartizipantin angedacht. Dies deckt sich auch mit den bereits gemachten Erfahrungen aus Clasivir 1.0. Denkbar sind auch mehrere 5 Minuten-Blöcke von Simulationsnutzung, die von Feedbackschleifen unterbrochen werden.

Weiterhin ist eine organisationale Einbindung von Lehrkräften, Fachdidaktikerinnen und Bildungswissenschaftlerinnen zur Szenariengenerierung

nötig. Diese Vernetzung ist von größter Wichtigkeit, da Clasivir ohne Szenarien nicht mehr als ein ungenutztes Grundgerüst bleiben wird. Ohne Szenarien ist Clasivir zu vergleichen mit einer Tech-Demo aus dem Computerspielbereich: Eine Zurschaustellung der technischen, grafischen Möglichkeiten einer Software ohne narrative Komponente. Es ist demzufolge wichtig, Kontakte zu dem Personal (z.B. Dozierende und Ausbilderinnen) herzustellen, das sich für einen Simulationseinsatz interessiert. Diese Personen können an der spezifischen Szenariengenerierung aktiv partizipieren und die Simulation als Multiplikatoren in ihre Lehrveranstaltungen tragen.

Feedback-Komponenten einer Simulation sind von zentraler Wichtigkeit. Es ist daher geplant, eine BAR (*before action report*) und eine AAR-Feedbackkomponente (*after action report*) in jede universitäre Simulationsnutzung verpflichtend einzubauen. Während der BAR-Komponente entwerfen die Partizipantinnen einen Plan, welches didaktische Ziel sie mit der Simulationserfahrung umsetzen wollen und was die geplanten Strategien sind um dieses Ziel zu erreichen. Es ist wünschenswert dies mit anderen, bereits existierenden, *blended learning* Komponenten zu verknüpfen.

In der AAR-Komponente ist es wichtig, die soziale Komponente des Lernens zu verstärken, indem eine Gruppendiskussion angeleitet wird. Diese sollte zweigeteilt sein, eine möglichst kurzfristig nach der Simulationserfahrung, eine ca. eine Woche später nach einer längerfristigen Nachbereitung. Die erste Gruppendiskussion sollte auf (a) emotionale Erfahrungen und (b) wahrgenommene Lernzuwächse der Simulationspartizipantinnen fokussieren. Beispiele für geeignete Fragen sind (a) „Wie habe ich mich vor, während und nach der Simulation gefühlt?“ und (b) „Was konnte ich durch die Simulationserfahrung mitnehmen?“. Die Simulation sollte in der Lage sein automatisierte Auswertungen in Form von Diagrammen bereitzustellen, um den Reflexionsprozess der Partizipantinnen anzuregen.

Letztlich ist auch geplant, die Simulationspartizipantinnen während ihrer Simulationserfahrung per Video aufzuzeichnen, sofern sich diese dazu bereit erklären. Diese Videos könnten während einer zweiten AAR-Gruppendiskussion zum Ausgangspunkt genommen werden um über intendierte Ziele, Umsetzung und Handlungsalternativen zu sprechen. Beide AAR-Komponenten sind wichtig, damit die Lernerfahrung nicht nur ein individueller Prozess ist, sondern der soziale Aspekt des Lernens im Sinne einer sozialkonstruktivistischen Perspektive nicht vernachlässigt wird.

Letztlich ist anzumerken, dass die aufgezeichneten Videos in Form von Ausschnitten auch in Lernportfolios verwendet werden und geteilt werden können –im Gegensatz zu realen Unterrichtsvideos, bei denen dies aus datenschutzrechtlichen Gründen nicht möglich ist. Insbesondere unter den neuen gesetzlichen Vorgaben, die die Novellierung des HLbG (Hessisches Lehrerbil-

dungsgesetz) bezüglich verpflichtender Portfolios macht, ist ein solcher Ansatz interessant.

Zusätzlich können weitere automatisierte Feedback-Komponenten der Simulation als Diskussionsgrundlage dienen. Denkbar sind beispielsweise:

- Auswertung des Aufrufverhaltens
- Log-Daten mit Aufschlüsselung der kognitiven Aufgabendimensionen, der Wissensdimensionen und des Schwierigkeitsgrads sowie der OCEAN- und Konzentrationswerte der SuS.
- *Heatmaps* der Position der Lehrkraft im Raum
- *Eye-Tracking* und automatisierte Analyse, wie oft zu welchen Entitäten geblickt wurde (Tafel, SuS, bestimmte Schülergruppen, etc.)

5.1.3 DOIT-Modell: Individuum

Auf der Individuum-Ebene des DOIT-Modells stellt sich die Frage, wie es gelingen kann, keine Benachteiligungen bei verschiedenen Nutzerinnengruppen entstehen zu lassen. Eine Reflexion, wie man die Akzeptanz der Simulation erhöhen kann, ist im Rahmen dieser Komponente auch sinnvoll. Es muss Begleitforschung initiiert werden, die etwaige Benachteiligungen von Personen ohne VR-/bzw. ohne digitale Spielerfahrungen untersucht. Da einige Personen unter der Simulatorkrankheit leiden, ist es wünschenswert, für diese spezielle statische Bewegungsmodi anzubieten, die das Auftreten von Übelkeit verhindern oder eindämmen können. Durch die Verwendung von Engines wie Unity ist es recht einfach möglich auch bildschirmbasierte Versionen von Clavisir anzubieten. Es ist darüber hinaus für alle Partizipantinnen erforderlich, dass entweder eine ausreichend große Fläche zum Agieren in einem virtuellen Klassenraum bereitgestellt wird oder die Bewegungen über Teleportationen realisiert werden, wie es bei anderen VR-Anwendungen üblich ist. Bezüglich der Akzeptanz ist außerdem darauf zu achten, dass etwaige genderspezifische Unterschiede analysiert und die Simulation ggf. angepasst wird. Ziel ist es zu vermeiden, dass die Simulation implizit auf eine bestimmte Gruppe hin konstruiert wird. Deswegen muss gerade in der Implementationsphase darauf geachtet werden diese Faktoren in den BAR- und AAR-Komponenten zu thematisieren, so Probleme auftauchen.

Weiterhin ist es im Individualkontext notwendig, die Simulation so zu konzipieren, dass ein *gaming-the-system*-Verhalten (Baker et al., 2008; Herz, 2002; Baker et al., 2006) nicht durch Simulationsmechaniken gefördert wird.

Dieses Verhalten kennzeichnet sich dadurch, dass Partizipantinnen die Simulationsmechaniken derart manipulieren, dass am Ende ein artifizierlicher Maximalpunktstand bzw. ein optimales Szenario erreicht wird, jedoch nichts über das didaktische Ziel der Simulation gelernt wird. Bestimmte Implementationen begünstigen ein derartiges ‚Lernen von Mechaniken‘ gegenüber dem erstrebten ‚Lernen von Prozessen‘. Wäre das Simulationsmodell von Clasivir beispielsweise so konstruiert, dass es keine negativen, sondern ausschließlich positive Auswirkungen gibt, wenn die Lehrkraft die gesamte Klasse ‚konstruktiv ermahnt/zurechtweist‘, so kann diese Mechanik, sobald das Funktionsprinzip verstanden wurde, ausgenutzt werden. Die Partizipantinnen könnten diese Handlungsalternative jedes Intervall, also alle 2 Sekunden, einsetzen um das Konzentrationsniveau der virtuellen SuS künstlich hoch zu halten. Ein konstruktives Ermahnen einer einzelnen Person wäre dann auch nicht zielführend, da eine Adressierung der ganzen Klasse immer ein besseres Ergebnis bringen würde. An diesem Beispiel kann man deutlich erkennen, dass dieses Verhalten nur noch wenig mit den didaktischen Zielen und den konkreten Prozessen des Lehrens zu tun hat. *Gaming the system* ist zwar nie gänzlich zu vermeiden, aber es gibt Möglichkeiten dem Problem zu begegnen. Zum einen kann man Interaktionen komplexer gestalten, so dass keine falschen Anreize gesetzt werden. Im konkreten Beispiel ist das ‚konstruktive Ermahnen/Zurechtweisen‘ in Clasivir so gestaltet worden, dass diese Handlung einen negativen Einfluss auf die Konzentration der SuS hat, wenn diese kein Störungsverhalten zeigen. So wäre ein ‚konstruktives Ermahnen/Zurechtweisen‘ der ganzen Klasse nur dann effektiv, wenn ein Großteil der SuS Störungsverhalten zeigt. Dies wäre ein intendierter Effekt, der sich im Simulationsmodell widerspiegelt.

5.1.4 DOIT-Modell: Technik

Die technische Komponente des DOIT-Modells betont, dass sich die Technik immer den didaktischen Zielen unterzuordnen hat. Deswegen ist kritisch zu hinterfragen, warum die Umsetzung als VR-Simulation und eine komplexe Interaktionsmodellierung erforderlich sind. Die Gestaltung von Clasivir als VR-Simulation ist ein wissenschaftliches Experiment. Es ist ungewiss, ob sich die VR-Technik für die konkrete Simulationsanwendung bewährt. Es gibt jedoch viele Hinweise darauf, dass VR-Lernumgebungen das *presence*-Gefühl steigern und weitere positive motivationale Effekte haben (siehe auch Kapitel [Abschnitt 2.5.7](#), S. 30). Dies wäre auch im konkreten Anwendungsfall weiter zu beforschen.

Zuletzt soll noch vorgestellt werden, wie die generelle Funktionsweise und Benutzerinnensteuerung von Clasivir angedacht ist. Da Clasivir als VR-Si-

mulation konzipiert ist, beginnen die Nutzerinnenaktionen mit dem Anlegen einer VR-Brille und dem Start eines Orientierungs- und Einführungsprogramms in den virtuellen Raum. Eine Instruktorin kann auf einem Desktop-PC die Klasse und den Blickwinkel der Partizipantin verfolgen, ähnlich wie dies in *Breaking Bad Behaviors* gelöst ist. Eine Anwesenheit der Instruktorin ist jedoch nicht zwingend erforderlich, sie greift – abgesehen von der Bereitstellung eines Szenarios – nicht in die Funktionsweise der Simulation ein. Wird die Simulation im Seminarkontext eingesetzt, so ist die Anwesenheit von Ausbildungspersonal intendiert, da diese die Handlungen der Simulationspartizipantinnen beobachten und anschließend in den Feedbackkomponenten thematisieren können. Es ist wünschenswert, dass auch eine Fernbeobachtung via Internet möglich ist, um verschiedene *blended-learning*-Szenarien zu ermöglichen.

5.2 Funktionsweise von Clasivir

Clasivir ist als VR-Simulation geplant. Partizipantinnen benutzen entweder ein HMD (*head-mounted display*) oder agieren in einer DOME-Umgebung und nehmen die Rolle einer Lehrkraft an. Die virtuellen SuS werden durch einen Algorithmus bzw. ein Simulationsmodell gesteuert, welches im Rahmen dieser Arbeit entwickelt wurde und dem Typ 5 der Interaktionsmodellierungen zugeordnet ist (siehe [Abschnitt 3.3](#), S. 54). Die Simulationsinteraktionen in Clasivir sollen nicht abstrakt, wie z.B. in *simSchool*, sein. Der entwickelte Prototyp (siehe [Abschnitt 5.4](#), S. 159) stellt allerdings alle Handlungen abstrakt dar, da er nur zur Überprüfung des Simulationsmodells dient. In der fertigen VR-Version von Clasivir stellen Partizipantinnen beispielsweise keine Frage mit der kognitiven Dimension ‚Erinnern‘, der Wissensdimension ‚Faktenwissen‘ und der Aufgabenschwierigkeit ‚Einfach‘, sondern sie wählen eine von vier von Bildungswissenschaftlerinnen, Lehrkräften oder Fachdidaktikerinnen vorselektierten und mit Simulationswerten versehenen Handlungsalternativen aus. Eine solche Frage könnte lauten: „Wer von euch weiß noch, was der Begriff ‚car‘ auf Deutsch bedeutet?“ Diese Frage wäre mit den Werten [Erinnern|Faktenwissen|Einfach] versehen (siehe [Abb. 10](#), S. 154 und [Abb. 11](#), S. 155). Die Partizipantinnen sehen dabei nicht, wie die Handlungsalternativen kodiert wurden. Eine Ausnahme sind alle *classroom-management*-relevanten Lehrkraftaktionen, die aufgrund ihrer Abstraktheit jederzeit ohne Vorformulierung ausgewählt werden können. So können alle SuS konstruktiv ermahnt werden, wenn sie Störungsverhalten zeigen. Da Störungsverhalten durch den Algorithmus bestimmt wird und damit unvorhersehbar eintritt, ist es nicht möglich alle Handlungsalternativen diesbezüglich vorzuselektieren. Auch können konkrete Handlungsalternativen mit *classroom-manage-*

ment Tags versehen werden, wenn ein Szenario dies erfordert.

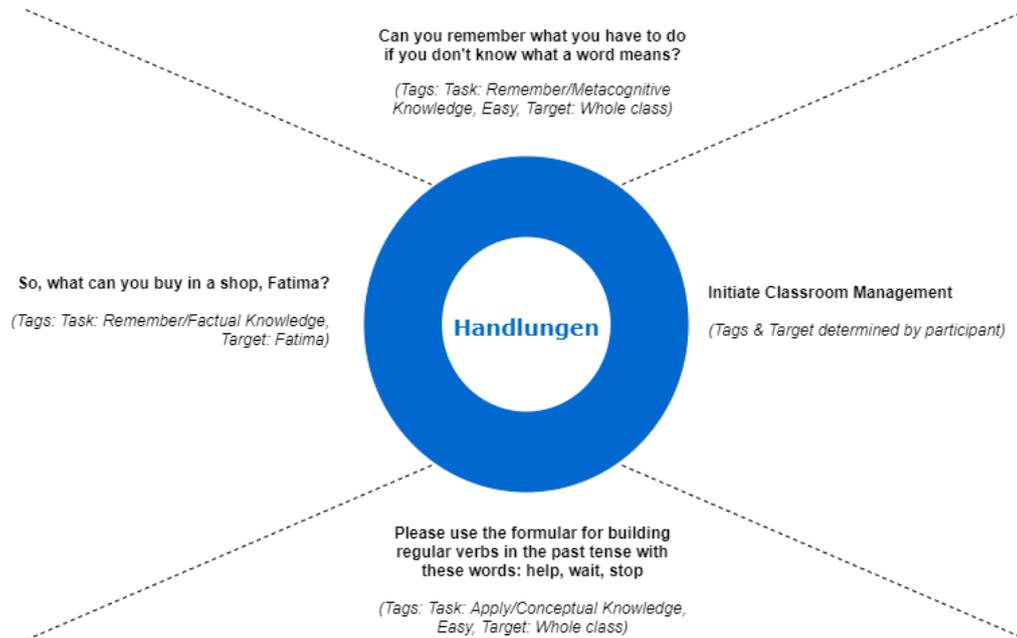


Abbildung 10: Schaubild eines Beispiels für Handlungsinitiiierungen durch die Simulationspartizipantinnen, Quelle: Autor

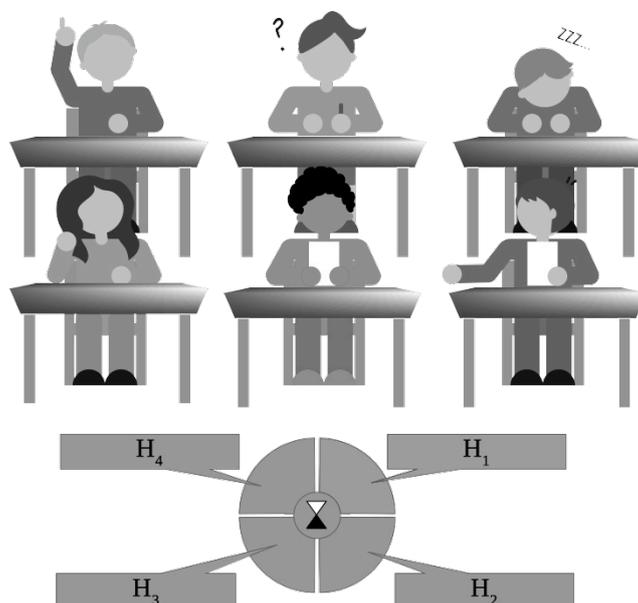


Abbildung 11: Schematische Darstellung der Partizipantinnenperspektive in Clasivir, Quelle: Autor

Die Erstellung von mehreren Fragen, Informationen, *classroom-management*-relevanten Vorgängen und SuS-Antworten soll an dieser Stelle Szenariengenerierung genannt werden. Welche SuS dann im Szenario wie reagieren, wird durch den Algorithmus bestimmt. Die Simulation läuft für die Partizipantinnen scheinbar in Echtzeit ab. Intern erstellt der Algorithmus jedes Intervall, also alle zwei Sekunden, einen neuen Zustand der situationsbezogenen Schülermerkmale, also beispielsweise, ob sich SuS zum Durchdenken einer Aufgabe entscheiden, ob sie Störverhalten zeigen, wie stark ihre Konzentration in diesem Moment ist oder ob sie sich melden.

5.3 Generelle Überlegungen zum Simulationsmodell

Das Szenario der Simulation beschränkt sich auf den öffentlichen Unterricht. Dedizierte Einzel-, Partner- und Gruppenarbeitsphasen werden aus den folgenden Gründen von der Simulation ausgeschlossen:

- In diesen Phasen findet z.T. keine direkte Interaktion mit der Lehrkraft statt. Wenn eine solche Interaktion stattfindet, handelt es sich dabei häufig um Hilfe auf individueller Ebene, welche technisch nur über Interaktionsmodellierungen des Typs 1–2 modellierbar ist.
- In diesen Phasen ist eine Modellierung des gesamten Klassengeschehens

didaktisch nicht zielführend: Die Lehrkraft beschränkt ihren Aufmerksamkeitsfokus auf einen Teilbereich der SuS.

- Eine Beschränkung ist sowohl aus forschungspragmatischen Gründen nötig, als auch, um überkomplexe Modelle („*bloated models*“) zu verhindern.
- Die Mehrzahl der vorgestellten Simulationen in der Lehrkräftebildung verzichtet auf die Darstellung von Einzel-, Partner- und Gruppenarbeitsphasen.
- Andere Formen der Aus- und Weiterbildung, z.B. Fallanalysen, *micro-teachings* oder Videoanalysen sind geeigneter um situationsrelevantes Wissen in diesem Bereich zu vermitteln.
- Auch neueste videoanalytische Studien (Böheim et al., 2020, S.7 mit N=56 analysierten Unterrichtsstunden) stellen fest, dass ca. 80% des Unterrichts in der Sozialform des öffentlichen Unterrichts verbracht wird. Dieser nimmt also den größten Anteil an der sozialen Situation Schulunterricht ein. Dies deckt sich auch mit den Ergebnissen von Appel (2015).

Diese Gründe implizieren aber weder, dass eine Modellierung von Einzel-, Partner- oder Gruppenarbeitsphasen didaktisch nicht sinnvoll ist, noch, dass diese Phasen generell unmöglich zu modellieren sind. Ein angedachtes Interaktionsmodell des Typs 5 eignet sich jedoch nur unter großem Aufwand zur Modellierung dieser Phasen.

Das Simulationsmodell bildet ausschließlich Unterrichtsgeschehen ab, weitere Aspekte des Lehrkrafthandelns wie Pausenaufsichten, Elterngespräche, Kommunikation mit Schulleitung, etc. werden nicht modelliert. Distale Faktoren nehmen nach dem Angebots-Nutzungs-Modell von Helmke (Helmke, 2015) Einfluss auf die Lernbewertung der SuS durch die Lehrkraft. Solche distalen Faktoren sind z.B. soziokultureller Hintergrund oder elterliche Unterstützung. Da jedoch kaum empirische Studien vorliegen, die kausale Beziehungen zwischen diesen Faktoren und Lernoutput postulieren und mediiierende Variablen schwer ausgeschlossen werden können, wird auf die Modellierung von distalen Effekten verzichtet.

Das Simulationsmodell übernimmt das bereits erfolgreich in anderen Simulationen wie *simSchool*, *SimClass* und *V.C.S.* (siehe [Abschnitt 4.16](#), S. 105, [Abschnitt 4.17](#), S. 121, [Abschnitt 4.15](#), S. 103, [Abschnitt 4.23](#), S. 134) etablierte Vygotskij'sche Paradigma der Zone der proximalen Entwicklung (Vygotskij, 1987b; Vygotskij, 1987a) zur Modellierung der generellen Aufgabenstruktur. In der programmiertechnischen Umsetzung wurde sich darum

bemüht, den nicht gänzlich zu vermeidenden Einfluss des Behaviorismus so gering wie möglich zu halten. Forschungsparadigmatisch lässt sich das Simulationsmodell von Clasivir dem Kognitivismus zuordnen.

Clasivir basiert neben dem Vygotskij'schen Paradigma der Zone der proximalen Entwicklung auf dem konzeptionstheoretischen Framework von Kunter und Trautwein (2013), das Unterrichtshandeln als Zusammenspiel von drei Dimensionen beschreibt: (1) Kognitive Aktivierung, (2) *classroom management* und (3) Unterstützung. Alle Variablen in Clasivir lassen sich einer oder mehrerer dieser Dimensionen zuordnen (siehe Abb. 12, S. 157) Dieses Konzeptionsmodell ist Ausgangspunkt sämtlicher weiterer Modellierungsschritte. Andere Konzeptionsmodelle, die zur Modellierung der spezifischen Variablen herangezogen wurden, werden in den jeweiligen Kapiteln vorgestellt und beschrieben.

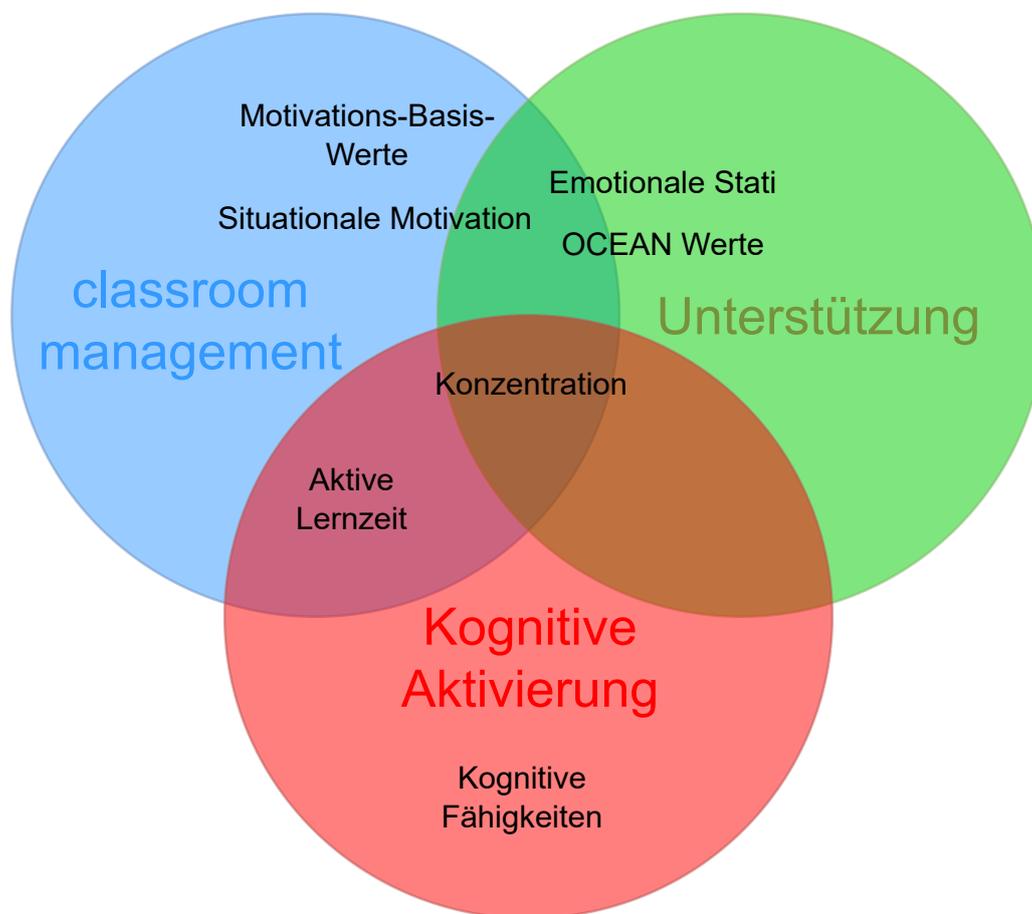


Abbildung 12: Zuordnung wichtiger Clasivir-Variablen zum verwendeten Framework, Quelle: Kunter und Trautwein (2013) und Autor

Das Verhalten der SuS im Simulationsmodell wird durch vier Teile ermittelt: (a) Durch das Lehrkrafthandeln, (b) durch die personenbezogenen Variablen der SuS, (c) durch die situationsbezogenen Variablen und (d) durch verschiedene Randomisierungsfunktionen, die bei gleichen Inputs nicht immer dieselben Outputs liefern und damit die Unsicherheit menschlichen Verhaltens approximieren.

Die situativen Variablen ändern sich primär (aber nicht ausschließlich) durch den Wert des vorangegangenen Intervalls. Dies ist mit Mehan (1979) zu begründen:

A stochastic model of behavior is Markovian in nature. It assumes that the next event is affected most by the immediately prior event. An interactional model recognizes, first of all, that behavior can be influenced by immediately prior events, those in the distant past, and those that have yet to occur (Mehan, 1979, S.76f.).

Das Lehrkrafthandeln umfasst folgende Kategorien, die dem kommunikativen Zweig einer IRE-Kaskade zuzuordnen sind:

- Eine Aufgabe stellen
- Informationen geben
- Eine/n SuS aufrufen
- Feedback zu einer Aufgabenlösung geben (loben/*praise*, bejahen/*affirm*, verneinen/*negate*, kritisieren/*criticize*) (J. E. Brophy & Good, 1970; J. Brophy, 1981)
- *classroom management*-bezogene Aktionen (loben, im Verhalten bestärken, konstruktiv kritisieren, destruktiv kritisieren, Schülerin in Parallelklasse oder zum Direktor schicken)

Auf die Modellierung organisationaler Lehrkraft-SuS-Interaktionen wie Öffnen/Schließen der Fenster, Organisation eines Klassenausfluges, Klassenbucheintragen, etc. wurde verzichtet. Die Inklusion dieser Aspekte hätte das Simulationsmodell überbordend komplex gemacht (*bloated model*) und hätte nur geringen didaktischen Mehrwert. Außerdem ist empirischen Studien zu entnehmen, dass diese Art der Interaktion im Unterrichtsgeschehen selten vorkommt (Appel, 2015).

Das in dieser Arbeit entwickelte Simulationsmodell läuft im Hintergrund der Simulation. Simulationspartizipantinnen können in einer zukünftig fertiggestellten Simulation aus vier – ausformulierten – Handlungsalternativen auswählen. Dieses Verfahren wurde aus drei Gründen gewählt:

- Durch die Beschränkung auf vier Handlungsalternativen wird der *cognitive load* der Simulationspartizipantinnen minimiert. Dies ist vor dem Hintergrund einer unerfahrenen Zielgruppe wichtig.
- Im Gegensatz zu Simulatoren wie V.C.S. und simSchool entsteht in Clasivir nicht der Eindruck eines Experimentalsettings. Damit ist die Hoffnung verbunden, dass dieser Ansatz die *presence* erhöht.
- Es ist möglich gezielte Szenarien zu generieren, um thematische und fachdidaktische Fragestellungen einzubringen.

5.4 Implementation eines Prototyps

Im Rahmen der Entwicklung und Testung des interaktionalen Simulationsmodells entwickelte der Autor einen text- bzw. tabellenbasierten Prototypen (siehe Abb. 13).

The screenshot shows the 'Clasivir Prototype' application window. At the top, it displays 'Schulform: Realschule', 'Kognitives Niveau der SuG umwandern: Realschule', 'Simulationsintervall: 4', and 'Jahrgangsstufe: 6. Klasse'. Below this, there are fields for 'Schüler: 21', 'Zeitpunkte: 252', and 'Angezeigtes Intervall: 4'. A checkbox 'Zeige nur aktives Intervall (empfohlen)' is checked.

The main part of the window is a table with columns: eid, Vorname, Nachname, Geschlecht, Schüleralter, Land, Staatsangehörigkeit, O, C, E, resen, Name, concentrationStat, _emotionalState, _AnimState, _Verarbeitung, and _Stoerverhalten. The table lists 17 students with their personal details and simulation-related data.

Below the table, there are several control buttons: 'Lehrkrafthandlung', 'Simulationsintervall inkrementieren (keine Handlung)', 'Continue', and 'SUS-Tabelle generieren'. There is also an 'Einstellungen' button and a 'Test' button.

At the bottom, a log window shows simulation events such as 'Chance für Verarbeitung von Chiara Reinhard: 0.889285', 'Task-Status für Maximilian Knorr: Task zu einfach', and 'Maximilian Knorr entscheidet sich zur Verarbeitung, trotz zu einfachem Task (+0.10 Con.)'.

Abbildung 13: Screenshot des programmierten Prototyps, Quelle: Autor

Dieser Prototyp ist nicht der geplante VR-Simulator und ist nicht zum Einsatz für Lehramtsstudierende gedacht. Er dient als Werkzeug bei der internen Konzeption und Evaluation des Simulationsmodells. Die mathemati-

sche und algorithmische Modellstruktur ist von einer Komplexität, die es unmöglich macht, das Simulationsmodell über einfache Techniken wie Makros oder Funktionen in Tabellenkalkulationen umzusetzen. Der Prototyp wurde in QT/C++ entwickelt und benutzt eine SQL-Datenbank zur Speicherung aller situations- und personenbezogener Variablen. Der Prototyp wurde verwendet, um die Datengrundlage der Videomockups zu generieren, die im empirischen Teil verwendet wurden (siehe [Abschnitt 6.3](#), S. 200).

Viele der in den kommenden Kapiteln angegebenen Zahlenwerte, insbesondere jene, die stärker auf Ausprobieren der Simulation basieren und eine schwächere empirische Basis haben, können von Usern beeinflusst werden. Clasivir speichert im User-Ordner des jeweiligen Betriebssystems die Clasivir.ini Datei. Diese kann via Editor oder direkt in Clasivir bearbeitet werden. Dort sind auch Beschreibungen aller editierbaren Variablen aufgeführt. Eine vollständige Liste dieser Variablen ist auch im Anhang zu finden ([Abschnitt 8.3](#), S. 266). Gewünschte Änderungen an den Formeln (z.B. Änderung der biexponentiellen Funktion in eine lineare) müssen im Quelltext geändert werden.

5.5 Persönlichkeitsbezogene Variablen: Namen der SuS / Diversity

Die meisten Simulatoren in der Lehrkräftebildung messen den Namen der SuS keine große Bedeutung zu. Gerade, wenn es darum geht, einen heterogenen und diversen Lernraum zu schaffen, ist es unerlässlich sich mit der Generierung von SuS-Namen zu beschäftigen. Es existieren verschiedene Ansätze, um Namen in Simulatoren zu generieren:

- Es existiert eine fixe Anzahl an festen Vor- und Nachnamenkombinationen, die normativ gesetzt wurden. Diese dürfen aus datenschutzrechtlichen Gründen nicht auf realen SuS basieren, weswegen die Namen fiktiv sein müssen. Dieser Ansatz kann durch die normative Setzung zu Problemen in der Darstellung von Diversität führen. Zudem wiederholen sich SuS-Namen, wenn Namen randomisiert gezogen werden. Werden Namen nicht randomisiert gezogen, ist nur eine Klasse modellierbar.
- Vor- und Nachname werden getrennt aus einer normativ gesetzten Liste von möglichen Vor- und Nachnamen randomisiert gezogen. Nachteil dieser Methode ist wie bei der vorherigen, dass man unter Umständen Diversität nicht in dem Maße abbildet, wie es der Realität entspricht. Zudem können Vor- und Nachnamenkombinationen aus unterschiedlichen Sprachbereichen kombiniert werden, die aufgrund der kompletten

Randomisierung nicht schlüssig wirken können.

- Vor- und Nachnamenkombinationen werden gewichtet oder ungewichtet randomisiert aus einer oder mehreren Listen gezogen, die auf empirischen Daten fußen.
- Vor- und Nachname werden getrennt gewichtet oder ungewichtet randomisiert aus einer oder mehreren Listen gezogen, die auf empirischen Daten fußen.

Vor- und Nachnamenkombinationen sind in öffentlich zugänglichen Datenbanken so gut wie nie hinterlegt. Die gewählte Methode die Namen der virtuellen SuS zu bestimmen, orientiert sich daher am letzten Ansatz. Grundidee ist es, mehrere Listen unterschiedlicher Sprach-/Kulturräume (Migrationshintergrund) zu definieren, aus denen passende Vornamen und Nachnamen randomisiert gezogen werden. Die unterschiedlichen Sprach-/Kulturräume werden über Nationen approximiert. Die Chance dafür, dass eine Schülerin einen bestimmten, nach Nation getrennten, Migrationshintergrund besitzt, wurde empirisch wie folgt gewonnen: Im statistischen Jahrbuch der Stadt Köln 2018 (Stadt Köln, 2018) wird die Bevölkerung recht feingliedrig aufgeschlüsselt. So ist zu entnehmen, welchen Migrationshintergrund, aufgeschlüsselt nach 101 Nationen, den 94.044 Einwohnerinnen und Einwohnern der Stadt Köln zuzuordnen sind (Stadt Köln, 2018, S.44f.). Insgesamt 7709 Einwohnerinnen und Einwohner sind nicht nach Nation aufgeschlüsselt, sondern in mehrere „Übrige“-Kategorien verteilt, weswegen diese Einwohnerinnen und Einwohner nicht näher betrachtet werden können. Die verwendeten Zahlen sind in [Tabelle 15](#) (S. 266) im Anhang aufgeführt. Nicht mehr existierende Länder wurden auf deren Rechtsnachfolger übertragen. Dies betrifft aber nur eine zu vernachlässigende Anzahl von Einwohnerinnen und Einwohnern.

Dem statistischen Jahrbuch ist zu entnehmen, dass der Anteil der Einwohnerinnen und Einwohner mit Migrationshintergrund unter 18 Jahre 53.7% beträgt, was bedeutet, dass 46,3% der Einwohnerinnen und Einwohner unter 18 Jahre in die Kategorie Deutsche ohne Migrationshintergrund fallen. Von den 53.7% Einwohnerinnen und Einwohnern unter 18 Jahre sind wiederum 12.6% Ausländerinnen und Ausländer, was bedeutet, dass sie eine andere als die deutsche Staatsbürgerschaft besitzen. In Clasivir wird nun die randomisierte Ziehung einer Einwohnerin bzw. eines Einwohners simuliert: Zunächst wird bestimmt, ob die Person einen Migrationshintergrund hat. Dies ist der Fall, wenn ein randomisierter Wurf zwischen 0 und 1 sich über 0.463 befindet. Hat die Person keinen Migrationshintergrund wird ihr ein deutschsprachiger Vor- und Nachname zugewiesen. Hat sie einen Migrationshintergrund, wird ein weiterer randomisierter Wurf zwischen 0 und 1 durchgeführt, der die

Nation des Migrationshintergrunds bestimmt. Liegt der Wert zwischen dem Edge-Wert von $n-1$ und n , wird der Person ein Migrationshintergrund aus dem Land mit der Nummer n zugewiesen. Es folgt nun die Zuweisung eines Vor- und Nachnamens. Der Nachname bezieht sich dabei immer auf die zugewiesene Migrationsnation, der Vorname ist hingegen entweder deutsch oder der korrespondierenden Migrationsnation zugeordnet, falls die Person über die deutsche Staatsbürgerschaft verfügt. Verfügt die Person nicht über eine deutsche Staatsbürgerschaft, wird immer ein Vorname der korrespondierenden Nation gewählt. Die Chance für die Wahl eines deutschen Vornamens im Migrationsfall ist abhängig von der zugewiesenen Nation. Bei sämtlichen Vergaben von Vornamen ist das Geschlecht (weiblich, männlich, divers) zu berücksichtigen.

Es war notwendig, die entsprechenden weiblichen und männlichen Vornamenslisten und der Nachnamenslisten für alle der 101 einzeln aufgeführten Nationen zu erstellen und als Tabelle in das SQL-Format zu überführen. Dabei wurde folgendermaßen vorgegangen: Falls möglich wurden aktuelle Listen von beliebigen weiblichen und männlichen Babyvornamen herangezogen und in eine Excel-Tabelle überführt. Dies hat den Vorteil, dass die Namen aus diesen Listen aktuell und nach Geschlecht geordnet sind. Es gibt auch einen Nachteil: Die Listen sind fast immer nur ordinal geordnet, es ist also unklar, wie viele Personen diesen Vornamen führen. In der Regel waren solche Daten nur für reiche Industrienationen vorhanden, bei anderen Nationen musste ein anderes Verfahren gewählt werden. Hierzu wurde die Datenbank von <https://forebears.io> genutzt. Diese Datenbank enthält die Vor- und Nachnamen von aktuell ca. 4.04 Mrd. Menschen und kennt 27.7 Millionen einzigartige Vor- und 27,2 Millionen einzigartige Nachnamen. Wählt man ein Land aus, so kann man sich eine Liste der am häufigsten vergebenen Vor- und Nachnamen anzeigen lassen. Die Vornamensliste ist nicht getrennt nach Geschlecht. Um diese Namen nun der männlichen oder weiblichen Tabelle zuzuordnen zu können musste jeder Name einzeln nochmal in forebears.io aufgerufen werden. Dort kann man dann auf eine Liste zugreifen, zu welchem Prozentsatz der Name an eine weibliche Person oder an eine männliche Person vergeben wurde, aufgeschlüsselt nach Land. Um die Datenbankarbeit etwas zu erleichtern wurde ein Excel-Skript geschrieben, das die Vornamenauswahl der ersten 100 Namen per URL-Erstellung automatisiert. Aufgrund der Webseitenstruktur und Dateninkonsistenzen war es nicht möglich die Vergabe des Namens nach Geschlecht und Nation zu parsen, weswegen diese Aufgabe manuell durchgeführt worden ist. Die am häufigsten auftretenden Dateninkonsistenzen waren das Nicht-Vorhandensein der Kategorisierung von weiblichen Namen, insbesondere in einigen, aber nicht allen, kleinen arabischen oder afrikanischen Staaten, wo diese Daten scheinbar nicht erhoben oder zu-

gänglich gemacht wurden. In diesem Fall wurde eine Nachbarnation gewählt, für die die entsprechende Geschlechtszuordnung übernommen wurde. Eine zusätzliche Herausforderung stellten Unisex-Namen dar. Überstieg der Wert der Vergabe des Namens zu einem Geschlecht den Wert 70% wurde der Name nur dieser Liste zugeordnet. Für alle Werte von 30 bis 70% wurde der Namen beiden Listen zugeordnet. Es ist zwar wünschenswert noch separate Unisex-Listen zur Vergabe an Personen mit diversem Geschlecht zu entwickeln, jedoch musste auf Ressourcen Gründen darauf verzichtet werden. Personen mit diversen Namen werden durch Zufall entweder einem weiblichen oder einem männlichen Vornamen zugeordnet. Bei der Erstellung sämtlicher Tabellen wurde die ordinale Position der Namen, sowie, falls vorhanden, eine absolute Anzahl des Auftretens der Namen gespeichert. Lagen Daten zum absoluten Auftreten vor, so wurde auch eine Edge-Spalte, analog zum Verfahren bei Auswahl der Nation, erstellt. Aktuell nutzt der Prototyp diese Daten nicht, Vornamen werden aus den Nationslisten vollständig randomisiert gezogen. Eine Inklusion dieser Daten ist allerdings von Vorteil, da zukünftige Versionen von Clasivir berücksichtigen können, dass bestimmte Vornamen überproportional häufig vergeben werden (z.B. Maria oder Muhammed).

Die Zuordnung der Nachnamen war weniger problematisch, hier konnten die Namen entweder direkt aus forebears.io entnommen werden oder aus anderen, teils sogar ausführlicheren Listen. Insgesamt wurden 303 SQL-Tables erstellt, jeweils eine Tabelle für weibliche und für männliche Vornamen und eine für Nachnamen, getrennt nach den 102 Nationen (Deutschland mitgerechnet). Die weiblichen Vornamentabellen enthalten insgesamt 9187 Namen, im Durchschnitt sind 51 pro Staat vorhanden. Die männlichen Vornamen enthalten insgesamt 13.931 Namen mit durchschnittlich 52,5 Namen pro Staat. Die Nachnamentabellen enthalten insgesamt 20.360 Nachnamen mit durchschnittlich 101 pro Staat.

5.6 Persönlichkeitsbezogene Variablen: Emotionsmodell

Ähnlich wie in simSchool (siehe [Abschnitt 4.16](#), S. 105 & [Abschnitt 4.17](#), S. 121), Simfta ([Abschnitt 4.18](#), S. 124) und simClass ([Abschnitt 4.15](#), S. 103) verwendet auch Clasivir zur Modellierung der emotionalen Persönlichkeitsstruktur der SuS das sogenannte OCEAN-Modell, auch Big Five genannt (McCrae & Costa Jr., 1996; Howard & Howard, 2000). Alle SuS verfügen demnach über eine Offenheits-, eine Gewissenhaftigkeits-, eine Extraversion-, eine Verträglichkeits- und eine Neurotizismusvariable. Im Gegensatz zu simSchool handelt es sich bei diesen Variablen aber um *traits*, sie ändern sich nicht während der Simulation. Zur generellen Beschreibung des OCEAN-Modells sei an [Abschnitt 4.16.4](#) (S. 109) verwiesen. Das OCEAN-Modell wurde

aus drei Gründen zur Implementation in das Simulationsmodell ausgewählt: Zum einen wurde es in anderen Simulationen bereits erfolgreich eingesetzt, zum anderen ist es ein in der Psychometrie akzeptiertes und häufig eingesetztes Modell. Letztlich liegen für dieses Modell, im Gegensatz zu dem HEXACO-Modell (Lee & Ashton, 2012), offen zugängliche empirische Daten vor, die zur Modellierung der SuS-Persönlichkeitsprofile genutzt werden können (Schmitt et al., 2007).

Die Werte werden nicht randomisiert verteilt, sondern basieren auf einer gauss'schen Normalverteilung, deren Standardabweichung und Mittelwert für deutsche Versuchspersonen empirisch ermittelt worden sind (Schmitt et al., 2007). Aktuell modelliert der Prototyp von Clasivir zwar nur deutsche SuS, es wäre aber mit wenig Aufwand möglich, z.B. eine Klasse mit chinesischen, französischen oder amerikanischen SuS zu modellieren, da hierzu ebenfalls empirische Ergebnisse zur Merkmalsausprägung vorliegen. Um Einheitlichkeit zu gewährleisten wurden die OCEAN-Variablen auf Werte zwischen 0..1 transformiert. 0 bedeutet, dass das Merkmal kaum vorhanden ist, 1 hingegen, dass das Merkmal sehr stark vorhanden ist. Die Variablen werden programmintern als Fließkommazahl kodiert.

Da die reinen Zahlenwerte nicht ohne weiteres zu interpretieren sind, zeigt der Prototyp von Clasivir an, in welchen Bereichen der Standardabweichung sich der jeweilige Wert befindet. Diese ‚relativen OCEAN-Werte‘ reichen von „---“ (drei Standardabweichungen unter dem Mittelwert) bis zu „+++“ (drei Standardabweichungen über dem Mittelwert) und werden intern als *enums* (Ganzzahlstati) gespeichert.

Die OCEAN-Werte bestimmen und beeinflussen viele situative Variablen im Clasivir Simulationsmodell. Diese Beeinflussung ist in den jeweiligen Kapiteln zu den situativen Variablen nachzulesen.

5.7 Persönlichkeitsbezogene Variablen: Kognitive Fähigkeit

Im aktuellen Simulationsmodell von Clasivir existiert nur eine Variable zur Modellierung von kognitiven Fähigkeiten. Konzeptionell zerfällt diese Variable in verschiedene Formen von Intelligenz (siehe [Abschnitt 4.16.4](#), S. 109) und domänenspezifisches (Vor-)Wissen (Seidel, 2006), jedoch würde eine Implementierung auf dieser feingranularen Ebene das Modell sehr komplex machen. Insbesondere das Aufgabenmodell müsste erheblich erweitert werden, wobei zweifelhaft ist, ob eine Ausdifferenzierung z.B. in verschiedene Intelligenzformen wirklich etwas Signifikantes zur Simulation beiträgt. Würde man Clasivir benutzen, um wirklich sehr viele Unterrichtsstunden in verschiede-

nen Fächern zu modellieren, ist eine Einbeziehung des domänenspezifischen Wissens sicher sinnvoll, allerdings ist Clavivir nicht für diesen Einsatzzweck konzipiert. In Zukunft ist nicht ausgeschlossen domänenspezifisches Wissen als Variable zu ergänzen, allerdings wurde zum jetzigen Zeitpunkt darauf verzichtet. Im Verlauf dieser Arbeit wird die kognitive Fähigkeit mit KF abgekürzt.

5.8 Situative Variablen: Konzentration

Um eine realistische Approximation von Aufgabenadaptivität und Störungsverhalten zu generieren, ist es notwendig, eine Variable zu implementieren, welche die Fähigkeit der SuS zur willentlichen Fokussierung ihrer Aufmerksamkeit beschreibt. Diese Variable ist situativ, ändert sich also über die Zeit der simulierten Unterrichtsstunde. Es handelt sich explizit nicht das transsituativ stabile Eigenschaftskonzept von Aufmerksamkeit als *trait*, wie es in der Aufmerksamkeitsdiagnostik betrachtet wird (Moosbrugger & Goldhammer, 2006, S.84).

Im Gegensatz zur allgemeinen Bevorzugung von englischsprachigen Konstrukten in anderen Teilen dieser Arbeit, ist hier der Begriff Konzentration (von lat. *concentra* „zusammen zum Mittelpunkt“) zu bevorzugen, der auf einen spezifisch deutschen Forschungsdiskurs rekurriert (Goldhammer & Moosbrugger, 2006, S.29f.). Im Folgenden sollen kurz verwandte Begriffe erläutert und der Unterschied zum Begriff der Konzentration herausgearbeitet werden.

Das Konstrukt *alertness* beinhaltet die Regulation der physischen und psychischen Reaktionsbereitschaft bzw. Wachheit und ist eine Voraussetzung für die Aktivierung von Aufmerksamkeit (Goldhammer & Moosbrugger, 2006, S.24), während Konzentration die Willensleistung des Individuums betont und damit für eine Modellierung von Unterrichtsinteraktionen besser geeignet ist. Äquivalent dazu beziehen nach Goldhammer und Moosbrugger (2006, S.30) einige Autorinnen Aufmerksamkeit nur auf das Wahrnehmen, während Konzentration eine bestimmte Art der kognitiven Anstrengung ist, die man als „energetischen Aspekt der Informationsverarbeitung“ beschreiben kann. Damit eignet sich das Konstrukt Konzentration dazu, auch situationale Motivation als medierende Variable miteinzubeziehen (Vgl. Bunce et al., 2010, S.1438).

Arousal hingegen scheint als aggregiertes Konstrukt zunächst gut zu passen, geht aber davon aus, dass es ein optimales *Arousal*-Niveau gibt und zu hohe Werte zu Nervosität führen, so dass die Leistungen abnehmen (Bligh, 1972, S.67). Dies ist mathematisch nur kompliziert modellierbar und schlechter zur kausalen Verbindung von Störungsverhalten geeignet. Es ist davon auszugehen, dass Störungsverhalten ein Coping-Mechanismus ist, der *arousal*

erhöht, da der Unterricht nicht genug Reize liefert. Dann wiederum ist die Variable als Mittel, um zu überprüfen, wie sehr eine Schülerin der aktuellen Unterrichtsaufgabe folgen kann, nicht geeignet.

Das Konzept von Aufmerksamkeit (*attention*) ist konzeptionell im Simulationsmodell nicht optimal nutzbar, auch wenn in der Modellierung von Aufmerksamkeitsverfall (*attention decay*) mathematisch darauf Bezug genommen wird. Diese mathematische Übertragung ist damit zu rechtfertigen, dass der Begriff *attention*, insbesondere in Verbindung mit *attention decay*, oft unscharf verwendet wird. Aufmerksamkeit ist konzeptionell ein Wahrnehmungsfokus und spielt sich im psychologischen Grundverständnis innerhalb von Zeiträumen im Millisekundenbereich ab (Ansorge & Leder, 2017, S.59ff.). Dies spricht gegen eine Modellierung über eine Stunde hinweg. Zudem gibt es ein ähnliches konzeptionelles Problem wie in der Benutzung des Konstruktes *arousal*. Es ist davon auszugehen, dass im Übergang zum Störungsverhalten lediglich ein Aufmerksamkeitswechsel (Goldhammer & Moosbrugger, 2006, S.24) stattfindet. Anders gesagt: Wendet sich eine Schülerin einem Störverhalten zu, spielt also beispielsweise mit ihrem Handy, so wechselt ihre Aufmerksamkeit nur von einem Gegenstand zum anderen, sinkt aber nicht. Diese Vermischung von unterschiedlicher Aufmerksamkeitsfoci macht das Konstrukt zur Implementation ungeeignet.

Lösung des Konstruktproblems ist die Verwendung von Konzentration, welche im spezifischen Clasivir-Simulationskontext definiert ist als „Konzentration auf die sozial erwünschte Form der jeweiligen Unterrichtssituation“.

Konzentration ist im Clasivir-Simulationsmodell eine Fließkommazahl-Variable mit dem maximalen Wert 1.0 (=hochkonzentriert der sozial erwünschten Form der jeweiligen Unterrichtsinteraktion folgend) und 0.0 (=kein Interesse an der Nachverfolgung der sozial erwünschten Form der jeweilige Unterrichtsinteraktion). Dieser Ansatz ist äquivalent zu dem von Jewett und Kronauer (1999), auch wenn diese ein anderes Konstrukt verwenden:

For ease of understanding, the models of subjective alertness and cognitive throughput presented here have been scaled so that 1.0 represents the maximum measurable level of alertness or throughput and 0.0 represents the minimum measurable level of alertness or throughput (Jewett & Kronauer, 1999, S.590).

Alle SuS verfügen über unterschiedliche personenbezogene Ressourcen, um ihre Aufmerksamkeit willentlich zu steuern (Goldhammer & Moosbrugger, 2006, S.19). Um diese verschiedenen Ausgangslagen adäquat zu modellieren, wird der Anfangswert der Konzentrations-Variable auf eine konzeptionell

sinnvolle Aggregation der OCEAN-Variablen gesetzt.

Ausgangswert

Konzeptionell ist davon auszugehen, dass vor allem die Gewissenhaftigkeit darüber entscheidet, ob jemand die Willensanstrengung aufbringt, seine Aufmerksamkeit ausschließlich der sozial erwünschten Interaktion zu widmen, weswegen der C-Faktor doppelt gewichtet in die Formel einfließt. Erwähnenswert ist, warum Offenheit als Faktor ausgeklammert wurde: Ein hoher Offenheitsfaktor spricht dafür, dass die Person neuen Dingen gegenüber neugierig eingestellt ist. Daher könnte man argumentieren, dass SuS die neu vorgestellten Aufgaben auch gegenüber einem Störverhalten bevorzugen würden. Dagegen ist allerdings einzuwenden, dass auch ein Störverhalten, z.B. das Malen von Comicfiguren, eine neue Erfahrung sein kann und eine Aufgabe nicht zwangsläufig einen Neuheitswert für die SuS besitzt. Daher wurde auf eine Inklusion des Offenheitsfaktors an dieser Stelle verzichtet.

Die Modellierung der mathematischen Funktion der Konzentration über den Zeitrahmen von einer Unterrichtsstunde im Schulgeschehen wurde bislang kaum empirisch beforscht, es existieren aber einige theoretische Annahmen darüber. Es gibt einen gewissen Forschungskonsens darüber, dass die Konzentration im Unterricht über die Zeit hinweg abnimmt (Wilson & Korn, 2007; Bradbury, 2016, S.510 erwähnt eine empirische Studie, basierend auf Selbstauskünften, dass dieser Verfall erst nach 10–20 Minuten einsetzt).

Die Vorlesungsanalysen von Bligh (1972) werden häufig zitiert und auch in modernen Forschungsartikeln zum Thema immer wieder herangezogen (S.511 Bradbury, 2016, z.B. Gibbs et al., 2011). Bligh postuliert darin eine Abnahme von Aufmerksamkeit über die Zeit (Bligh, 1972, S.72), welche die mathematische Form der Dichtefunktion besitzt. Bligh geht nicht darauf ein, warum diese Form der Modellierung gewählt wurde. Es ist wichtig zu betonen, dass es sich bei dieser mathematischen Funktionsform lediglich um eine Hypothese handelt, wie Bligh (1972, S.72) selbst eingesteht: „Empirical confirmation of this hypothesis is incomplete“.

Liegt für die jeweilige Schülerin im aktuellen Simulationsintervall keine kognitive Aufgabe vor oder wird die Verarbeitung der Aufgabe durch die Schülerin abgelehnt, so ist davon auszugehen, dass die Schülerin die Situation als monoton erlebt. Sie muss unter Aufbringung von Vigilanzkapazitäten (Goldhammer & Moosbrugger, 2006, S.24) gezielt daran arbeiten, ihre Konzentration aufrecht zu erhalten. Es ist theoretisch plausibel, dass diese Aufrechterhaltung von Konzentration bei fehlenden Reizen immer schwerer wird (Moosbrugger & Goldhammer, 2006, S.87). Der Verfall der Konzentrationsleistung konnte auch empirisch in einer Studie mit N=200 SuS zu 10

Messzeitpunkten à 20 Sekunden (N=745 SuS zur Modellvalidierung in derselben Studie) nachgewiesen werden (Lorenzi, 2013, S.11, 25, 45).

2018 erschien in Nature: Human Behavior eine Studie von Candia et al. (2018), die sich mit dem gut quantifizierbaren Aufmerksamkeitsverfall bezüglich kultureller Artefakte wie Filmen, Musik, Biographien, Patenten und Forschungsartikeln beschäftigten. Sie fanden heraus, dass sich bei allen Artefakten ein nicht-linearer biexponentieller Verfall des kollektiven Gedächtnisses und der kollektiven Aufmerksamkeit zeigt. Da die Zeiträume dieser Studie Jahre umfassen, ist es unwahrscheinlich, dass man dies auf ein Konstrukt wie klassenraumrelevante Konzentration direkt übertragen kann. Allerdings ist es augenfällig, wie universell der Aufmerksamkeitsverfall über alle Artefakte hinweg war. In der Studie von Jewett und Kronauer (1999) modellierten die Autorinnen durch die Datenanalyse von 15 *forced desynchrony* (FD) Experimenten eine mathematische Formel für *subjective alertness*, die ebenfalls nicht-linearer, exponentieller Natur ist. Die Autorinnen fordern:

Thus, any mathematical model of alertness or performance must incorporate these nonlinear interactions [...] (Jewett & Kronauer, 1999, S.589)

Da die FD-Experimente bei Jewett und Kronauer mit Schlafentzug arbeiten, sind auch hier die untersuchten Zeiträume nicht länger als in der Modellierung einer Unterrichtsstunde. Zusammengenommen, scheinen beide Studien darauf hinzuweisen, dass auch Unterrichtskonzentration in einem nicht-linearen Verhältnis zur Zeit steht. Die Studie von Lorenzi (2013) kommt mittels mathematischer Analyse von bayes'schen Netzwerken zum Ergebnis, dass lineare Modelle den Aufmerksamkeitsverfall, approximiert durch On- und Off-Taskverhalten, am besten beschreiben, aber exponentielle Modelle auch gute Werte beim Devianz-Informationskriterium erreichen (Lorenzi, 2013, S.24).

Zusammengefasst gibt es einige Anhaltspunkte dafür von einem biexponentiellen Verfall der Konzentration über die Zeit auszugehen. Bei der konkreten mathematischen Umsetzung im Simulationsmodell wurde die Dichtefunktion benutzt:

$$Con(x) = \sigma \cdot e^{-\frac{1}{2} \cdot (\frac{x-u}{l})^k}$$

wobei:

u=Intervall

σ =aktueller Konzentrationswert (Ausgangswert)

l=Agreeableness · GV_DIVISOR_CONCENTRATION_DECAY

GV_DIVISOR_CONCENTRATION_DECAY (Standardwert) = 32 (zur Erklärung der globalen Variablen, auch GV-Werte, siehe [Tabelle 28](#), S. 279)

Höhere u-Werte führen in der Formel dazu, dass der Graph auf der x-Achse nach rechts verschoben wird. Für σ ist der aktuelle Konzentrationswert

im Intervall zu wählen, von dem aus man den Konzentrationsabfall berechnen möchte. Der zweite Exponent k gibt an, wie stark die Stauchung vom absoluten Extrema zum Wendepunkt und vom Wendepunkt zur Asymptote zur x -Achse ist. Der l -Wert bestimmt die Länge der Verfallsfunktion bzw. die Streckung oder Stauchung der gesamten Kurve. Der Wert wird modifiziert durch den OCEAN-Wert der Verträglichkeit (*Agreeableness*).

Grund für die Verwendung der Dichtefunktion statt einer einfachen exponentiellen Funktion ist, dass damit der biexponentielle Anteil der nicht-linearen Beziehung modelliert wird (Candia et al., 2018; Vgl. auch Jewett & Kronauer, 1999, S.594). Dies ist auch konzeptuell zu begründen: Würde die Konzentration der SuS exponentiell steil abfallen, wären sie innerhalb weniger Intervalle sofort auf niedrigem Konzentrationsniveau, was für unrealistisches Verhalten sorgen würde. Eine konzeptuell sinnvollere Modellierung ist es, die Aufmerksamkeit zunächst nur leicht und dann stärker abfallen zu lassen, falls keine Inputs gegeben werden.

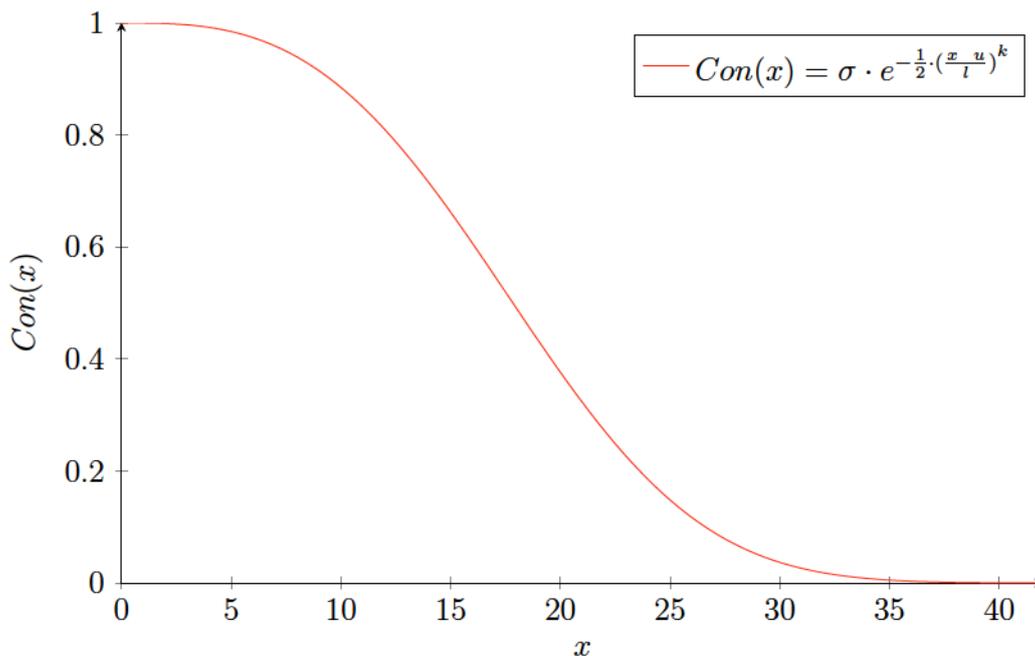


Abbildung 14: Graphische Darstellung der Konzentrationsverfall-Funktion

Damit ist die Modellierung der Konzentration noch nicht vollständig. Würde nur diese Formel benutzt werden, so würden alle SuS, egal, ob sie erfolgreich mit der Lehrkraft interagieren oder Aufgaben durchdenken und lösen, nach einer gewissen Anzahl von Intervallen ihre Konzentration fast vollständig verlieren. Wie auf dem Graph in Abb. 14 zu erkennen, ist das bei

einer Schülerin mit einem Agreeableness-Wert von 0.5 nach ca. 30 Intervallen (=60 Sekunden) der Fall.

Dieses Problem wird dadurch gelöst, dass die Konzentrationsfunktion zu allen Zeitpunkten, an denen eine sozial erwünschte Interaktion mit der Lehrkraft stattfindet, neu berechnet wird, wobei das aktuelle Intervall als Nullpunkt interpretiert wird. (siehe [Abb. 15](#)) (S. 170) Die Steigungstangente beträgt im Falle einer sozial erwünschten Interaktion also immer 0. Die Annahme einer Wiederherstellung der Konzentrationsleistung durch eine kognitiv fordernde Aufgabe oder eine Pause wird auch von anderen Autorinnen theoretisch angenommen (Lorenzi, 2013, S.26; Bligh, 1972, S.73). Zusätzlich wird der Konzentrationswert um eine geringe Konstante nach oben verschoben, wenn eine sozial erwünschte Interaktion vorliegt und um eine größere Konstante, wenn eine erfolgreiche Interaktion stattfand

Eine Interaktion wird als sozial erwünscht angesehen, wenn die SuS sich zur kognitiven Verarbeitung entschließen. Eine erfolgreiche Interaktion ist vorhanden, wenn ein SuS-Aufruf nach einer Meldung erfolgt oder von der Lehrkraft Feedback zu einer Antwort gegeben wird. Die Größe dieser Konstanten lässt sich nicht aus den empirischen Untersuchungen oder konzeptuellen Überlegungen herleiten, sondern wurde forschungspragmatisch durch mehrfaches Durchspielen des Modells sinnvoll gewählt.

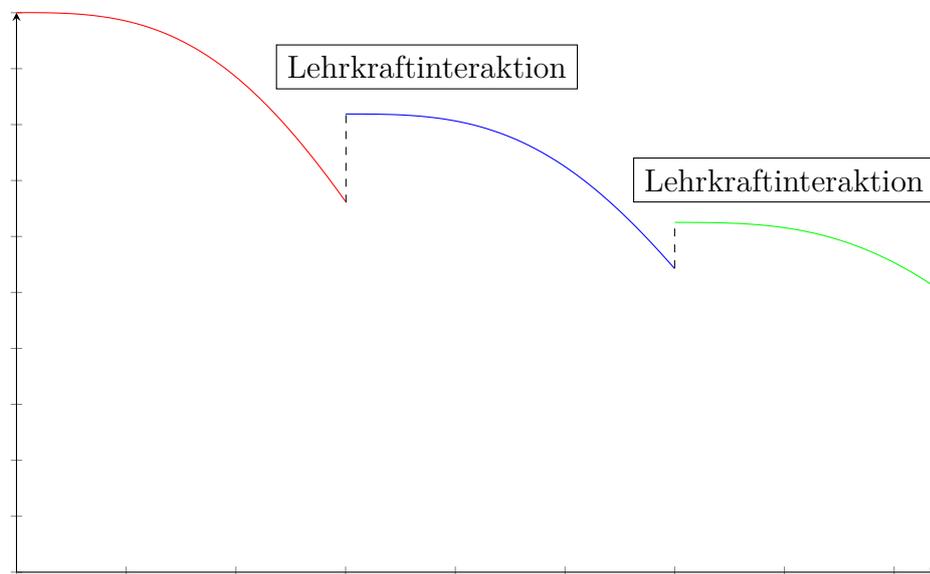


Abbildung 15: Beispiel der Konzentrationsverfall-Funktion nach Lehrkraftinteraktionen

5.9 Situative Variablen: Emotionaler Status

Die in Clasivir verwendeten emotionalen Stati orientierten sich an den sechs Grundemotionen ‚Freude‘, ‚Traurigkeit‘, ‚Angst‘, ‚Wut‘, ‚Überraschung‘ und ‚Ekel‘ (Mohanalan et al., 2017, S.2). Da im Unterrichtsgeschehen die Emotion ‚Ekel‘ eher selten und nur sehr kontextspezifisch auftritt, wurde auf eine Modellierung dieser Emotion verzichtet. Stattdessen wurde ein weiterer emotionaler Status implementiert, der ‚Neutral‘ genannt wurde und sich durch das Fehlen einer überwiegenden Grundemotion kennzeichnet. Dieses emotionsfreie abwartende Verhalten ist im Unterrichtsgeschehen oft bei SuS zu beobachten.

Emotionen verändern sich in Clasivir nach zwei verschiedenen Algorithmen: Ein hauptsächlich verwendeter Markov-Ketten-Algorithmus und ein zielspezifischer Algorithmus. Der Markov-Ketten-Algorithmus sucht nach der Emotion des letzten Intervalls und macht die Emotion des folgenden Intervalls davon abhängig. Die Wahl der Emotion erfolgt stochastisch, wobei die Wahrscheinlichkeiten für eine Beibehaltung oder einen Wechsel der Emotion unterschiedlich sind (siehe [Tabelle 16](#), S. 269 für eine Übersicht über diese Wahrscheinlichkeiten). Ist der emotionale Status einer Schülerin im vorherigen Intervall ‚Traurigkeit‘, so wird dieser Status mit der größten Wahrscheinlichkeit auch im kommenden Intervall beibehalten. Ein Wechsel zu ‚Neutral‘, ‚Angst‘ oder anderen Emotionen findet seltener statt. Begründung für diese Modellierung ist, dass die meisten Emotionen innerhalb der kurzen modellierten Zeitphasen dazu tendieren stabil zu bleiben.

Wird eine Emotion gewählt, bleibt diese Emotion für 10 Intervalle (=20 Sekunden) stabil, außer der zweite zielspezifische Algorithmus wird angewendet. Erst nach 10 Intervallen wird eine Folgeemotion bestimmt. Dieses Feature wurde aufgrund der gemachten Erfahrungen erst später in Clasivir implementiert. Gründe dafür sind folgende:

- Wird die Wahrscheinlichkeit für das Beibehalten der Emotionen extrem hoch gewählt, bleiben die meisten SuS während der Stunde fast ausschließlich in einem gleichbleibenden emotionalen Zustand, was nicht realistisch ist.
- Ist die Wahrscheinlichkeit für das Beibehalten der Emotionen zu niedrig gewählt, bleiben die Emotionen nicht stabil, sondern ändern sich aufgrund der Wahrscheinlichkeitsverkettungen (jedes Intervall wird eine neue Berechnung angestellt) schnell und erratisch.

Um diese Probleme zu beseitigen, wurde eine weitere Variable eingeführt, die die gewählte Emotion für 10 Intervalle stabil hält. Die einzige Ausnahme

davon ist die Emotion ‚Überraschung‘ die sich nur kurzfristig manifestiert: In Clasivir hat sie immer die Länge eines Intervalls.

Wird eine Schülerin Ziel eines aggressiven Störverhaltens einer weiteren Schülerin, folgt der Aufruf eines zweiten Algorithmus’. Der Markov-Ketten-Algorithmus wird für dieses Intervall nicht durchgeführt. Die Schülerin wechselt in diesem Intervall zu einem neuen emotionalen Status, der entweder ‚Wut‘, ‚Trauer‘ oder ‚Angst‘ sein kann, in Abhängigkeit von dem gezeigten Störverhalten. Beleidigt eine Schülerin eine andere, so reagiert diese auf die Beleidigung mit ‚Wut‘ oder mit ‚Trauer‘. Schlägt eine Schülerin eine andere, so reagiert diese mit ‚Angst‘ oder ‚Wut‘. Das konkrete Störverhalten wird aus diesen Listen randomisiert gezogen, bei einer Beleidigung wird beispielsweise zu 50% die Emotion ‚Trauer‘ und zu 50% die Emotion ‚Wut‘ vergeben.

5.10 Situative Variablen: Störungsverhalten

Fällt der Konzentrationswert einer Schülerin unter einen bestimmten Schwellenwert, so gibt es die Chance, dass diese Schülerin Störverhalten zeigt. Welches Störverhalten gezeigt wird, ist abhängig von einer Grundwahrscheinlichkeit für die einzelnen Störungsverhalten und einer Modifikationsvariable in Abhängigkeit von den OCEAN-Werten. Die Entscheidung Störverhalten in Abhängigkeit von der Konzentration zu implementieren, ist eine wichtige paradigmatische Entscheidung, die konzeptionstheoretisch darauf basiert, dass gelangweilte SuS eher zu Störungsverhalten neigen. Scherzinger et al. (2018) erwähnen zu diesem Thema unter Referenz auf J. Brophy (1999) und Doyle (1986): „Wenn Schülerinnen und Schüler genügend kognitiv gefordert werden, treten weniger Störungen auf“ (Scherzinger et al., 2018, S.389). Diese Aussage unterfüttern sie mit den empirisch gewonnenen Erkenntnissen ihrer Videostudie mit N=83 fünften und sechsten Klassen im deutschsprachigen Kanton Bern. Der Katalysator für das Auftreten von Störungsverhalten bei fehlender Konzentration ist Langeweile, insbesondere die Transformation in reaktante Langeweile (Götz & Frenzel, 2006; Götz et al., 2007). Diese Transformation tritt im Gegensatz zu anderen Coping-Verhalten vor allem bei SuS mit hoher Extraversion mit einer Wahrscheinlichkeit von 5–10% auf (Götz & Frenzel, 2006; Götz et al., 2007). Es ist theoretisch naheliegend Langeweile konzeptionell mit Konzentration zu verbinden, da Autorinnen wie Götz et al. (2006, S.129) die Entstehung von Langeweile zu einem Großteil mit dem spezifischen Fach oder der spezifischen Aufgabe verbinden („Inhalt Thema uninteressant/trocken“, „Inhalt/Thema zu schwer/zu einfach“, „Inhalt/Thema nutzlos“, „Fach uninteressant/trocken/nutzlos“).

Eine Reihe von bereits bestehenden Beobachtungsinstrumenten wurden daraufhin analysiert, ob sie Anhaltspunkte dafür geben können, welche Stör-

verhalten am häufigsten vorkommen. Diese wären dann konzeptionell-theoretisch in der Modellierung zu bevorzugen. Untersucht wurden (a) das SESBI (Sutter-Eyberg Student Behavior Inventory) (Floyd et al., 2004, S.68), (b) ein Fragebogen-Instrument von Pfitzner und Schoppek (2000, S.375ff.), (c) BASYS (Beobachtungssystem zur Analyse aggressiven Verhaltens in schulischen Settings) von Wettstein (2008), sowie (d) der Vorgängiger BAVIS von Humpert und Dann (1984) und das damit zusammenhängende Konstanzer Trainingsmodell (KTM)(Dann & Humpert, 2002). Die in dieser Arbeit getroffene Unterscheidung zwischen nicht-aggressiven Störverhalten, aggressiven Störverhalten und passiven Störverhalten ist von Scherzinger et al. (2018, S.390) abgeleitet.

Eine Liste aller implementierten Störungsverhalten ist im Anhang unter [Tabelle 17](#) (S. 270) zu finden.

Als Standardwert für den Basis-Konzentrations-Schwellenwert hat sich 0.3 als sinnvoll erwiesen. Inhaltlich formuliert: Sind die SuS nur zu einem Drittel konzentriert (und zudem wenig motiviert), besteht die Möglichkeit, dass Störverhalten auftritt. Der Konzentrations-Schwellenwert wird mit relativen OCEAN-Werten (also +++ bis ---) von Neurotizismus (stark), Offenheit (weniger stark) und Gewissenhaftigkeit (stark, umgekehrte Reihenfolge) modifiziert. Die exakten Werte wurden durch exzessives Durchspielen der Simulation gewonnen und sind im Quellcode nachzulesen. Eine Schülerin, die extrem neurotisch ist, sehr offen und kaum gewissenhaft, erreicht so einen maximalen Konzentrations-Schwellenwert von .495. Eine Schülerin, die kaum neurotisch, wenig offen und sehr gewissenhaft ist, erreicht einen minimalen Konzentrations-Schwellenwert von .1512.

Fällt die Konzentration in einem Intervall unter diesen Schwellenwert, berechnet ein zweiter Algorithmus, ob tatsächlich Störungsverhalten stattfindet. Die Wahrscheinlichkeit, dass eine unkonzentrierte Schülerin Störverhalten zeigt, setzt sich aus einer Basis-Chance und einer personenbezogenen Modifikation durch die OCEAN-Variablen für einige der 22 möglichen Störungsverhalten zusammen. Der Prozentwert für aggressives Störverhalten (verbal, physisch gegen Menschen, physisch gegen Gegenstände) wird mit den relativen OCEAN-Werten von Neurotizismus und Extraversion modifiziert, so dass die Wahrscheinlichkeit größer ist, dass eine extrovertierte und stark neurotische Schülerin aggressives Störverhalten zeigt. Nach der Multiplikation der Modifikationsvariable werden die Wahrscheinlichkeiten für alle einzelnen Störverhalten so transformiert, dass sie in der Summe wieder 1.0 ergeben. Somit verändern sich also auch alle nicht-aggressiven Störungswahrscheinlichkeiten leicht. Die Grundwahrscheinlichkeiten für einzelne Störverhalten (siehe [Tabelle 17](#), S. 270, Spalte 4) wurden durch Erproben der Simulation gewonnen und basieren auf folgenden empirischen Erkenntnissen:

- Kounin (1976) gibt folgende durch Beobachtung gewonnene Störungshäufigkeiten an, die auch in Seidel (2009) genannt werden:
 - Unterhaltungen 30%. In Clasivir ist die Grundwahrscheinlichkeit für Unterhaltungen 15% und wird durch die Extraversionsvariable modifiziert.
 - Kaugummikauen 6.8%. In Clasivir beträgt die Wahrscheinlichkeit für Kaugummikauen 4%.
 - Lärm, Geräusche, laute Unterhaltungen 25%. In Clasivir beträgt die Wahrscheinlichkeit für SV_GERAEU 10%. Dabei ist zu beachten, dass Beleidigungen und Bedrohungen, die in Clasivir separat modelliert werden, von Kounin in diese Kategorie eingruppiert werden.
 - Sachfremde Orientierungen 17.2%. In Clasivir sind diese sachfremden Orientierungen durch folgende interne Variablen abgebildet: $SV_ZEICHN + SV_HANDYT + SV_HANDYS + SV_SCHLAF = 19.5\%$. Für eine Übersicht der Variablennamen siehe [Tabelle 17](#), (270).
 - Restkategorie 21%, u.a. unerlaubtes Herumlaufen, Zuspätkommen, Hausaufgaben oder Hilfsmittel vergessen. In Clasivir wird nur unerlaubtes Herumlaufen mit 2% modelliert. Kounin füllt diese Kategorie mit vielen Verhaltensweisen, deren Eingruppierung als Störverhalten fraglich ist. Andere Verhaltensweisen sind in einer digitalen Simulation nicht operationalisierbar.
- Siestrup (2010) gibt folgende durch Beobachtungen gewonnene Störungshäufigkeiten an:
 - Aggressives Störungsverhalten tritt nur „sehr selten“ auf (Siestrup, 2010, S.145).
 - Verbales Störungsverhalten ist mit 42.4% das am häufigsten anzutreffende Störungsverhalten (Siestrup, 2010).
 - Passivität, Desinteresse, Opposition oder Angst sind zu 16.7% im Unterricht anzutreffen. In Clasivir wird lediglich Passivität und Desinteresse modelliert, subsumiert unter SV_BLICKFEN und SV_BLICKTISCH = 22%.
 - Verletzungen moralischer Normen wie Abschreiben oder Lügen 6,4%. Diese Variablen werden in Clasivir nicht modelliert, da sie schlecht operationalisierbar sind.

- Non-Verbale Aktivitäten 19.2%. Dies entspricht zu einem großen Teil Kounins sachfremder Orientierungen und beträgt in Clasivir 19.5%)
- Die Untersuchungen von (Appel, 2015) zeigen eine recht starke Ausprägung von motorischer Unruhe in 6. Klassen. In Clasivir beträgt die Wahrscheinlichkeit für das Auftreten motorischer Unruhe 12.5%.

Es ist offensichtlich, dass die detaillierten prozentualen Störungswahrscheinlichkeiten und die bereits erwähnten Konzentrations-Schwellenwerte oft unter theoretisch-konzeptionellen Prämissen gestaltet wurden. Es wäre wünschenswert, diese Variablen in Zukunft durch weitere empirische Forschungen zu verbessern. Bislang liegen Ergebnisse in der dafür nötigen Granularität noch nicht vor. Eine empirische Erhebung ist aufgrund der vielen miteinander verstränkten Variablen mitunter sehr schwierig.

Zusätzlich zu den durch die OCEAN-Werte modifizierten Grundwahrscheinlichkeiten ändern sich die Störungswahrscheinlichkeiten adaptiv während der Simulation, abhängig von den emotionalen Stati der SuS. (siehe [Tabelle 18](#), S. 272) Dies führt zu einem komplexeren Simulationsmodell, approximiert allerdings besser realistisches SuS-Verhalten. So ist beispielsweise die Chance, dass eine wütende Schülerin aggressive Störungsverhalten zeigt höher als bei anderen emotionalen Stati.

Die adaptiven Störungswahrscheinlichkeiten können im Prototyp von Clasivir über ein Tortendiagramm angezeigt werden. (siehe [Abbildung Abb. 16](#))

Findet als Störungsverhalten eine Unterhaltung statt, so greift ein Subalgorithmus, der bestimmt, ob die angesprochene Schülerin nun ihrerseits die Unterhaltung erwidert. Diese Modellierung wurde durchgeführt, damit SuS-Gespräche realistischer wirken und nicht nur von einer Schülerin geführt werden. Kommunikationskaskaden sind so möglich. Die Chance für die Erwidern einer Unterhaltung ist von der Extraversion und der doppelten Gewissenhaftigkeit abhängig:

Kommunikationserwiderung

5.11 Situative Variablen: Aufgabenmodellierung

5.11.1 Adaptive Aufgabenpassung

Das Stellen und Lösen von Aufgaben macht einen Großteil des Unterrichtsgeschehens aus. Daher ist der Modellierung dieser komplexen Interaktion besondere Aufmerksamkeit zu schenken. Wie bereits erwähnt, folgt die theoretische

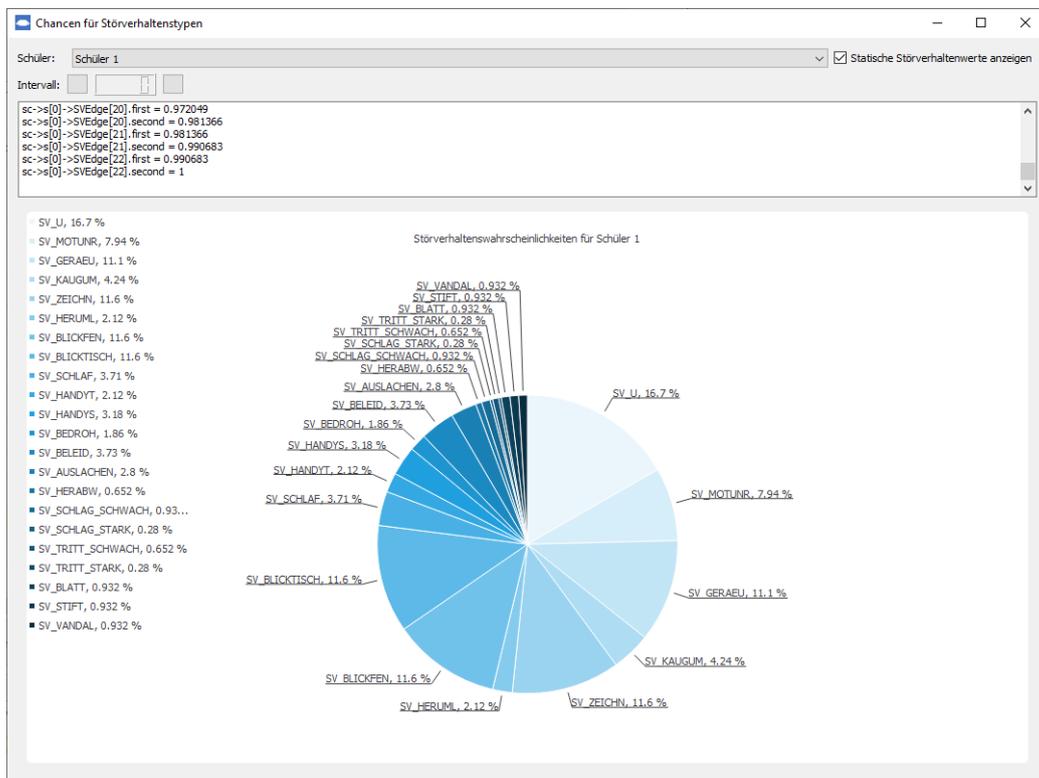


Abbildung 16: Screenshot des Clavivir Prototyps, Veranschaulichung der Wahrscheinlichkeiten für Störungsverhalten eines Schülers

Modellierung der Aufgaben dem Vygotskij'schen Paradigma der Zone der proximalen Entwicklung (Vygotskij, 1987a; Vygotskij, 1987b). Diesem Paradigma folgend können vier logische Kategorien adaptiver Aufgabenschwierigkeit unterschieden werden:

- (a) Aufgaben, die für die spezifische Schülerin zu einfach sind.
- (b) Aufgaben, die für die spezifische Schülerin normal lösbar sind.
- (c) Aufgaben, die im optimalen Lernbereich der Schülerin liegen, d.h. gerade noch lösbar sind.
- (d) Aufgaben, die außerhalb der Zone der proximalen Entwicklung liegen und deswegen unlösbar sind.

In welche dieser Kategorien eine Aufgabe für eine spezifische Schülerin fällt ist kausal abhängig von ihrer kognitiven Fähigkeit. Aufgaben kann man in unterschiedlicher Art und Weise klassifizieren. Eines der am häufigsten genutzten taxonomischen Modelle ist das von Anderson et al. (2000). Da diese Taxonomie bereits in Versatzstücken in simSchool 1.0 und 2.0 verwendet wurde, kann von einer Erprobung im Simulationsbereich gesprochen werden, wenngleich im simSchool-Simulationsmodell die Wissensdimension der Aufgaben nicht berücksichtigt wird.

Clasivir nutzt ebenfalls die Taxonomie von Anderson et al. (2000) und modelliert sie vollständig, d.h. jede Aufgabe wird mit einer kognitiven Dimension (*remember, understand, apply, analyze, evaluate* und *create*) und einer Wissensdimension (*factual, conceptual, procedural, metacognitive*) versehen. Zusätzlich gehen Anderson et al. (2000) davon aus, dass die Schwierigkeit einer Aufgabe ordinal von diesen Dimensionen abhängt, wie aus Abb. 17 ersichtlich wird.

Neben diesen beiden Dimensionen wurde in Clasivir eine dritte Aufgabenvariable hinzugefügt, nämlich die der spezifischen Schwierigkeit. Die spezifische Schwierigkeit modelliert den Umstand, dass bestimmte Aufgaben des Types *remember/factual knowledge* sehr einfach sein können („Welchen Wochentag haben wir heute?“), auf der anderen Seite aber auch sehr schwer sein können („In welchem Jahr wurde die Kaiserpfalz in Gelnhausen von Friedrich, I. gegründet?“). Um das Modell nicht überkomplex zu machen, werden drei ordinal skalierte Werte für die spezifische Aufgabenschwierigkeit modelliert: ‚einfach‘, ‚mittel‘ und ‚schwer‘. Der Algorithmus von Clasivir berechnet jetzt ausgehend von diesen drei Aufgabenvariablen eine Aufgabenschwierigkeit (siehe Tabelle 19, S. 273). Die exakten Werte orientieren sich zum einen an der theoretischen ordinalen Ordnung. Zum anderen wurde die Simulation mehrfach durchgespielt, um exakte Werte zu bestimmen.

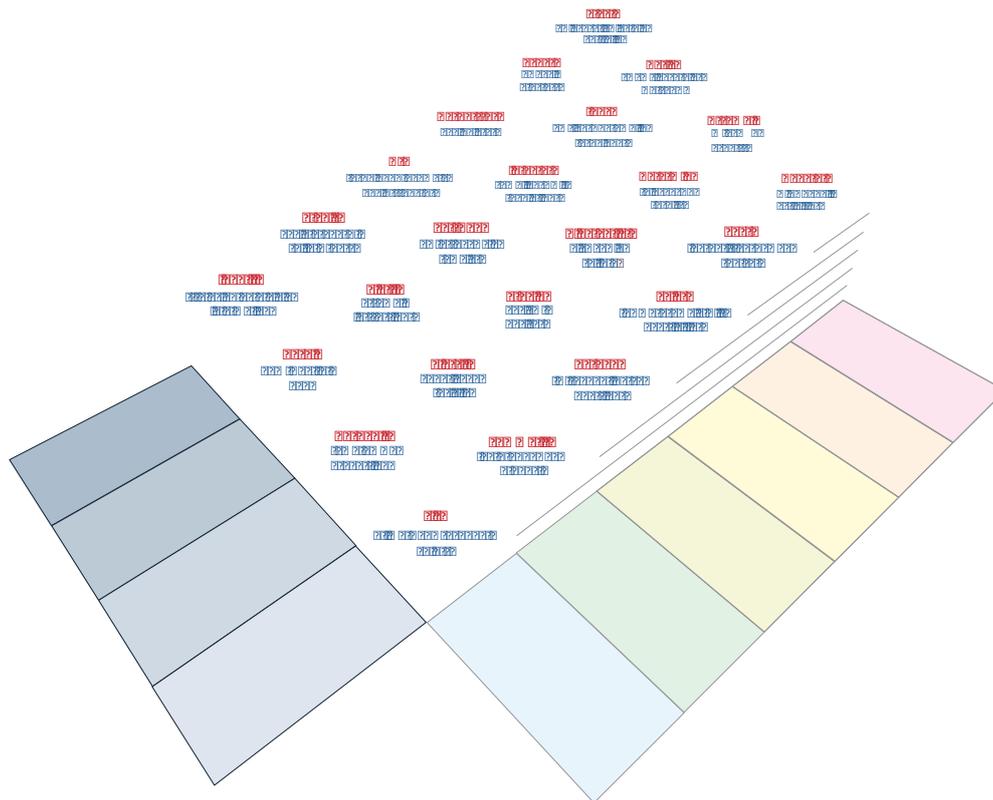


Abbildung 17: Graphische Darstellung der Anderson/Krathwohl'schen Taxonomie. Originaltitel: A Model of Learning Objectives-based on A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, von Rex Heer, Center for Excellence in Learning and Teaching, Iowa State University. Lizenziert unter Creative Commons Attribution-ShareAlike 4.0 International License

Die Aufgabenschwierigkeit wird im Folgenden mit der gewichteten kognitiven Fähigkeit der jeweiligen SuS verglichen.

- (a) Liegt sie unter 50% der kognitiven Fähigkeit wird die Aufgabe als zu einfach kategorisiert.
- (b) Liegt sie über 50%, aber unter 90% der kognitiven Fähigkeit wird die Aufgabe als normal lösbar kategorisiert.
- (c) Liegt sie über 90%, aber unter 100% der kognitiven Fähigkeit wird die Aufgabe als optimal lösbar kategorisiert.
- (d) Liegt die Aufgabe über 100% der kognitiven Fähigkeit wird die Aufgabe als nicht lösbar kategorisiert.

Dieser Vorgang wird in dieser Arbeit mit dem Begriff der adaptiven Aufgabenpassung beschrieben.

Die kognitive Fähigkeit wurde im Rahmen der Simulation mit einem konstanten Faktor multipliziert (aktuell 1.22) und anschließend wieder auf ein Intervall von 0 bis 1 komprimiert. Dieser Wert wird als kognitiver Schwellenwert bezeichnet. Dies wurde nachträglich durchgeführt, um eine bessere Passung der errechneten Schwierigkeitsgrade zu erreichen. Ohne diese Anpassung würden bestimmte höchst komplexe Aufgaben, wie z.B. *create/metakognitive knowledge* selbst bei kognitiv höchst fähigen Schülern so gut wie immer über der Schwelle ihrer kognitiven Fähigkeit liegen und wären somit unlösbar. Mit der Anpassung können zwar nur die fähigsten SuS diesen Aufgabentyp lösen, aber die Aufgabe ist nicht generell unlösbar. Eine Beispieltabelle für eine durchschnittliche Schülerin mit einer kognitiven Fähigkeit von .5 ist im Anhang in [Tabelle 20](#) (S. 274) zu finden.

5.11.2 Verarbeitungsalgorithmus

Nachdem eine Aufgabe gestellt und in eine adaptive Lösbarkeitskategorie für alle SuS eingeteilt wurde, berechnet ein Algorithmus für alle SuS, ob sich diese zu einer kognitiven Verarbeitung der Aufgabe entschließen. Die Chance dafür ist abhängig von der adaptiven Aufgabenpassung, der Konzentration und der gewichteten OCEAN-Werte, die einen Proxy für die Motivation darstellen. Die Konzentration wird mit den gewichteten OCEAN-Werten folgendermaßen verrechnet, um die Motivation im Intervall zu bestimmen:

Die Rationale für dieses Vorgehen ist, dass die vorhandene Motivation in erster Linie situativ ist, also am stärksten davon abhängt wie konzentriert die jeweilige Schülerin der Lehrkraft aktuell folgt. Weiterhin ist entscheidend, wie offen die Schülerin für die Verarbeitung neuer Eindrücke ist, was durch die Openness-Variable modelliert wird. Die Inklusion der Extraversion ist erklärungsbedürftig: Es wird davon ausgegangen, dass die Extraversion andeutet, wie begeisterungsfähig eine Schülerin für interpersonelle Kommunikation ist, die bei der Aufgabenstellung einen wichtigen Faktor darstellt. Nachdem die Motivation im Intervall bestimmt wurde, wird eine Chance zur Verarbeitung der Aufgabe berechnet. Diese ist abhängig von der adaptiven Aufgabenpassung:

Vearbeitung, zu einfach	_____
Vearbeitung, normal	_____
Vearbeitung, optimal	_____
Vearbeitung, nicht lösbar	_____

Da die mentale Anstrengung, die für eine Aufgabe aufgebracht werden muss, von deren Schwierigkeit abhängt (Goldhammer & Moosbrugger, 2006, S.19) , ist die vorgeschlagene Operationalisierung einer kausalen Abhängigkeit von Aufgabenschwierigkeit und Konzentrationsvariable plausibel.

Wurde einer Schülerin bereits eine Aufgabe zugewiesen, während diese noch eine andere Aufgabe verarbeitet, so wird aufgrund von *cognitive load* (Sweller, 1988; Sweller, 2011) und ungünstiger Unterrichtsführung ein kleiner Wert (aktuell 0.1) von der Konzentration der Schülerin abgezogen. Die aktive Aufgabe der Schülerin wird danach auf die neu gestellte Aufgabe festgelegt. Will man von dieser Schülerin eine Antwort zur vorherigen Fragestellung, so muss die Aufgabe erneut gestellt werden. Diese Modellierung setzt Anreize SuS nicht mit zu vielen Aufgaben in zu kurzer Zeit zu überfordern.

5.11.3 *functional off-task*

Wenn sich eine Schülerin zur Verarbeitung einer Aufgabe entscheidet, kann es sein, dass sie *functional off-task* (Helmke et al., 1986) zeigt, also passives Störungsverhalten, obwohl die interne kognitive Verarbeitung läuft. Dies trägt zur Modellierung des Umstands bei, dass SuS durchaus motorische Unruhe zeigen oder aus dem Fenster blicken können, obwohl sie über die gestellte

Aufgabe nachdenken. Diese Modellierung soll Partizipantinnen dazu anhalten, darüber zu reflektieren, was Störverhalten auszeichnet und dass es nicht immer negativ zu bewerten ist oder einer Intervention bedarf.

Es ist für die Simulationspartizipantinnen unklar, ob ein Störverhalten wie motorische Unruhe entstanden ist, weil die Konzentration unter den Schwellenwert gefallen ist und sich aus Langeweile Störverhalten manifestiert oder ob es sich beim passiven Störverhalten lediglich um *functional off-task* handelt.

Das Zeigen von *functional off-task* wird randomisiert vergeben (siehe [Tabelle 22](#), S. 275).

5.11.4 Verarbeitungsalgorithmus

Entscheidet sich die jeweilige Schülerin nicht zur Verarbeitung ist die potentielle Antwort auf die Aufgabe immer falsch, denn es wurde ja nicht über die Aufgabe reflektiert. Entscheidet sich die Schülerin zur Verarbeitung, so benötigt sie eine gewisse Zeit, um zu einem Ergebnis zu kommen. Diese Verarbeitungszeit wird in Intervallen angegeben und ist abhängig von der adaptiven Aufgabenpassung:

- 1 Intervall, falls die Aufgabe zu einfach ist
- 2-3 Intervalle, falls die Aufgabe normal lösbar ist.
- 3-5 Intervalle, falls die Aufgabe optimal lösbar ist.
- 2-6 Intervalle, falls die Aufgabe nicht lösbar ist.

Wird eine Schülerin bereits während der Verarbeitungszeit aufgerufen, so kann sie noch keine Antwort geben und es wird eine Konstante von ihrer Konzentration abgezogen. Dies setzt für Simulationspartizipantinnen den Anreiz die auch in der Realität vorhandenen Verarbeitungszeiten einzuhalten. Die Länge der Verarbeitungszeiten ist abgeleitet von den Untersuchungen von Sacher (1995, S.102), Lipowsky et al. (2007, S.127) und Borich (2014, S.142).

5.11.5 Lösungsalgorithmus

Nach der Wartezeit greift ein weiterer Algorithmus, der die Chance bestimmt zu der die Aufgabe korrekt gelöst wird:

$$Ch_{\text{korrekt}} = \frac{1 - 1.5^{\frac{1}{0.087} \cdot C} + Gw}{2}$$

wobei $Gw = \text{Grundwert}$

Der erste Faktor beschreibt dabei mathematisch ein exponentielles Abklingen der Wahrscheinlichkeit für die korrekte Lösung in Abhängigkeit von der Gewissenhaftigkeit der SuS (siehe [Abbildung 18](#)). Theoretische Rationale für den kausalen Zusammenhang ist, dass die Gewissenhaftigkeit der SuS darüber entscheidet, ob sie die Aufgabe nur flüchtig durchführen oder genau durchdenken. Die mathematische Funktion wurde gewählt, da Gewissenhaftigkeits-Werte über 0.5 nur kleinen Einfluss auf die korrekte Lösungswahrscheinlichkeit nehmen. Bei Personen, die eine signifikante negative Gewissenhaftigkeit aufweisen, bestimmt dieser Faktor hingegen stärker die Wahrscheinlichkeit einer korrekten Lösung. Der zweite Faktor, der Grundwert, bestimmt zur Hälfte die korrekte Lösungswahrscheinlichkeit.

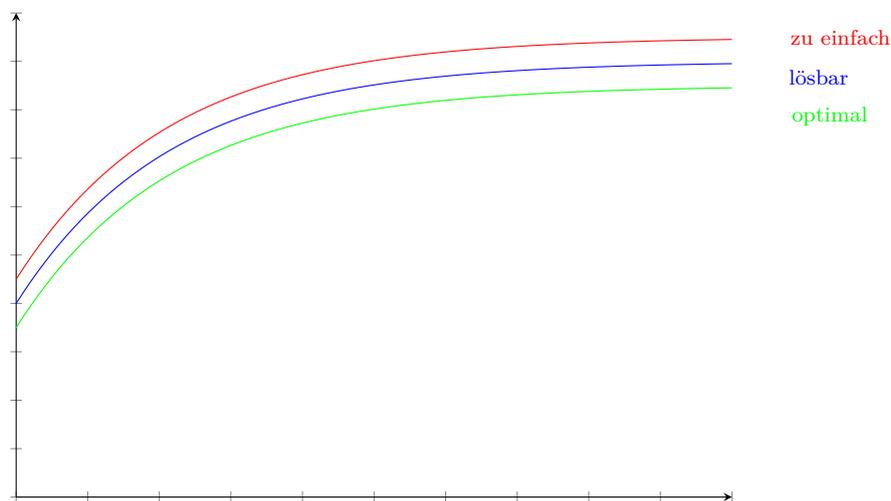


Abbildung 18: Exponentielles Abklingen im Lösungsalgorithmus

Dieser Grundwert ist von der adaptiven Aufgabenpassung abhängig:

- zu einfach
- normal lösbar
- optimal lösbar

Diese Werte wurden durch Durchspielen der Simulation gewonnen, bis die Ergebnisse im Durchschnitt die postulierten Werte von Sacher ([1995](#)) abbildeten.

5.11.6 Hereinrufen-Algorithmus

Ist die Verarbeitungszeit einer Schülerin abgelaufen, ermittelt ein Algorithmus, ob die Schülerin die Antwort in die Klasse hereinruft ohne aufgerufen worden zu sein. Die Chance dafür ist abhängig von der adaptiven Aufgabenpassung und den OCEAN-Werten für Extrovertiertheit und umgekehrter Verträglichkeit.

$$\frac{+(1 \quad)}{2}$$

wobei:

- adptiv, zu einfach
- adptiv, normal lösbar
- adptiv, optimal lösbar
- adptiv, nicht lösbar

Damit steigt die Chance für das Hereinrufen bei extrovertierten und wenig verträglichen SuS. Die Chance ist aufgrund der Divisor-Werte gering und steigt nur signifikant an, wenn die Aufgabe zu einfach ist. Sacher (1995, S.105) ermittelte, dass 9.68% aller von ihm untersuchten SuS-Äußerungen ohne Autorisierung erfolgten. Dabei ist zu beachten, dass viele von diesen gezählten Äußerungen ein abkürzendes Verfahren im Unterricht darstellten, wenn sich ohnehin niemand meldete. Dieses Verfahren wird in Clasivir nicht modelliert. Es ist davon auszugehen, dass 1–3% der Wortmeldungen ohne Autorisierung kein abkürzendes Verfahren darstellen und damit ein sanktionierbares Störverhalten abbilden. Ruft eine Schülerin daher eine Antwort herein, so wird dies von Clasivir als Störverhalten gewertet (im Code als SV_DIRECTANSWER bezeichnet). Dieses Störungsverhalten kann nur unter den hier beschriebenen Bedingungen auftreten und nicht in der allgemeinen Auswahl der Störungsverhalten, wenn die Konzentration unter ein bestimmtes Niveau fällt. Das Störungsverhalten der unautorisierten Antwort ist Pfitzner und Schoppek (2000, S.375ff.) entnommen.

Meldet sich eine Schülerin unautorisiert, so ist die Aufgabe selbst bei korrekter Antwort noch nicht beendet. Die Simulationspartizipantinnen können die Schülerin mit unautorisierter Meldung entweder zur Antwort autorisieren oder eine weitere Schülerin aufrufen. Die Antwort einer weiteren aufgerufenen Schülerin ist nicht immer automatisch richtig, wenn die richtige Antwort schon hereingerufen wurde. Die SuS wissen zu diesem Zeitpunkt nämlich nicht, ob die hereingerufene Antwort korrekt war. Eine Beispieltabelle der Hereinrufenswahrscheinlichkeiten für eine spezifische Schülerin ist im Anhang in [Tabelle 23](#) (S. 276) zu finden.

5.11.7 Meldealgorithmus

Ist der Verarbeitungsprozess zu Ende und wurde eine korrekte oder inkorrekte Antwort generiert, bestimmt der Meldealgorithmus, ob sich SuS melden. Die Wahrscheinlichkeit dafür ist abhängig von (a) allen OCEAN-Variablen, besonders stark von der Extraversion des SuS und (b) von der adaptiven Aufgabenpassung. Introvertierte Schüler melden sich daher seltener. Zudem ist die Meldewahrscheinlichkeit bei schwierigeren Aufgaben niedriger. Folgende Formel wird zur Bestimmung der Meldewahrscheinlichkeit verwendet:

$$\text{Meldung} = \frac{+ +7 + (1 \quad) + (1 \quad)}{20}$$

wobei:

- adptiv, zu einfach
- adptiv, normal lösbar
- adptiv, optimal lösbar
- adptiv, nicht lösbar

Die Meldewahrscheinlichkeit ändert sich nicht in Abhängigkeit davon, ob die Aufgabe korrekt gelöst wurde. Dies ist Folgendermaßen konzeptionell zu begründen: Denkt eine Schülerin über die Aufgabe nach und kommt zu einem Ergebnis, so geht sie davon aus, dass dieses Ergebnis richtig ist. Zögerliche Meldungen bei einer schwereren Aufgabe werden über die Inkorporation der adaptiven Aufgabenpassung realisiert. Eine Beispieltabelle der Meldewahrscheinlichkeiten für eine spezifische Schülerin ist im Anhang in [Tabelle 24](#) (S. 277) zu finden.

Der Algorithmus wurde so implementiert, dass näherungsweise der empirisch von Böheim et al. (2020) gewonnene Schulstundenmittelwert für Meldungen durch das Simulationsmodell reproduziert wird. Im arithmetischen Mittel melden sich SuS fünf Mal in einer Schulstunde (SD=5.5). Entscheidet sich eine Schülerin zur Meldung, bestimmt ein weiterer Algorithmus die Art der Meldung. Es gibt in Clasivir drei Arten der Meldung:

- Undeutliche Meldung
- Aufdringliche Meldung
- Normale Meldung

Diese drei Arten der Meldung sind Sacher (1995, S.88) entlehnt. Die Wahrscheinlichkeiten sind aufgrund der unterschiedlichen Voraussetzungen für die Ar-

ten der Meldungen etwas komplexer:

	Aufdringliche Meldung	Normale Meldung
Undeutliche Meldung		
	$\frac{2(1 \quad)+}{3}$	Normale Meldung
Aufdringliche Meldung	Undeutliche Meldung	
	Aufdringliche Meldung	Undeutliche Meldung
Normale Meldung		

Auf die Implementierung von gegenseitigen Aufrufen der SuS, z.B. in einer Diskussion (Sacher, 1995, S.129), wird verzichtet, da es die Komplexität exponentiell steigert und zudem selten vorkommt.

Eine Tabelle der Wahrscheinlichkeiten für die verschiedenen Meldungstypen ist im Anhang unter [Tabelle 25](#) (S. 277) zu finden.

5.11.8 Meldungsabbruchalgorithmus

Melden sich SuS eine lange Zeit, ohne dass sie von der Lehrkraft beachtet werden, brechen diese zum Teil ihre Meldung ab. Es ist konzeptionell sinnvoll, die Wahrscheinlichkeit mit der Anzahl der Intervalle ohne Aufruf steigen zu lassen. Zudem soll die Wahrscheinlichkeit, dass der Meldeabbruch schon in den ersten Intervallen stattfindet sehr gering sein, da dies in der Realität eher selten vorkommt. Um das Modell nicht unnötig komplex zu machen, wurde ein einfacher linearer Zusammenhang modelliert: Mit jedem Intervall steigt die Wahrscheinlichkeit, dass ein Meldungsabbruch stattfindet. Im ersten Meldungsintervall beträgt die Wahrscheinlichkeit 0.02 und steigt um 0.01 bis der Wert nach 8 Intervallen einen Maximalwert von 0.1 erreicht. Die Wahrscheinlichkeit, dass nach 8 Intervallen ein Meldungsabbruch stattfindet oder in den vorherigen Intervallen bereits stattgefunden hat beträgt somit 36.5%. Die Meldungsabbruchswahrscheinlichkeit in 3 Intervallen beträgt nur 8.7%, Meldungsabbrüche in den ersten Intervallen nach der Meldung sind also sehr selten, aber nicht ausgeschlossen. Nach 11 Intervallen beträgt die Meldungsabbruchswahrscheinlichkeit 53.7%.

5.12 Situative Variablen: Lehrkraftaktionen

Die Simulationsteilnehmerinnen können über eine Reihe von Handlungen mit den virtuellen SuS interagieren. Diese werden an dieser Stelle nacheinander beschrieben.

5.12.1 Aufgaben/Fragen

Aufgaben sind ein Kernmerkmal von unterrichtlichem Handeln und gehören damit zu den wichtigsten Funktionen in Clasivir. Die Partizipantinnen können wie bereits in [Abschnitt 5.11](#) (S. 175) aufgeführt eine Aufgabe mit einer kognitiven Dimension, einer Wissensdimension und einer spezifischen Aufgabenschwierigkeit wählen. In der geplanten VR-Version von Clasivir sind diese Aufgabenstellungen durch Szenarienerstellerinnen (z.B. Fachdidaktikerinnen) inhaltlich ausformuliert und vorkodiert. Zusätzlich kann festgelegt werden, wer das Ziel der Aufgabe ist. Dies kann entweder die gesamte Klasse, eine einzelne Schülerin oder eine Gruppe von SuS sein. Clasivir ist so konzipiert, dass ohne Probleme mehrere Aufgaben adaptiv vergeben und durch verschiedene SuS parallel bearbeitet werden können. Schülerin 1 kann also an einem spezifischen für sie geeigneten Problem arbeiten, während der Rest der Klasse eine andere Aufgabe bearbeitet.

5.12.2 Informationen geben

Bei dieser Handlung hält die Lehrkraft einen Lehrkraftvortrag zu einem fachlichen Thema. Die Aufgabentaxonomie von Anderson et al. (2000) ist auf Lehrkraftvorträge logisch nicht anwendbar, so dass diese Komponente sehr vereinfacht über einen Slider zur Bestimmung der Schwierigkeit des Vortrags modelliert wird. Dabei steht der Wert 0 für eine extrem einfache Information, die fast keine kognitiven Ressourcen zur Verarbeitung benötigt, während der Wert 1.0 einen Vortrag bezeichnet, der für den jeweils modellierten Jahrgangstyp nur von den kognitiv fähigsten SuS zu lösen ist. Der Wert der Informationsschwierigkeit wird mit der kognitiven Fähigkeit bzw. der Motivation im Intervall der SuS verglichen und den Stati ‚Information zu einfach‘, ‚Information normal verarbeitbar‘, ‚Informationsschwierigkeit optimal‘ und ‚Informationsschwierigkeit zu schwer‘, analog zur Aufgabenmodellierung, zugeordnet. Die Informationswerte werden intern auf den Wertebereich 0.1–0.8 komprimiert. Würde man das nicht tun, könnten selbst die kognitiv fähigsten SuS einer sehr komplexen Information fast nie folgen.

Daneben kann analog zur Aufgaben/Fragestellung angegeben werden, wer das Ziel des Vortrags ist. Letztlich ist definierbar, wie viele Intervalle der Lehrkraftvortrag dauert. Entscheidet sich eine Schülerin zur Verarbeitung der Informationen bekommt sie einen geringen Bonus auf ihre Konzentration und dieser Konzentrationswert wird für die Dauer der Information eingefroren. Entscheidet sich eine Schülerin gegen die Verarbeitung fällt der Konzentrationswert entsprechend biexponentiell ab und es kann sich Störverhalten manifestieren. Aus der Konstruktion der Informationskomponente ist ersichtlich,

dass sie eher als Mikroaktion konzipiert wurde, also ein ganzer Lehrkraftvortrag in einzelne Informationsabschnitte einzuteilen ist. Eine Information mit einer Länge von z.B. 60 Intervallen würde unrealistisches SuS-Verhalten provozieren.

Ein *functional off-task* für informationsverarbeitende SuS ist bislang nicht implementiert.

Ähnlich wie beim *classroom management* (siehe [Abschnitt 5.12.3](#), S. 187) existiert eine Variable, die registriert, wann die letzte Information gegeben wurde. Wird innerhalb von 5 Intervallen seit der letzten Information eine zusätzliche Information gegeben, so wird allen adressierten SuS ein Konzentrationswert von 0.1 abgezogen. Dies unterbindet *gaming-the-system*-Verhalten (Baker et al., 2006, siehe auch [Unterunterabschnitt 5.1.3](#), S. 151).

Simulationspartizipantinnen sind dazu angehalten, für die Intervalle während der sie ‚Informationen geben‘ keine anderen Handlungen zu initiieren. Tun sie es trotzdem und die folgende Handlung ist eine Aufgabe oder eine Information, so erhalten die adressierten SuS einen Konzentrationsmalus von 0.1. Dieser Wert ist äquivalent zum Abzug der stattfindet, wenn SuS zusätzliche Aufgaben gestellt werden, während sie sich noch in der Verarbeitungszeit befinden. Wird eine *classroom-management*-Handlung während der Informationszeit initiiert, so wird der Konzentrationswert der SuS nicht verändert. Diese Modellierung ermöglicht es, SuS während Informationsanteilen zu ermahnen, ohne dass andere SuS dafür unter Konzentrationseinbußen leiden. Es ist jedoch unklar, wie gut dieser Ansatz realen Informationsunterbrechungen gerecht wird.

5.12.3 *classroom management*

Die Simulationspartizipantinnen haben die Möglichkeit verschiedene *classroom-management*-bezogene Handlungen zu initiieren. Diese bestehen zum einen aus einem vierstufigen Handlungssystem, dass von der Logik des Feedbacks (siehe [Abschnitt 5.12.5](#), S. 190) abgeleitet ist:

- Loben
- Im Verhalten bestärken
- Konstruktiv Ermahnen
- Destruktiv Ermahnen

Ziel dieser Handlungen kann eine einzelne Schülerin, eine Gruppe oder die gesamte Klasse sein. Sie beziehen sich immer auf die An- oder Abwesenheit von Störungsverhalten. Wird eine Schülerin, die Störungsverhalten zeigt,

konstruktiv oder destruktiv ermahnt, so wird das aktuelle Störungsverhalten beendet. Während konstruktives Ermahnen bei Störverhalten einen Konzentrationsbonus auslöst, ist dies bei destruktiven Ermahnungen nicht der Fall. Das Störverhalten wird zwar beendet, aber die Konzentration fällt weiter. Wird eine Schülerin gelobt oder im Verhalten bestärkt, so wird die Konzentration leicht gesteigert, wenn kein Störungsverhalten vorliegt. Andererseits wird die Konzentration gesenkt, wenn Störungsverhalten vorliegt. Wird die gesamte Klasse als Ziel festgelegt, so sind die Konzentrationsboni weniger stark ausgeprägt und die Konzentrationsmali stärker negativ. Dies soll die Partizipantinnen dazu incentivieren eher spezifische SuS anstelle der gesamten Klasse zu adressieren. Die exakten Werte sind im Anhang in [Tabelle 26](#) (S. 278) nachzulesen.

Der aufmerksamen Leserin wird auffallen, dass dieses System zu *gaming-the-system*-Verhalten (Baker et al., 2006, siehe auch [Abschnitt 5.1.3](#), S. 151) einladen kann. Wäre das System exakt so implementiert, würde nichts dagegen sprechen, die ersten 15–20 Intervalle keine andere Aktion durchzuführen als die gesamte Klasse ständig im Verhalten zu bestärken oder zu loben. Da in den ersten Intervallen für gewöhnlich noch kein Störverhalten auftritt, könnte man so den Wert für Konzentration künstlich auf dem Maximalwert von 1.0 halten. Natürlich müsste man dafür die Funktionsweise des Clasivir-Simulationsmodells kennen, aber der hier vorgestellte Zusammenhang ist recht schnell intuitiv herauszufinden, wenn nicht im ersten Simulationsdurchlauf, dann zumindest, wenn man Clasivir mehrfach benutzt hat.

Um diese Manipulation zu unterbinden, wurde eine weitere Variable eingeführt, die registriert, wann die letzte *classroom-management*-Handlung durchgeführt worden ist. Wird innerhalb von 5 Intervallen eine weitere *classroom-management*-Handlung durchgeführt, so sinkt die Konzentration der gesamten Klasse um 0.04. Danach muss wieder 5 Intervalle gewartet werden, bis eine valide *classroom-management*-Handlung erneut durchgeführt werden kann.

Punitive Handlungen (Strafarbeiten schreiben, Drohen) werden in Clasivir vorrangig über die Kategorie ‚Destruktives Ermahnen‘ abgebildet. Es wurde sich bewusst dafür entschieden punitive Handlungen nicht ausdifferenziert zu modellieren, da sich die meisten Expertinnen einig sind, dass diese weder zur kurzfristigen noch zur langfristigen Vermeidung von Störverhalten geeignet sind (Siestrup, 2010, S.67). Die einzigen explizit modellierten punitiven Handlungen sind:

- Verweis einer Schülerin zum Direktor
- Verweis einer Schülerin in eine Parallelklasse.

Diese Extremmaßnahmen sind deswegen sinnvoll, da sie als ultima ratio bei aggressivsten körperlichen Störverhalten dienen können. Die Reflexion darüber, wann diese Maßnahmen angebracht sind und das Abwägen zwischen den verschiedenen *classroom-management*-Handlungen ist ein wichtiger Bestandteil der BAR- und AAR-Feedbackkomponenten. Wird eine solche *classroom-management*-Handlung durchgeführt, wird die Anwesenheitsvariable der Zielschülerin auf 0 geändert und sämtliche situativen Werte auf den Stand des Intervalls 0 gesetzt. Da nicht geplant ist, Clasvir-Szenarien länger als 10 bis 15 Minuten zu simulieren, ist aktuell kein Rückkehralgorithmus der SuS in die Klasse implementiert. Bei zukünftigem Bedarf wäre dies aber einfach umzusetzen.

Die veränderten Werte der *classroom-management*-Handlungen in Abhängigkeit vom gewählten Ziel (Einzelziele, Gruppen oder global über die Klasse) orientieren sich an den von Dawson (2016, S.14f.) berichteten Ergebnissen, wonach an Einzelpersonen adressiertes Lob oder konstruktive Kritik einen stärkeren Einfluss auf die SuS hat.

Da geplant ist, dass Simulationspartizipantinnen die *classroom-management*-Aktionen in der VR-Simulation immer einsetzen können, ist es notwendig, das abstrakte, vierstufige System in konkrete Handlungsalternativen zu überführen. Ansonsten ist es unwahrscheinlich, dass sich Simulationspartizipantinnen dafür entscheiden würden eine ‚destruktive Ermahnung‘ durchzuführen, zumindest wenn ihr Ziel ist die Simulation erfolgreich zu bewältigen. Dies erfordert eine enge Abstimmung mit Feldexpertinnen. Dies könnte sich deswegen als schwierig herausstellen, weil beispielsweise die Vergabe einer Strafarbeit bei einigen Störverhalten angemessen ist, bei anderen aber nicht. Im ersten Falle wäre die Handlung als ‚konstruktive Kritik‘ zu klassifizieren, im zweiten Falle als ‚destruktive Kritik‘.

5.12.4 SuS aufrufen

Die Simulationspartizipantinnen können zu jeder Zeit SuS zur Beantwortung einer Aufgabe aufrufen. Das Ziel für diese Aktion muss ein Einzelziel sein. Existiert keine valide Frage, so erleidet die adressierte Schülerin einen Konzentrationsmalus von 0.1. Existiert eine valide Frage, erfahren die Simulationspartizipantinnen, ob die Schülerin richtig oder falsch geantwortet hat. Es existieren im Quellcode einige Bausteine zur Implementierung partiell richtiger Antworten, jedoch würde eine vollständige Integration die Komplexität des Simulationsmodells stark erhöhen, weswegen darauf verzichtet wurde. Ein Instrument zur Analyse von Lehrkraftfeedback, das von Borich (2014, S.206, S.290) entwickelt wurde, teilt SuS-Antworten in die Kategorien „korrekt“, „partiell korrekt“ und „falsch“ ein.

Gibt eine Schülerin die korrekte Antwort, so wird die Aufgabe abgeschlossen und die nächste Fragestellung kann ohne negative Konsequenzen folgen. Wird die falsche Antwort gegeben, so wird der Task so lange weitergeführt, bis er korrekt beantwortet wurde oder die Lehrkraft die Aufgabe ohne Beantwortung abbricht. Diese Interaktionslogik ist den Analysen von Borich (2014, S.144) entlehnt.

Das Beantworten einer validen Aufgabe allein modifiziert nicht den Konzentrationsstatus der entsprechenden Schülerin. Erst durch das Feedback der Lehrkraft wird dieser Effekt ausgelöst.

5.12.5 Feedback auf Antwort geben

Die Lehrkraft kann zu jeder Zeit Feedback auf eine Antwort der SuS geben. Dies folgt der Logik der IRE-Kaskaden (Mehan, 1979).

Das Ziel dieser Handlung muss immer ein Einzelziel sein. Hat eine Schülerin bislang noch keine Frage beantwortet, so ist diese über das Lehrkraftfeedback verwirrt und verliert 0.1 ihres Konzentrationswerts. Hat die Schülerin eine Frage beantwortet, so wird nach dem Feedback ein bestimmter Wert von der Konzentration der Zielschülerin abgezogen oder hinzugefügt, abhängig von dem Typ des Feedbacks und davon, ob eine korrekte Antwort gegeben wurde. Eine Tabelle zu diesen Werten ist im Anhang unter [Tabelle 27](#) (S. 278) zu finden. Die vorhandenen Typen von Feedback sind abgeleitet von J. Brophy (1981) und Borich (2014, S.204, S.206, S.290). Dass *affirm*-Feedback gegenüber *praise*-Feedback zu bevorzugen ist, ist von Humphreys (1995) abgeleitet.

6 Methodische Anlage der Untersuchung

Ziel des empirischen Teils dieser Arbeit ist es das in den vorherigen Kapiteln beschriebene, neu entwickelte Simulationsmodell quasi-experimentell zu evaluieren. Grundsätzlich liegt der Beforschung die Annahme zugrunde, dass es wichtig ist, dass die Zielgruppe, nämlich Studieninteressierte und Lehramtsstudierende, das Simulationsmodell für realistisch hält. Außerdem sollen mittels einer qualitativen Analyse von Freitextantworten Spezifika in der Simulationsmodellfidelitätsbewertung dieser Zielgruppe erschlossen werden.

Sowohl quantitative als auch qualitative Daten der Untersuchung speisen sich aus einem Online-Fragebogen, den die Partizipantinnen ausgefüllt haben. In diesem Online-Fragebogen sind graphisch aufbereitete Videos von drei unterschiedlichen Simulationsmodellen (real, Clasivir, randomisiert) vorhanden, die miteinander verglichen werden sollten. Aufgrund Corona-beding-

ter Schwierigkeiten in der Akquirierung von Partizipantinnen wurden alle Untersuchungsgruppen, die ein randomisiertes Modell beobachten sollten, in andere Gruppen eingeteilt. Im Fragebogen wurden neben simulationsfidelitätsbezogenen Variablen demographische Daten und Daten zur Technikaffinität, VR- und Computerspielnutzung und Einstellungen zu digitalen Medien erhoben. Weitere Informationen zu den genutzten Instrumenten finden Sie in Kapitel [Abschnitt 6.2.2](#) (S. 197).

6.1 Forschungsdesign

Die Validierung des (pädagogischen) Simulationsmodells von Clasivir ist ein komplexes Unterfangen, da keine erprobten Instrumente existieren um die wahrgenommene Fidelität eines pädagogischen Simulationsmodells (siehe [Abschnitt 2.5.12](#), S. 38) zu erfassen. Es gilt zudem zu beachten, dass die methodische Anlage – etwa die Anwendung von Mockup/Videos – generell auf die Validierung anderer Simulationsmodelle übertragbar, aber die konkrete Ausgestaltung des Instruments nicht generalisierbar ist. Im Falle von Clasivir fokussiert das Instrument die wahrgenommene Simulationsmodellfidelität unter anderem durch Analyse von Störungs/, Melde/ und Antwortverhalten. In einer Sekundärnutzung des Instruments würde es aber schon bei der Anwendung in einer Simulation für Gruppenarbeiten kaum mehr anwendbar sein und bei Situationen wie Elterngesprächen vollständig an die Grenzen stoßen. Es ist wünschenswert in zukünftigen Forschungsarbeiten im Sinne der Studienvergleichbarkeit einen Kriterienkatalog zu entwickeln, der in abstrakter Form Konstrukte vorschlägt, wie man die Fidelitätsaspekte einer Simulation gewinnbringend evaluieren kann.

Eine Herausforderung im konkreten Fall dieser Arbeit ist, dass nicht eine bereits existierende Simulation evaluiert werden soll, sondern nur das Simulationsmodell. Bei Vorliegen einer fertigen Simulation könnte auf Instrumente wie das USE Questionnaire (Lund, 2001) zurückgegriffen werden, auch wenn diese primär auf Usability Aspekte der Simulation und nicht auf das Simulationsmodell fokussieren. Daher musste in dieser Arbeit ein eigenes Instrument entwickelt werden um die wahrgenommene Fidelität des Simulationsmodells zu bestimmen. Zusätzlich wurde überprüft ob sich die Befragten latent in ihrer Fidelitätswahrnehmung unterscheiden.

6.1.1 Hypothesen

- $H_{A,0}$: Es zeigen sich keinerlei Unterschiede in der Bewertung des Clasivir/Simulationsmodells und des realen Simulationsmodells durch die Partizipantinnen. Es ist zu beachten, dass die Nullhypothese hier der Forschungshypothese entspricht.
 - $H_{A,1}$: Partizipantinnen bewerten das Clasivir-Simulationsmodell unterschiedlich zum realen Simulationsmodell.
- $H_{B,1}$: Partizipantinnen bewerten das Clasivir-Simulationsmodell besser als das randomisierte Simulationsmodell.
 - $H_{B,0}$: Partizipantinnen bewerten das Clasivir-Simulationsmodell schlechter oder gleich gut wie das randomisierte Simulationsmodell.
- $H_{C,1}$: Partizipantinnen mit hohem technologischem Wissen (TPACK-Komponente) schätzen den Schwierigkeitsgrad der Bewertung geringer ein.
 - $H_{C,0}$: Partizipantinnen mit hohem technologischem Wissen schätzen den Schwierigkeitsgrad der Bewertung gleich hoch oder höher ein.
- $H_{D,1}$: Personen, die in eine Gruppe mit 3D-Darstellungsform zugeteilt waren, berichten über einen niedrigeren Schwierigkeitsgrad der Bewertung als Personen, die der Gruppe der 2D-Darstellungsform zugeteilt waren.
 - $H_{D,0}$: Personen, die in eine Gruppe mit 3D-Darstellungsform zugeteilt waren, berichten über einen gleich hohen oder höheren Schwierigkeitsgrad der Bewertung als Personen, die der Gruppe der 2D-Darstellungsform zugeteilt waren.
- $H_{E,1}$: Personen, die den Computer für ein nützliches Werkzeug halten (Subskala FIDEC), bewerten die Nützlichkeit des Einsatzes von Simulationen in der Lehrkräftebildung höher.
 - $H_{E,0}$: Personen, die den Computer für ein nützliches Werkzeug halten, bewerten die Nützlichkeit des Einsatzes von Simulationen in der Lehrkräftebildung schlechter oder gleich gut als Personen, die den Computer nicht für ein nützliches Werkzeug halten.

- $H_{F,1}$: Die bewertete Fidelität des gleichen Simulationsmodells unterscheidet sich nicht vom ersten zum zweiten präsentierten Schulstundenausschnitt.
 - $H_{F,0}$: Die bewertete Fidelität des gleichen Simulationsmodells unterscheidet sich vom ersten zum zweiten präsentierten Schulstundenausschnitt.
- $H_{G,1}$: Personen, die häufig Videospiele auf einem Computer oder auf einer Konsole spielen, bewerten den Schwierigkeitsgrad in der Bewertung des Simulationsmodells geringer.
 - $H_{G,0}$: Personen, die häufig Videospiele auf einem Computer oder einer Konsole spielen, bewerten den Schwierigkeitsgrad in der Bewertung des Simulationsmodells höher oder gleich hoch wie Personen, die nicht häufig Videospiele auf dem Computer spielen.
- $H_{H,1}$: Männliche Partizipanten bewerten die Nützlichkeit von Simulationen in der Lehrkräftebildung positiver als weibliche Partizipantinnen.
 - $H_{H,0}$: Männliche Partizipanten bewerten die Nützlichkeit von Simulationen in der Lehrkräftebildung schlechter oder gleichwertig zu weiblichen Partizipantinnen.

Wird sowohl $H_{A,0}$ (als auch $H_{B,1}$) akzeptiert, kann eine Anwendung des Clasivir-Simulationsmodells in der Lehrkräftebildung empfohlen werden. Zum einen würde sich damit ein naiver, randomisierter Ansatz als weniger fidel als das Clasivir-Simulationsmodell erweisen, zum anderen würde das Clasivir-Simulationsmodell damit als äquivalent zur Realität wahrgenommen werden.

Bei Ablehnung der $H_{A,1}$ kann davon ausgegangen werden, dass sich das Clasivir-Simulationsmodell nicht statistisch signifikant von der Übertragung einer realen Schulsituation in die Simulationsumgebung unterscheidet. Selbst bei einer statistischen Signifikanz kann bei kleiner Effektstärke immer noch von einer guten Passung für die intendierte Anwendung ausgegangen werden. Zusätzlich wurde auch geprüft, ob sich das Clasivir-Simulationsmodell von einem randomisierten Simulationsmodell unterscheidet. Die Ergebnisse dieser Untersuchung sind aufgrund des ohnehin explorativen Vorgehens mit besonderer Vorsicht zu interpretieren, da die Teilnehmenden im Verlauf des Forschungsprojektes nicht mehr in Gruppen mit randomisiertem Simulationsmodell eingeteilt wurden. Dennoch ist eine Untersuchung dieser Hypothese

sinnvoll: Aus einer Ablehnung von $H_{B,1}$ kann zunächst gefolgert werden, dass das Clasivir-Simulationsmodell für den intendierten Einsatz schlecht geeignet ist. Würden gleichzeitig die Fidelitätswerte des randomisierten Simulationsmodells über den Fidelitätswerten des realen Simulationsmodells liegen, so würde dies gegen die Methodik des gewählten Instruments sprechen, da in diesem Fall anzunehmen ist, dass Studierende Simulationsmodelle von Unterricht nicht sicher bewerten können.

Sollte $H_{A,0}$ verworfen werden und das Clasivir-Simulationsmodell als fiderler bewertet werden als das Modell, das auf der realen Schulstunde basiert, so könnte dies dafür sprechen, dass Clasivir allgemein akzeptierte Vorstellungen, Annahmen und vielleicht auch Klischees über Schulunterricht reproduziert, jedoch nicht akkurat die Wirklichkeit abbildet. Es könnte auch implizieren, dass Studierende zu diesem Zeitpunkt in ihrer professionellen Karriere die Fidelität eines Simulationsmodells noch nicht genau einschätzen können. Liegen die Fidelitätswerte des Clasivir-Simulationsmodells und des realen Simulationsmodells eng beieinander, spricht dies stark für eine Implementierung des Simulationsmodells.

6.2 Quantitative Validierung des Modells

Um die wahrgenommene Simulationsfidelität zu bestimmen wurde ein mixed/methods Ansatz gewählt. Zunächst soll der quantitative Teil beschrieben werden, bei dem Lehramtsstudierende einen Online/Fragebogen absolvieren. Dieser Online-Fragebogen enthält verschiedene sogenannte Mockup-Videos, die das jeweilige Simulationsmodell audiovisuell aufbereitet präsentieren. Der Grad der audiovisuellen Aufbereitung wird in der Studie variiert, ebenso wie die verwendeten Simulationsmodelle (siehe [Tabelle 4](#), 195). Weil die Partizipantinnen immer zwei Simulationsmodelle miteinander vergleichen sollen, werden ihnen zwei Videos präsentiert. Da das Lehrkraftverhalten in allen Simulationsmodellen weitestgehend identisch ist, können die Videos der ca. 8 Minuten langen Unterrichtsausschnitte in zwei Blöcke à ca. 4 Minuten aufgeteilt werden. Daraus ergeben sich insgesamt zwölf Untersuchungsgruppen - später auf 4 Untersuchungsgruppen reduziert, in die die Teilnehmerinnen durch Urnenziehung mit Zurücklegen randomisiert eingeteilt wurden. Eine Übersicht über die in dieser Studie verwendeten Gruppen ist in [Tabelle 4](#) (S. 195) zu finden. Eine Übersicht zu den verwendeten Video-Mockups mit Hyperlinks zu den jeweiligen Videos ist [Tabelle 5](#) (S. 195) zu entnehmen.

Tabelle 4: Übersicht über die Einteilung der Gruppen

Gruppe	Simulationsmodelle		Darbietung
	Erster Videoabschnitt	Zweiter Videoabschnitt	
G1	real	Clasivir	2D
G2 [zu G1]	real	randomisiert	2D
G3	Clasivir	real	2D
G4 [zu G3]	Clasivir	randomisiert	2D
G5 [zu G3]	randomisiert	real	2D
G6 [zu G1]	randomisiert	Clasivir	2D
G7	real	Clasivir	3D
G8 [zu G7]	real	randomisiert	3D
G9	Clasivir	real	3D
G10 [zu G9]	Clasivir	randomisiert	3D
G11 [zu G9]	randomisiert	real	3D
G12 [zu G7]	randomisiert	Clasivir	3D

Tabelle 5: Übersicht über die verwendeten Mockup-Videos

Videokürzel	Verwendetes Sim-Modell	Timecode Start- und Endzeit	Darb.	Link
2D 1-1	real	00:01:42:00 – 00:05:18:00	2D	https://youtu.be/82_EYYTeGtc
2D 1-2	real	00:05:18:00 – 00:10:00:00	2D	https://youtu.be/_KzUala3kig
2D 2-1	Clasivir	00:01:42:00 – 00:05:26:00	2D	https://youtu.be/TVZR8Se3l-s
2D 2-2	Clasivir	00:05:26:00 – 00:09:26:00	2D	https://youtu.be/E592H6s2lEQ
2D 3-1	randomisiert	00:01:42:00 – 00:05:18:00	2D	https://youtu.be/fN03sx5MGCw
2D 3-2	randomisiert	00:05:16:00 – 00:10:00:00	2D	https://youtu.be/h-W0evHhZqc
3D 1-1	real	00:01:42:00 – 00:05:18:00	3D	https://youtu.be/reynpInhlM0
3D 1-2	real	00:05:18:00 – 00:10:00:00	3D	https://youtu.be/HLuvSjQu8D4
3D 2-1	Clasivir	00:01:42:00 – 00:05:18:00	3D	https://youtu.be/VMeJblETXr4
3D 2-2	Clasivir	00:05:26:00 – 00:09:26:00	3D	https://youtu.be/3zq2oYf_eyY
3D 3-1	randomisiert	00:01:42:00 – 00:05:18:00	3D	https://youtu.be/AXHUOwYO8cs
3D 3-2	randomisiert	00:05:16:00 – 00:10:00:00	3D	https://youtu.be/jS6pQVDrdag

6.2.1 Stichprobe

Die Stichprobe setzte sich aus insgesamt N=103 Studierenden vom ersten bis zum achten (und weiterer) Fachsemester zusammen. Im Median befanden sich die Studierenden im 3. Fachsemester. Die genaue Aufschlüsselung ist in [Tabelle 7](#) (S. 197) zu finden. Bis auf eine Person, die Erziehungswissenschaft studierte und zwei Personen, die sich in einem anderen Studiengang befanden, studierten alle Personen einen Lehramtsstudiengang. Auf L1 entfielen 30 Personen, auf L2 15, auf L3 41 und auf L5 16. 67 Personen ordneten sich dem weiblichen, 38 dem männlichen Geschlecht zu, siehe [Tabelle 6](#) (S.

196). Die Möglichkeit sich dem diversen Geschlecht zuzuordnen wurde nicht genutzt. Da die aussagekräftigere Variable für die Untersuchung das aktuelle Fachsemester war, wurde aus Datensparsamkeitsgründen auf eine Erhebung des Alters verzichtet.

Tabelle 6: Geschlechtsangaben der Partizipantinnen und Partizipanten

	absolut	relativ
Männlich	38	36%
Weiblich	67	64%
Divers	0	0%

Die Teilnehmerinnen wurden zum einen innerhalb dreier „Virtual Reality in der Lehrkräftebildung“-Seminare an der Goethe Universität Frankfurt am Main im Jahr 2021–2023 rekrutiert (N=90), die im Rahmen des Projektes Digi_Gap durchgeführt wurden. Zum anderen wurden Lehramtsstudierende aus früheren *digital learning* Seminaren angeschrieben und mit einer Prämie von 15 Euro zur Teilnahme incentiviert (N=15).

Da nur Lehramtsstudierende Zielgruppe der Simulation sind, wurden überwiegend Studierende dieser Studienrichtung rekrutiert, es handelt sich somit um ein bewusstes Auswahlverfahren. Es ist zu beachten, dass ein Großteil der Stichprobe aus Seminaren rekrutiert wurden, die bereits eine inhaltliche Affinität zu VR und *digital learning* vermuten lassen. Eine Repräsentativität für alle Lehramtsstudierende ist somit nicht gegeben. Da bei jedem Besuch der Fragebogenseite ein Datensatz protokolliert wurde (da direkt eine Gruppenzuteilung vergeben werden musste), wurden 306 Datensätze ausgeschlossen, die keine Frage beantwortet haben, also lediglich die Seite aufrufen. Zwei Datensätze brachen an verschiedenen Punkten innerhalb der Bearbeitung ab. Diese Datensätze werden abhängig von der Forschungshypothese benutzt, wenn die nötigen Daten vorhanden sind.

Der größte Anteil an Personen verglich das Clasivir-Simulationsmodell mit dem realen Simulationsmodell, entweder in einer 2D-Aufbereitung (Gruppen G1 und G3) oder in einer 3D-Aufbereitung (Gruppen G7 und G9). Wie bereits erwähnt, wurde im ursprünglichen Forschungsdesign angedacht, alle Gruppen mit gleichen Teilnehmendenzahlen zu besetzen. Dies wurde jedoch aufgrund der Corona-Pandemie und den damit verbundenen Schwierigkeiten Teilnehmende zu akquirieren geändert, so dass Teilnehmende die eigentlich den anderen Gruppen zugeteilt worden wären, anteilig auf G1, G3, G7 und G9 verteilt wurden. Die genauen Zahlen sind [Abb. 19](#) (S. 197) zu entnehmen.

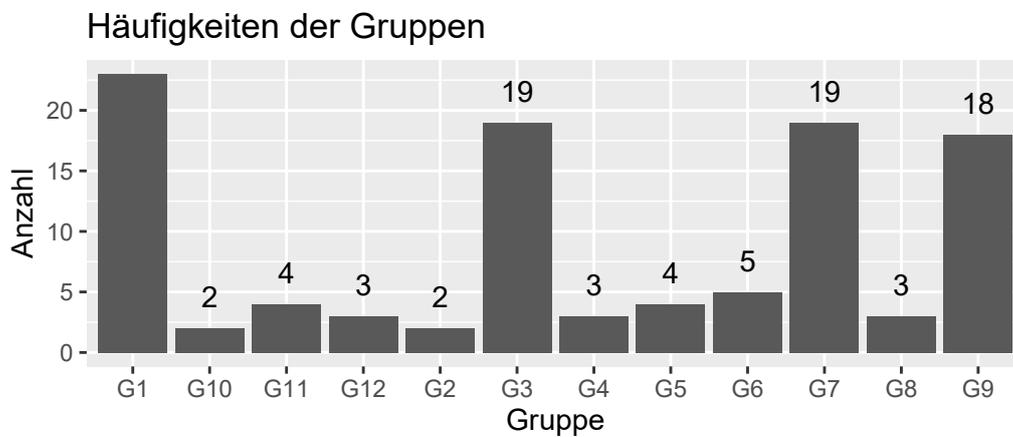


Abbildung 19: Balkendiagramm der Gruppeneinteilungen

Tabelle 7: Partizipantinnen nach Fachsemester

Fachsemester	Anzahl
1	1
2	18
3	12
4	17
5	22
6	10
7	9
8 oder mehr	12

6.2.2 Material

Die Untersuchung benutzt zwei etablierte Skalen zur Erfassung von (a) technischer Kompetenz und (b) computerbezogenen Einstellungen. Die technische Kompetenz wurde mithilfe des TK-Konstruktes innerhalb des TPACK-Modells gemessen.

Die computerbezogenen Einstellungen basieren auf dem INCOBI-R (Inventory zur Computerbildung) Instrument von Richter et al. (2010), welches eine Revision des INCOBI ist (Naumann et al., 2001; Richter et al., 2000). Mit dem INCOBI-R werden computerbezogene Einstellungen über acht verschiedene Skalen analysiert. Eine dieser Skalen ist der Fragebogens zur inhaltlich differenzierten Erfassung computerbezogener Einstellungen (FIDEC). Dieser wiederum ist untergliedert in die Subskalen „Persönliche Erfahrungen“ und

„Gesellschaftliche Folgen“. Beide Subskalen untersuchen in einer weiteren Untergliederung die Rolle des Computers als Lern- und Arbeitsmittel und die Rolle des Computers als Unterhaltungs- und Kommunikationsmittel. Da für diese Arbeit nur das Instrument zur Messung der persönlichen Erfahrungen mit dem Computer als Lern- und Arbeitsmittel relevant ist, wird in dieser Arbeit nur diese Subskala des FIDEC benutzt. Die 20 benutzten Frageitems sind in [Abschnitt 8.3](#) zu finden.

Technologische Kompetenz ist ein Unteraspekt des TPACK-Modells (Koehler & Mishra, 2009). TPACK steht für *Technological Pedagogical Content Knowledge* und beschreibt das für Lehrkräfte nötige Wissen, um Technologie erfolgreich im Unterricht einzusetzen. Es umfasst drei Komponenten: Inhaltskenntnis (*Content Knowledge (CK)*), pädagogisches Wissen (*Pedagogical Knowledge (PK)*) und technologisches Wissen (*Technological Knowledge (TK)*), sowie Integrationen dieser Wissensbezeichnungen (z.B. *Technological Pedagogical Knowledge*). Im Rahmen dieser Arbeit ist jedoch primär das technologische Wissen (TK) der Partizipantinnen von Interesse, weswegen nur dieser Teilaspekt untersucht wird. Technologisches Wissen wird über die 8 Items in Schmidt et al. (2009) angegebenen Items operationalisiert, wobei die Items ins Deutsche übersetzt wurden.

6.2.3 Durchführung

Die Datenerhebung erfolgte in Form eines Online-Fragebogens über die Plattform *soscisurvey*. Die Einladung zur Teilnahme erfolgte durch einen Link auf der Lernplattform OLAT für die verpflichtende Teilnahme im Rahmen der *Digi_Gap* Seminare und über E-Mailanschriften mit Link im Falle der incentivierten, freiwilligen Teilnehmerinnen. Die Erhebung fand ausschließlich online statt und war sowohl auf stationären Computern, Laptops, Tablets als auch auf Smartphones durchführbar. Die Teilnehmerinnen wurden jedoch gebeten, möglichst großformatige Bildanzeigen zu nutzen, da es ansonsten schwierig ist auf den Mockup-Videos sämtliche Informationen gut aufzunehmen.

Auf der ersten Seite des Fragebogens fand eine programmierte Urnenziehung ohne Zurücklegen statt, welche die Teilnehmerinnen in eine von 12 Gruppen einteilte (siehe [Tabelle 4](#), S. 195). Im Verlauf der Datenerhebung wurde dies aufgrund von zu wenig Teilnehmerinnen auf 4 Gruppen (nur Mockup-Videos von *Clasivir* und der realen Situation) reduziert. Des Weiteren wurden die Teilnehmerinnen aufgefordert sich mindestens eine Stunde Zeit zur Beantwortung des Fragebogens zu nehmen. Die Bitte zur Benutzung von großformatigen Bildschirmen erfolgte ebenfalls auf dieser Seite. Zuletzt folgten hier Hinweise zum Datenschutz und eine Kontaktadresse für Fragen,

Fehler im Fragebogen oder Anmerkungen. Die Daten wurden anonym gespeichert.

Der Fragebogen wurde für eine Bearbeitungszeit von maximal einer Stunde konzipiert. Die Bearbeitungszeit wurde für jede Frage automatisiert durch sosciturvy erhoben. Die Gesamtbearbeitungsdauer des Fragebogens betrug durchschnittlich $AM=17.62$ min ($SD=8$ min, $MD=20.07$ min) und liegt damit unter der konzipierten Bearbeitungszeit. Dabei ist zu beachten, dass das Anschauen der Mockup-Videos ($2 \times 5 \text{ min} = 10$ min, zuzüglich eventueller Wiederholungen) in einigen Browser-Konfigurationen nicht mitprotokolliert wurde.

Die zweite Seite fragte die personenbezogenen Attribute Geschlecht und Studiengang ab. Wurde als Studiengang Lehramt ausgewählt, sorgte ein Fragefilter dafür, dass das Fachsemester erfragt wurde (Seite 3). Da das Fachsemester als personenbezogenes Attribut für die Fragestellung ausreicht, wurde aus Gründen der Datensparsamkeit auf eine Erhebung des Alters verzichtet.

Auf Seite 4 wurde das technologische Wissen (TK), auf Seite 5 die Subskala „Einstellungen zum Computer als Lerninstrument“ erhoben (siehe [Abschnitt 6.2.2](#), S. 197 zur Beschreibung der Instrumente, siehe [Abschnitt 8.3](#), S. 284 zur Übersicht über die spezifischen Items der Instrumente). Für beide Skalen wurde eine fünfstufige Likert-Skala benutzt. Seite 6 erfragte weitere personenbezogene Attribute wie Vorerfahrungen mit pädagogischen Klassenraumsimulatoren, Videospieleerfahrungen, VR-Vorerfahrungen, den Besitz einer VR Brille und generelle Einstellung zur Nutzung von VR-Technik innerhalb des Lehramtsstudiums. Details zur Ausgestaltung der Items sind in [Abschnitt 8.3](#) (S. 284) zu finden.

Auf Seite 7 wurde ein Einführungsvideo verlinkt, das die Teilnehmenden mit dem Verfahren der folgenden Analysen der Mockup-Videos vertraut machte. Das Einführungsvideo existierte in zwei Varianten für die 2D-Gruppen und die 3D-Gruppen. Die Videos unterschieden sich nur an einigen wenigen Stellen, in denen es um die grafische Aufbereitung der Daten ging. Die Videos können über folgende Links eingesehen werden:

- 2D – <https://youtu.be/pjsXTWjTLHE>
- 3D – https://youtu.be/1Td_bUclDSo

Nach Beendigung des Erhebungszeitraumes (April 2023) wurde der Datensatz von der Plattform heruntergeladen und in die Software RStudio importiert, dort aufbereitet und ausgewertet.

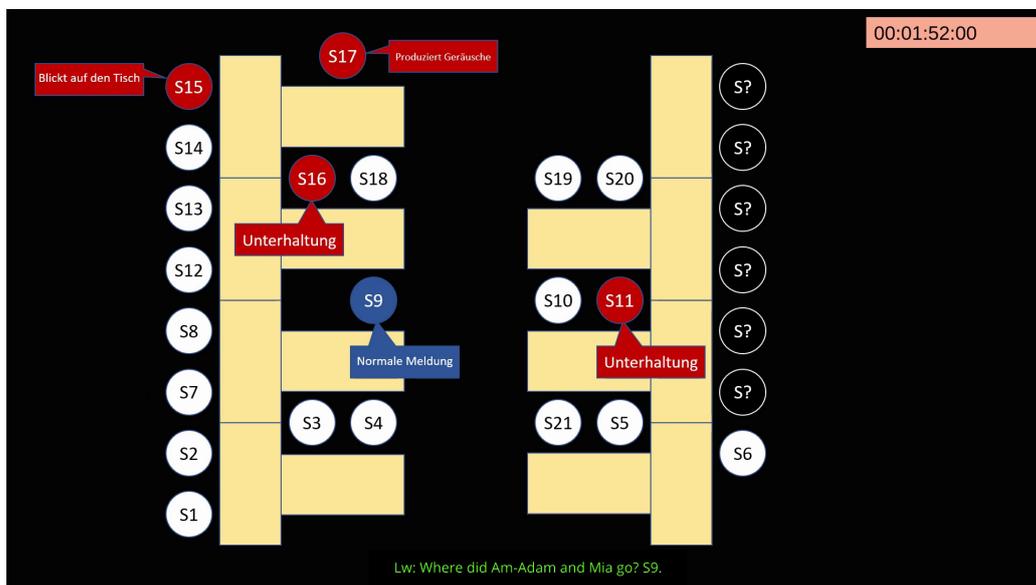


Abbildung 20: 2D-Mockup Video der Simulationsmodelle

6.3 Mockup-Videos

Unter dem Begriff ‚Mockup-Video‘ soll an dieser Stelle ein Video verstanden werden, das den Verlauf des Klassenraumgeschehens grafisch aufbereitet darstellt, ohne dass die Simulation dafür komplett fertiggestellt sein muss. Anders formuliert: Ein Mockup-Video zeigt festgelegte Lehrkraftinteraktionen, die in einem vollständig implementierten Simulator frei wählbar wären. Im Rahmen der Studie wurden zwei unterschiedliche Typen von Mockup-Videos erstellt:

- (a) 2D-Mockup-Videos. Diese Videos zeigen den Klassenraum aus der Vogelperspektive (siehe [Abb. 20](#), S. 200). Wechseln SuS ihr Verhalten, indem sie sich melden, Störverhalten zeigen oder eine Antwort geben, so wird dies durch eine Texteinblendung neben der anonymisierten SuS-Kennung angezeigt. SuS, die nicht aufgezeichnet werden durften, weil Sie keine Einverständniserklärung abgegeben haben, werden in dieser Darstellungsform als S? oder Sx bezeichnet.
- (b) 3D-Mockup-Videos. Bei diesen Videos wird das Verhalten der SuS mittels animierter 3D-Figuren gezeigt (siehe [Abb. 21](#), S. 201). Die Konstruktion dieses im technischen Verfahren aufwendig herzustellen Videos wird in [Abschnitt 6.3.2](#) (S. 204) beschrieben.

Beide Arten von Videos enthalten Sprache der Lehrkraft und Sprache der SuS. Bei dem realen Simulationsmodell wurde ein Transkript erstellt und



Abbildung 21: 3D-Mockup Video der Simulationsmodelle

dieses Wort für Wort, sowohl für die Lehrkraft als auch für SuS nachträglich durch den Autoren eingesprochen und stimmverfremdet, um verschiedene Personen zu indizieren. Zusätzlich wurden aus Störverhalten resultierende Geräusche (z.B. motorische Unruhe) aufgenommen, so dass die Partizipantinnen den Eindruck einer in Echtzeit ablaufenden Schulstunde haben.

Bei den randomisierten Videos und Clavivir-Videos wurden die Interaktionen (hauptsächlich Fragestellungen) der Lehrkraft zum Großteil aus dem Transkript übernommen. Da, wo das Verhalten der Lehrkraft abweicht, wurden sinnvoll gewählte Alternativformulierungen gefunden. Dies trifft zum Beispiel dann zu, wenn eine Lehrkraft in einem simulierten Modell eine Schülerin ermahnt, die in der realen Situation kein Störverhalten gezeigt hat. Da das Verhalten der SuS das Unterscheidungsmerkmal der Simulationsmodelle darstellt, konnte nur fragmentarisch auf Aussagen von SuS aus dem realen Transkript zurückgegriffen werden. Das bedeutet, dass unter anderem die gesprochenen Beleidigungen im Clavivir- und randomisierten Modell konstruiert sind. Das Clavivir-Modell sagt lediglich voraus, dass eine Beleidigung stattgefunden hat, nicht aber den konkreten Inhalt. Die Konstruktion der SuS-Sprache basierte auf den Kenntnissen des Autors in der Videografie von ca. 600 Unterrichtsstunden im Rahmen des Projektes Level/The Next Level. Alle sprachlichen Anteile wurden als Untertitel am unteren Bildrand im Video angezeigt.

Dieses Vorgehen ist nicht problemfrei. Es ist nicht ausgeschlossen, dass

Partizipantinnen sich von der konkret gesprochenen SuS-Sprache mehr beeinflussen lassen als von den visuell dargestellten Störverhalten. Die konkreten sprachlichen Ausgestaltungen von SuS-Aussagen können für sich genommen realistisch oder unrealistisch wirken. In allen Einführungsvideos wurde dieser Umstand explizit angesprochen und problematisiert. Die Partizipantinnen wurden angewiesen nicht auf die konkrete sprachliche Ausgestaltung zu achten.

Als Alternative zum gewählten Vorgehen hätten keine sprachlichen Elemente verwendet werden können. Dies würde allerdings zu zwei gravierenden Problemen führen: (1) Damit wäre die beobachtete Simulation so weit von einer realen Unterrichtsbeobachtung entfernt, dass eine Einschätzung des Kernaspekts dieser Arbeit, der Fidelität, nur schwer möglich wäre, insbesondere durch Novizinnen. (2) Mit einer abstrakten Implementation ohne Sprachanteile würde sich der benötigte Leseanteil für die Partizipantinnen stark erhöhen. Um überhaupt zu verstehen, was konkret in der Klasse passiert, muss bei jeder Handlung der Lehrkraft das Video gestoppt werden, da realistischerweise keine Partizipantin ihre Aufmerksamkeit auf 21 SuS in 2 Sekundenintervallen aufsplitten kann. Da dies eine ohnehin kognitiv höchst anspruchsvolle Aufgabe noch komplexer machen würde, wurden sprachliche Anteile in das Video eingewebt, um so den den *cognitive load* der Partizipantinnen zu reduzieren.

Sämtliche Simulationsmodelle (real, Clasivir, randomisiert) basieren auf den Lehrkraftaktionen einer realen Unterrichtsstunde, die im Rahmen der QLB-geförderten Projekte Level/The Next Level entstanden sind. Da der Autor für die Produktion aller ca. 600 Unterrichtsstunden, die im Rahmen von Level/The Next Level aufgezeichnet worden sind, verantwortlich war, konnte ein Auswahlprozess zu passenden Videos zur Fragestellung eingeleitet werden. Folgende notwendige Kriterien wurden dabei angewendet:

- Der verwendete Unterrichtsausschnitt darf nur aus öffentlichem Unterricht (keine Partner-, Gruppen- oder Einzelarbeit) bestehen, weil Clasivir nur dieses Unterrichtsgeschehen simulieren kann.
- Der öffentliche Unterricht muss primär von der Lehrkraft geleitet/geführt werden.
- Die Jahrgangsstufe muss zwischen 5 und 6 liegen.

Zusätzlich wurden folgende hinreichende Kriterien beachtet:

- Das Unterrichtsgeschehen sollte ein gewisses Maß an Störungsverhalten zeigen.

- Lehrkraftinformationen liegen nicht in langer, monologartiger Form vor. Ein Unterrichtsgespräch mit Aufgaben findet statt.
- Lehrkraftfragen sind im besten Falle gut zu kodieren, d.h. eindeutig zuzuordnen und nicht zu lang.
- Möglichst alle SuS sollten bei der Aufzeichnung sichtbar sein (d.h. möglichst alle SuS sollten eine Einverständniserklärung zur Aufzeichnung abgegeben haben).

Nach genauerer Evaluation mehrerer Videos fiel die Entscheidung auf eine videografierte Englischunterrichtsstunde einer Realschulklasse der sechsten Jahrgangsstufe, welche im Rahmen des Fächerverbundes Sprachen des Projektes Level 2016 entstanden ist.

Im Kapitel Methodenkritik (siehe [Abschnitt 6.4](#), S. 206) ist genauer aufgeführt, welche Probleme die verwendete Stunde beinhaltet, welche Entscheidungen getroffen worden sind, um mit diesen Problemen umzugehen und welche Probleme bei der Kodierung der Verhaltensweisen auftraten.

6.3.1 Kodierung

Der erste Schritt in der Erstellung der Mockup-Videos war die Analyse der realen Unterrichtsstunde. Da das Simulationsmodell von Clasivir diskret aufgebaut ist und sich die Verhaltensweisen der SuS in zweisekündigen Abständen ändern, wurde eine Intervallkodierung mit zweisekündigen Abständen durchgeführt. Kodiert wurden sowohl Lehrkraftverhalten als auch SuS-Verhalten mit den jeweiligen Ausprägungen, die den in Clasivir verwendeten Verhaltensweisen entsprachen. Die Kodierung beginnt direkt nach der Begrüßung der SuS durch die Lehrkraft und dem Stellen der ersten inhaltlichen Frage. Folgende Verhaltensweisen wurden für das Lehrkraftverhalten kodiert:

- Information geben
- Frage stellen (Kognitive Dimension / Wissensdimension / Ziel)
- Aufruf von SuS (Ziel)
- classroom-management (loben / im Verhalten bestärken / konstruktiv kritisieren / destruktiv kritisieren)
- Feedback auf Antwort geben (loben / bejahen / verneinen / kritisieren)

Folgende Verhaltensweisen wurden für alle einundzwanzig SuS einzeln kodiert:

- Meldung (undeutliche Meldung / normale Meldung / aufdringliche Meldung)
- Antwort (richtig / falsch)
- Störverhalten (Unterhaltung / Motorische Unruhe / Geräusch- oder Lärmproduktion / Kaugummikauen / Ins Heft malen / Im Raum herumlaufen / Aus dem Fenster blicken / Auf den Tisch blicken / Schlafen / Mit dem Handy telefonieren / Mit dem Handy spielen / Bedrohung / Beschimpfung / Auslachen / Herabwürdigen / Schwacher Tritt / Starker Tritt / Schwacher Schlag / Starker Schlag / Blatt zerreißen / Stift zerstören / Tisch beschmieren)

Die Kodierung wurde mit dem Videoanalysetool Mangold Interact (Breitenbach & Appel, 2016) durchgeführt. Interact ermöglicht es einzelne Intervalle abzuspielen und so das Kodieren von umfangreichen Daten einfacher umzusetzen. Die in Interact erzeugten Daten wurden anschließend in Excel aufbereitet und bildeten die Grundlage für die Entwicklung des ersten Simulationsmodells (real). Das reale Simulationsmodell zeigt demzufolge alle Lehrkraftverhalten und SuS-Verhalten, die in der realen Unterrichtsstunde zu beobachten waren. Die Mockup-Videos überführen die kodierten Verhaltensweisen in eine andere Darstellungsform, damit sie mit den anderen Simulationsmodellen verglichen werden können. Eine Nutzung des originalen Videos war nicht möglich, da ansonsten direkt ersichtlich wäre, welches Simulationsmodell der Realität entnommen wurde.

6.3.2 Zur Konstruktion der Mockup-Videos

Die grafische Aufbereitung der 2D-Mockup-Videos wurde in Powerpoint, die der 3D-Mockup-Videos in Blender durchgeführt. Die Erstellung der 3D-Mockup-Videos war sehr viel komplexer und bedarf einer genaueren Beschreibung. Beide Videos wurden in die Videoschnittsoftware Da Vinci Resolve importiert, dort mit Sounddaten vervollständigt (siehe [Abschnitt 6.3](#), S. 200) und Untertitel im *.srt Format fest in das Video enkodiert. Die dazu nötigen *.srt-Daten wurden mittels des Programms Subtitle Edit erzeugt. Die Videos wurden exportiert, auf Youtube (nicht indexiert) hochgeladen und von dort aus mit dem Fragebogen in soscisurvy verknüpft.

Die Erstellung der 3D-Videos in Blender kann man in drei Phasen aufteilen: (1) Erstellung der Figuren und der Rigs¹³ (2) Erstellung der Animationen

¹³Die mit den Figurenvertices verknüpfte Knochenstruktur und nötige Kontrollmechanismen werden als ‚Rig‘ bezeichnet.

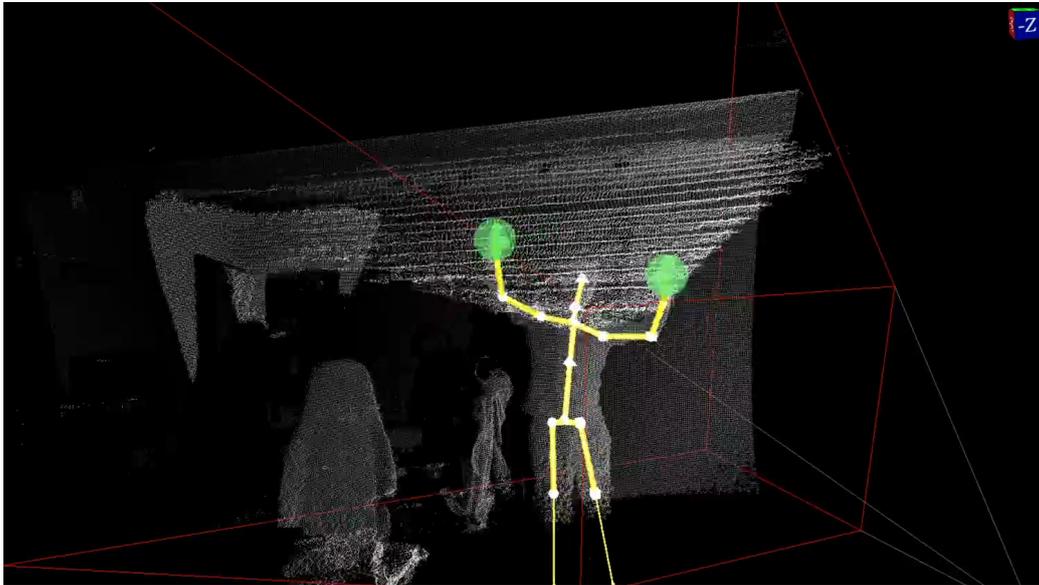


Abbildung 22: Nutzung des Kinect-Sensors zur Erstellung von motion capturing-Daten, Quelle: Autor

in Form von Actions und (3) Mapping der Actions zu den jeweiligen im Simulationsmodell vorhergesagten Handlungen in der NLE-Umgebung (*non-linear editor*) von Blender. Die Erstellung der Figures gestaltete sich vergleichsweise einfach, da hierfür die quelloffene Software *makehuman3D* verwendet werden konnte, mit der man menschliche 3D-Gitternetzmodelle, auch Meshs genannt, mit Texturen anhand verschiedener Merkmale wie Alter, Geschlecht usw. erzeugen konnte. Diese Merkmale sind über Slider kontrollierbar, so dass unendlich viele Zwischenstufen möglich sind. Durch diesen Prozess konnten 6 generische SuS-Meshs erstellt werden, die zudem im Rahmen des Teilprojektes VR-Lerngelegenheiten im QLB-Projekt *Digi_Gap* in der VR-Simulation *Clasivir 1.0* genutzt wurden (siehe [Abschnitt 8.3](#), S. 238). Bei 21 SuS bedeutet dies allerdings, dass sich das Aussehen der SuS zum Teil wiederholt. Da die Erstellung von 21 verschiedenen SuS-Modellen zu viele Ressourcen in Anspruch genommen hätte, insbesondere was die Erstellung der Animationen betrifft, wurde darauf verzichtet.

Einige Animationen wurden unter Zuhilfenahme von *motion capturing* mittels des 3D-Tiefensensors Microsoft Kinect 2 erstellt (siehe [Abb. 22](#)). Andere Animationen wurden manuell in Blender bearbeitet, da die *motion capturing* Technik für die Details der Handbewegungen und Handinteraktionen (Stift, Buch, etc.) zu ungenau war (siehe [Abb. 23](#)).

Alle 22 in *Clasivir* modellierten Störverhalten wurden für die 3D-Figuren



Abbildung 23: Darstellung des manuellen Animationsverfahrens in Blender, Quelle: Autor

erstellt und anschließend anhand einer Timeline in Blender (NLE=Umgebung) auf die Figuren gemappt.

6.4 Methodenkritik

Die gewählte Methode kann zunächst nicht alle Teile des entwickelten Simulationsmodells vollständig evaluieren. Unter anderem ermöglicht es der gewählte Ansatz nicht den Algorithmus zur Erstellung der Namen zu prüfen. Auch der Emotionsalgorithmus (siehe [Abschnitt 5.9](#), S. 171) kann nicht evaluiert werden.

Die Evaluation des Simulationsmodells über Mockup-Videos mit einer Vertonung der Lehrkraft- und SuS-Aussagen erzeugt methodologische Probleme, die sich nicht vollständig auflösen lassen. Generell zielt die Analyse der Mockup-Videos darauf ab, den Realitätsgrad der Häufigkeit und Abfolge der gezeigten SuS-Handlungen zu beurteilen.

Das erste Problem ist, dass es für die Partizipantinnen schwer ist, das gezeigte Verhalten der Lehrkraft zu ignorieren. Im Fragebogen wurden die Partizipantinnen dazu aufgefordert, sich eine unerfahrene Lehrkraft vorzustellen, wenn sie die Handlungen der Lehrkraft nicht plausibel fanden. Aber trotz dieser Anweisung ist anzunehmen, dass die Handlungen der Lehrkraft, etwa, ob auf eine Störung reagiert wird, Einfluss auf die Bewertung des Realitätsgrads des SuS-Verhaltens hat. Bei der Nutzung eines fertiggestellten Simulators initiieren Versuchspersonen selbst die Lehrkraftinteraktionen und

das Problem stellt sich in dieser Form nicht mehr. Damit wäre es aber methodisch unmöglich das Simulationsmodell mit einer realen Unterrichtssituation zu vergleichen.

Das zweite Problem liegt in der linguistischen Umsetzung des Simulationsmodells. Da das Clasivir-Simulationsmodell neue Handlungsalternativen erzeugt, müssen konkret gesprochene Aussagen ex nihilo generiert werden. Eine Vertrautheit mit der Beobachtung von Schulalltag hilft hierbei zwar, löst das Problem aber nicht vollständig.

Der gewählte Weg über einen Online-Fragebogen macht es für Partizipantinnen möglich sich nicht tiefergehend mit den Mockup-Videos auseinander zu setzen. Trotz Aufforderung, die Videos mehrfach anzusehen, ist es möglich die Videos lediglich einmal zu betrachten und danach die Fragen zum Simulationsmodell zu beantworten. Um zu erfassen, ob die Partizipantinnen dieses Verhalten tatsächlich zeigen, wurde der Fragebogen so konfiguriert, dass er die Zeit zur Beantwortung der Fragen protokolliert. Diese Maßnahme kann technisch wie inhaltlich umgangen werden. Es kann nicht geprüft werden, ob eine Partizipantin beispielsweise das Fragefenster lediglich eine Stunde geöffnet hält oder ob tatsächlich eine Arbeit am Video durchgeführt wurde. Diese Daten stellen nichtsdestotrotz einen gangbaren Proxy dar um festzustellen, ob ein umfassendes Abkürzen des Fragebogens stattfand.

6.4.1 Betrachtung des Kodierungsprozesses

Obwohl die ausgewählte Unterrichtsstunde als Grundlage für die verschiedenen Simulationsmodelle die geeignetste Aufzeichnung für den intendierten Einsatzzweck ist, so existieren gewisse Fallstricke in der Kodierung der videografierten Unterrichtsstunde.

Am problematischsten ist sicherlich, dass sechs SuS keine Einverständniserklärung abgegeben haben und demzufolge nicht in der Aufzeichnung zu sehen sind. Die Lehrkraft interagiert mit diesen SuS selbstverständlich dennoch, da diese im Raum verblieben. Bei einigen anderen Videoerhebungen im Projekt Level/The Next Level wurden SuS ohne Einverständnis in eine Parallelklasse geschickt, dies war in dieser Aufzeichnung jedoch nicht der Fall.

Obwohl die Klasse insgesamt recht unruhig ist, insbesondere im Vergleich zu anderen Aufzeichnungen, so finden keine gravierenden, extremen Störverhalten statt, wie Schläge oder Handynutzung. Selbst Beleidigungen, Herabsetzungen oder Auslachen waren in der Unterrichtsstunde nicht zu beobachten. Hierbei ist zu erwähnen, dass dies kein Spezifikum der gewählten Stunde war, sondern gravierende Störverhalten in dem gesamten Level-Videobestand so gut wie nie auftraten, was durch den Invasivitätseffekt der Kamerabeobachtung zu erklären ist. Diese gravierenden Störverhalten treten aber im Unter-

richtsgeschehen (wenn auch selten) auf, so dass ein Simulationstraining diese Störverhalten auch abbilden sollte.

Obwohl die Aufzeichnung mit vier Grenzflächenmikrofonen zur genaueren Analyse der SuS-Sprache während des öffentlichen Unterrichts durchgeführt wurde, so ist dennoch anzumerken, dass geflüsterte Interaktionen nur dann wirklich hörbar waren, wenn sich SuS in der Nähe eines Mikrofons befunden haben. Um eine Aussage als Beleidigung zu klassifizieren ist es nötig zu verstehen, was konkret gesagt worden ist. Damit ist es theoretisch möglich, dass Beleidigungen fälschlicherweise als Unterhaltungen kodiert wurden. Unterhaltungen selbst konnten gut kodiert werden, da diese rein visuell gut zu identifizieren waren.

Auffällig war, dass die ausgewählte Stunde weder Lob noch destruktive Kritik in den Lehrkraftfeedbacks auf SuS-Antworten enthielt. Dies könnte unter anderem damit zu erklären sein, dass die Lehrkraft über große Lehrererfahrung verfügt und zudem unter Beobachtung stand.

Ursprünglich war geplant SuS-Emotionen zu kodieren, um auch diesen Aspekt des Simulationsmodells zu evaluieren. Als klar wurde, dass Verhalten, die auf Emotionen verweisen (wie Lächeln als Indikator für Freude) nur selten beobachtbar waren und es auch in Clasivir noch keinen Algorithmus gibt, um (innere) Emotionen auf äußerliche Verhaltensweisen zu extrapolieren, wurde diese Idee wieder fallengelassen. Eine Modellierung von sporadisch auftretenden sichtbaren Verhaltensweisen zur Darstellung der inneren emotionalen Zustände hätte den Algorithmus von Clasivir zum einen exponentiell komplexer gemacht und zum anderen wären aus der realen Stunde kaum genug Beobachtungsdaten zu verwenden gewesen.

Wie bereits bei anderen Videoaufzeichnungen, die der Autor durchgeführt hat, war auch in dieser Stunde das Phänomen des ‚Puppenverhaltens‘ zu erkennen. Einige SuS, wie z.B. S6, zeigten fast während der gesamten Aufzeichnung kein Störverhalten, kein Meldeverhalten und kein Antwortverhalten. Sie folgten mit den Blicken der Lehrkraft, verhielten sich also rollenkonform, interagierten aber ansonsten nicht mit ihrer Umgebung. Dieses Verhalten stellt als reales Phänomen nicht notwendigerweise ein Problem dar. Da die quantitative Validierung des Simulationsmodells aber auf beobachtbaren Verhaltensweisen fußt, könnte dieses ‚Puppenverhalten‘ die Fidelitätsbewertung beeinflussen. Zeigt die Mehrheit der SuS dieses ‚Puppenverhalten‘ – was bei der genutzten Stunde glücklicherweise nicht der Fall war – ist eine Validierung des Simulationsmodells mit der hier vorgeschlagenen Methode nicht möglich, da keine ausreichende Anzahl an beobachtbaren Verhaltensweisen dargestellt werden kann.

Ein anderes beobachtbares Verhalten war das Phänomen der Dauermel-

dungen (Sacher, 1995), also Meldungen von SuS über eine längere Zeit, unabhängig von einer konkreten Aufgabe. Die Darstellung des realen Simulationsmodells im Mockup-Video ist zwar kein Problem, allerdings kann das Clasivir-Simulationsmodell (und das randomisierte Simulationsmodell) dieses Verhalten nicht abbilden.

In der realen Stunde ist häufig ein Abfolgemuster im Meldeverhalten zu beobachten. So folgen Meldungen bei bestimmten SuS dem Muster ‚Unauffällig-Normal-Unauffällig‘, ‚Unauffällig-Aufdringlich-Normal‘ und ‚Aufdringlich-Normal-Unauffällig‘. Das Interaktionsmodell in Clasivir ist noch nicht der Lage diese Muster zu modellieren, die Entscheidung für einen Meldetyp bleibt bis zum nächsten Fragetyp oder bis zum Abbruch der Meldung konstant.

Während die kognitive Dimension und die Wissensdimension der Aufgaben recht gut zu erfassen waren, war die Kodierung der spezifischen Schwierigkeit schwieriger, zumal hierfür strenggenommen spezialisiertes fachdidaktisches Wissen erforderlich gewesen wäre. Aus Ressourcengründen musste auf eine Miteinbeziehung von Fachdidaktikerinnen verzichtet werden. Abmildernd wirkte, dass die Fragen sehr eindeutig auf einem geringen Schwierigkeitsniveau über die gesamte Stunde lagen.

Die Zeit, die die Lehrkraft braucht, um eine Aufgabe zu stellen, variiert stark. Teilweise wurden auch Gesten als Aufgabestellung interpretiert (z.B. 00:06:46:00). Aufgabenstellungen überschreiten aber in der Regel die Kodierungs-Intervalllänge von zwei Sekunden. Aufgabenstellungen wurden wie folgt kodiert: Die Zeit zum Stellen einer Aufgabe wurde als „Informationen geben“ klassifiziert, mit demselben Schwierigkeitsgrad wie die Aufgabe, die später gestellt wird. Wurde die Aufgabe danach lediglich wiederholt und änderte sich die Zielperson der Aufgabe nicht, so wurde auf eine weitere Kodierung verzichtet. Der Grund dafür ist, dass man ansonsten Wartezeiten, die sich aus einer Aufgabenwiederholung ergeben, unterschätzt.

Selbst bei dem recht kurz gewählten Beobachtungsintervall von zwei Sekunden ist teilweise zu beobachten, dass SuS während dieser Zeit ihr Verhalten wechseln. Zumeist tritt dies auf, wenn motorische Unruhe in einen idle-Zustand zurückkehrt. Meistens wurde in dieser Situation ein Störverhalten kodiert; fanden zwei Störverhalten statt, wurde das invasivere kodiert.

Die Kodierung der Feedbackschleifen gestaltete sich in der Abwägung zwischen „konstruktiver Kritik“ und „Bejahung“ häufig schwierig:

- Bei ca. 7:48 gibt die Lehrkraft Feedback auf die Schülerantwort „He run out of the garden.“ Die Lehrkraft nickt stark und sagt im bestätigendem Ton „He ran out of the garden.“ Später fügt sie noch hinzu - was die Situation eindeutiger macht - „He ‚ran‘, nicht ‚run‘“. Das Feedback ist also trotz gestischem Signal und bestätigendem Ton als konstruktiv-

ve Kritik zu deuten, da mit der Aufgabe nicht nur eine Reproduktion von Faktenwissen verknüpft war, sondern ebenfalls eine grammatikalisch korrekte Umsetzung. Dies geht jedoch nicht vollständig aus der gestellten Frage hervor, ist damit also in gewissem Maße interpretativ.

- Ein gegensätzliches Verhalten ist in der Situation zu finden, die ca. 8:04 stattfindet. Hier antwortet der Schüler S12 auf die Frage „What did Adam say to the kids on the street?“ mit „Get lost.“, was die Lehrkraft erst mit einem emphatischen „Yes, exactly!“ belohnt. Nach einigen Sekunden korrigiert sie sich und sagt: „Actually, he said: „Hey, get lost“. Es ist ersichtlich, dass das Hinzufügen der Interjektion ‚Hey‘ die Antwort nicht falsch macht oder negativ beeinflusst. Die Aussprache war akzeptabel, die Satzkonstruktion (wenn auch sehr simpel) richtig und die inhaltliche Frage wurde korrekt beantwortet. Deswegen ist dieses Feedback als „Bejahung“ zu klassifizieren.
- Einige Feedbackaussagen sind auch nach intensiver Beschäftigung mit der Mikrosituation unklar. Ca. 9:20 gibt die Lehrkraft auf die Schülerantwort „They locket [sic] for Buster.“ das Feedback „They looked for Buster. [ohne Betonung]“, während die Lehrkraft das korrekte Antwortkärtchen an der Tafel befestigt. Selbstverständlich ist der Lehrkraft klar, dass ein Aussprachefehler begangen wurde, weswegen sie diesen Fehler in der Wiederholung der Schülerantwort korrigiert. Sie verzichtet aber darauf den Fehler zu thematisieren. Daher ist nicht ganz klar, ob die Antwort korrekt war und das Feedback dies bejaht oder ob die Antwort falsch war und das Feedback den Aussprachefehler kritisiert. Dies ist abhängig davon, was genau mit der Frage „Another word for searched?“ erfragt werden sollte: Ging es eher um das Erinnern des Wortes „looked“ oder ging es auch um die Aussprache? Es ist anzunehmen, dass es um alle Aspekte ging, aber inwiefern ist hier der Aussprachefehler zu werten? Unterminiert er die gesamte Antwort? Dieses Dilemma ist innerhalb des Simulationsmodells von Clasivir nicht lösbar, da man die Antwort als ‚teilweise korrekt‘ klassifizieren müsste und das Feedback sowohl bestätigt als auch konstruktiv kritisiert. Eine Implementierung solcher Feinheiten hätte das Simulationsmodell von Clasivir gesprengt und hätte zu einem *bloated model* geführt.

Ein weiteres wichtiges Problem in der Kodierung war, wie man mit beobachtbarem SuS-Störverhalten umgeht, das nicht im Katalog der Clasivir-Störverhalten zu finden ist. Es wurde die Entscheidung getroffen diese Verhaltensweisen mit den Clasivir-Störverhalten zu kodieren, die ihnen am ehesten entsprechen:

- Ca. 7:16 schlägt eine Schülerin das Heft auf und blättert darin, vermutlich um die Frage der Lehrkraft beantworten zu können. Die Lehrkraft wies die SuS an, die Bücher und Hefte geschlossen zu halten, während sie Fragen zu der in der letzten Stunde gelesenen Geschichte stellt. Dies wurde zu einem ähnlichen Störverhalten übertragen, nämlich „Ins Heft malen“.
- 2:10 ist eine Form von Kommunikation mit Gesten durch Zunicken, Blickkontakt und Mimik zwischen zwei Schülern zu beobachten. Dies wurde als Unterhaltung kodiert, auch wenn keine Worte gesprochen wurden.
- 3:00 ist zu beobachten wie S10 mit einem Gegenstand spielt. Es handelte sich dabei wohl um ein Gummiband. Dieses Verhalten enthält keine genaue Entsprechung im Clasivir-Störverhaltenssystem. Da es sich um ein passives, wenig invasives Störverhalten handelt, wurde es kodiert als „Im Heft malen“.
- Oftmals wurde das Störverhalten „aus dem Fenster schauen“ kodiert, wenn der Blick einer Schülerin einem anderen als den intendierten Gegenständen zugewandt war, selbst wenn es sich dabei nicht um das Fenster handelte. SuS blickten in einigen Fällen beispielsweise auf den an einer Korkwand aufgehängten Stundenplan.

Ein weiteres Problem ist das Phänomen, dass mehrere Störverhalten oder Störverhalten mit Meldung gleichzeitig auftreten:

- Ca. 9:28 schaut der Schüler S5 auf den Tisch während er motorische Unruhe zeigt. In Fällen von gedoppelt auftretenden Störverhalten wird immer die invasivere Störung kodiert, in diesem Fall die motorische Unruhe. Stoppt eines der Störverhalten im Verlauf der Zeit wird das andere Störverhalten weiterkodiert.
- Ein gemeinsames Auftreten von undeutlichen Meldungen und motorischer Unruhe war vereinzelt beobachtbar. Es wurde auf Einzelfallebene entschieden, welche der beiden Codes zu vergeben waren, unter Berücksichtigung der vorhergehenden und nachfolgenden Ereignisse sowie dem Grad der Invasivität.

Letztlich sind in der Stunde verschiedene Verhaltensweisen zu beobachten, die nicht kodiert wurden:

- 2:36 fällt S8 ein Mäppchen versehentlich herunter, was S9 leise aufhebt. Dies wurde nicht als Störverhalten kodiert, da die Intention erkennbar war, die Situation so schnell und unauffällig wie möglich zu regeln, die Situation nicht absichtlich herbeigeführt worden ist und sie keinen Einfluss auf das Unterrichtsgeschehen hat.
- 1:43 findet eine eine non-verbale Kommunikation statt, die jedoch kaum eine Sekunde andauert und nur aus einem Kopfnicken hin zu einem anderen Schüler besteht. Diese Situation wurde nicht kodiert, da sie keine Unterrichtsstörung konstituiert.
- Die Grenze zwischen motorischer Unruhe und nicht kodiertem Idle-Verhalten ist nicht immer trennscharf. So kann beispielsweise intensives Augenreiben mit geschlossenen Augen als motorische Unruhe kodiert oder als Idle-Verhalten gewertet und damit nicht kodiert werden. Dies wurde auf Einzelfallebene entschieden, abhängig davon wie abgelenkt die SuS bei diesen Aktionen wirkten. Tendenziell wurde ein eher strenger Maßstab angelegt, da die virtuellen Figuren im Mock-up sonst gar kein Verhalten zeigen würden (Vgl. Beschreibung des ‚Puppenverhaltens‘ in diesem Kapitel).

6.4.2 Betrachtung der Clasivir-Funktionsweise

Bei der Umsetzung des Simulationsmodells in Clasivir wurden die Lehrkraftfragen in derselben Abfolge und denselben kognitiven Dimensionen und Wissensdimensionen gestellt wie in dem realen Unterrichtsausschnitt. *Classroom-management* bezogene Aktionen und Feedback mussten der konkreten Situation in der Simulation angepasst werden. Da der Grund dafür vielleicht nicht ohne weiteres einleuchtend ist, soll dies mit einem kleinen Beispiel illustriert werden. Angenommen in der realen Unterrichtssituation kritisiert die Lehrkraft S15 konstruktiv, weil dieser ein Störverhalten zeigt, nachdem eine Aufgabe gestellt wurde. Wird dieselbe Situation in Clasivir simuliert kann es sein, dass S15 gar kein Störverhalten zeigt, sondern die Aufgabe wie gewünscht bearbeitet. Stattdessen könnte z.B. S1 ein Störverhalten zeigen. Es wäre deswegen sinnvoll, die konstruktive Kritik auf S1 zu übertragen. Wenn im Clasivir-Simulationsmodell keine SuS Störverhalten zeigen, wäre es eine realistische Handlungsalternative gar keine *classroom-management* bezogene Aktion durchzuführen.

Ein weiteres Problem betrifft die Interaktion mit S[?]/S_x-SuS, also Interaktionen mit SuS, die keine Einverständniserklärung zur Aufzeichnung abgegeben haben und deswegen nicht im Video zu sehen sind. Es wurde die Entscheidung getroffen, diese Schülerinnen weder in Clasivir noch im randomisierten

Modell zu simulieren und stattdessen an S[?]/S_x gerichtete Fragen auf andere SuS zu übertragen. Wären in Clavivir und im randomisierten Simulationsmodell S[?]/S_x modelliert worden, so könnte man deren Verhaltensweisen in den Mockup-Videos nicht abbilden, weil ansonsten das reale Unterrichtsmodell klar identifizierbar wäre. Eine Modellierung von Interaktionen mit den S[?]/S_x-SuS ohne sie visuell in den Mockup-Videos darzustellen, wie es im realen Unterrichtsvideo der Fall war, hätte zur Folge, dass viele SuS-Interaktionen, die bewertet werden sollen, verdeckt stattfinden, also gar nicht von den Partizipantinnen beobachtet werden könnten. Es wurde sich deswegen gegen eine solche Modellierung ohne Abbildung entschlossen.

7 Ergebnisse

7.1 Quantitative Ergebnisse der Modellvalidierung

7.1.1 Datenaufbereitung

Ein wichtiger erster Schritt in der Datenaufbereitung war die Trennung der Fidelitätswerte für die einzelnen Simulationsmodelle in eigene Datensätze. Während im allgemeinen Datensatz N=105 Personen vorhanden sind, musste für die Trennung der Datenwerte berücksichtigt werden, dass jede Person zwei Simulationsmodelle verglich und damit auch Fidelitätswerte für zwei verschiedene Modelle generierte. Die Separierung der Datensätze wurde direkt in R durchgeführt. Da die Fidelitätswerte für Störverhalten, Meldeverhalten und Antwortverhalten im originalen Datensatz in die falsche Richtung zeigen, wurden alle Items so umkodiert, dass sie von 1 (niedrigste Fidelität) bis zu 4 (höchste Fidelität) reichen. Fidelitätswerte wurden anschließend als das arithmetische Mittel aus Fidelität des Störverhaltens, des Meldeverhaltens und des Antwortverhaltens der beobachteten SuS berechnet. Demzufolge reicht auch der Fidelitäts-Gesamtwert von 1 (niedrigste Fidelität und damit am unrealistischsten) bis zu 4 (höchste Fidelität und damit am realistischsten).

Die einzelnen Komponenten der Fidelitätswerte, nämlich Störverhalten, Meldeverhalten und Antwortverhalten sind wie erwartet nur schwach miteinander korreliert (.19-.29).

Der FIDEC Subskalen-Wert zu persönlichen Erfahrungen mit Computern im Bereich Lernen und Arbeiten wurde durch Berechnung des arithmetischen Mittels der einzelnen Items gewonnen, wobei die Werte der Likert-Skala (1–5) des ursprünglichen Datensatzes in Werte von -2 bis +2 umgerechnet wurden. Dabei steht -2 für die Einstellung, dass es sich bei Computern um eine unbeeinflussbare Maschine handelt und +2 dafür, dass der Computer als nützlich

ches Werkzeug wahrgenommen wird. Dafür mussten die Items 13, 14, 17, 26, 27, 50, 56, 58, 62 und 68 invertiert werden. Obwohl der Fragebogen eigentlich so konfiguriert war, dass keine Items fehlen durften um zur nächsten Seite zu gelangen, fehlten im Datensatz bei 7 Personen einzelne der insgesamt 20 FIDEC-Items. Um diese Daten weiterhin verwenden zu können wurde eine Mean-Imputation durchgeführt.

Die TK (technische Kompetenz) aus dem TPACK Modell wurde durch Berechnung des arithmetischen Mittel aus den einzelnen Werten der Items berechnet, wobei die ursprüngliche Likert-Skala (1–5) in Werte von -2 (niedrige Kompetenz) bis +2 (hohe Kompetenz) umgerechnet wurde. Da die Items immer in eine Richtung zeigen waren Invertierungen hier nicht nötig.

Um die Hypothese zu prüfen wurden die Werte für Computerspiel-erfahrungen von Computern und Konsolen in eine Variable überführt, wobei der höhere Wert für diesen zusammenfassenden Wert übernommen wurde. Eine simple arithmetische Mittelung der Daten hätte für den Forschungsfragenbezug zu Verzerrungen geführt.

Gewähltes statistisches Verfahren ist der T-Test für unabhängige Stichproben.

Voraussetzungsprüfung

- Fidelität wurde durch Aggregation der Items Fidelität Störungsverhalten, Fidelität Meldeverhalten und Fidelität Antwortverhalten in eine metrisch skalierte Variable überführt, die Voraussetzung ist damit erfüllt.
- Die einzelnen Messwerte sind durch das Forschungsdesign voneinander unabhängig.
- Nach optischer Prüfung folgt die Verteilung der Fidelität nicht der Normalverteilung. Dies ist auch mit einem Shapiro-Wilk Test zu bestätigen, der mit $W = .92738$, $p = 7.363e-08$ ein signifikantes Ergebnis liefert. Die Voraussetzung ist daher verletzt. Da der T-Test jedoch robust gegen die Verletzung der Normalverteilungsannahme ist, der Levene-Test Homoskedastizität anzeigt (s.u.) und $N=179$ () Beobachtungen vorhanden sind, kann der T-Test trotzdem durchgeführt werden.
- Der Levene-Test zeigt Homoskedastizität ($F(1,179) = .201$, $p = .6545$). Die Voraussetzung ist daher erfüllt.

Gewähltes statistisches Verfahren ist der T-Test für unabhängige Stichproben.

Voraussetzungsprüfung

- Fidelität ist durch Aggregation der Items Fidelität Störungsverhalten, Fidelität Meldeverhalten und Fidelität Antwortverhalten eine metrisch skalierte Variable, die Voraussetzung ist damit erfüllt.
- Die einzelnen Messwerte sind durch das Forschungsdesign voneinander unabhängig.
- Nach optischer Prüfung folgt die Verteilung der Fidelität nicht der Normalverteilung. Dies ist auch mit einem Shapiro-Wilk Test zu bestätigen, der mit $W = .94453$, $p = .0001159$ ein signifikantes Ergebnis liefert. Die Voraussetzung ist daher verletzt. Da der T-Test jedoch robust gegen die Verletzung der Normalverteilungsannahme ist, der Levene-Test Homoskedastizität anzeigt (s.u.) und $N=116$ () Beobachtungen vorhanden sind, kann der T-Test trotzdem durchgeführt werden.
- Der Levene-Test zeigt Homoskedastizität ($F(1,114) = 2.0363$, $p = .1563$) an. Die Voraussetzung ist daher erfüllt.

Gewähltes statistisches Verfahren ist Kendall's Tau-B-Koeffizient, da bei den Daten Rangbindungen vorliegen.

Voraussetzungsprüfung

- Die Variable zur Bewertungsschwierigkeit ist ordinal skaliert und paarweise unabhängig von der metrisch skalierten technischen Kompetenz (TK). Damit sind alle Voraussetzungen erfüllt.

Gewähltes statistische Verfahren ist der Wilcoxon-Test.

Voraussetzungsprüfung

- Die Schwierigkeit des Ratings ist ordinal skaliert und paarweise unabhängig von der visuellen Darbietung. Die erste Voraussetzung ist damit erfüllt.
- Die optische Prüfung auf Vergleichbarkeit der Verteilungsform in den beiden Gruppen auf Grundlage von [Abb. 25](#) (S. 223) ergibt, dass zwar Unterschiede im Bereich „eher schwierig“ und „eher einfach“ bestehen, die Verteilungen aber im wesentlichen ähnlich sind. Die Voraussetzung ist damit erfüllt.

- Prüfung auf Gleichheit der Streuungen wird über den Levene-Test berechnet, der die Werte $F(1,101) = .15$, $p = .6994$ lieferte. Homoskedastizität wird damit angenommen, die Voraussetzung ist damit erfüllt.

Gewähltes statistisches Verfahren ist die Spearman-Korrelation.

Voraussetzungsprüfung

- Die FIDEC-Subkala ist metrisch, der wahrgenommene Nutzen der Simulation in der Lehrkräftebildung ordinal skaliert. Die erste Voraussetzung ist damit erfüllt.
- Aufgrund der deskriptiven Ergebnisse in [Abb. 30](#) (S. 228) kann ein monotoner Zusammenhang vermutet werden. Die zweite Voraussetzung ist damit erfüllt.
- Einzelne fehlende Werte in den FIDEC-Subskala-Items wurden durch Mean-Imputation ergänzt, so dass weder fehlende Werte noch Ausreißer vorliegen. Die dritte Voraussetzung ist somit erfüllt.

Gewähltes statistisches Verfahren sollte der T-Test für unabhängige Stichproben sein, nach Scheitern der Voraussetzungen wurde der Wilcoxon-Test gewählt.

Voraussetzungsprüfung 1: (T-Test)

- Fidelität ist durch Aggregation der Items Fidelität Störungsverhalten, Fidelität Meldeverhalten und Fidelität Antwortverhalten eine metrisch skalierte Variable, die Voraussetzung ist damit erfüllt.
- Die einzelnen Messwerte sind durch das Forschungsdesign voneinander unabhängig.
- Nach optischer Prüfung folgt die Verteilung der Fidelität nicht der Normalverteilung. Dies ist auch mit einem Shapiro-Wilk Test zu bestätigen, der mit $W = .92469$, $p = 5.672e-05$ ein signifikantes Ergebnis liefert. Die Voraussetzung ist daher verletzt.
- Der Levene-Test zeigt Heteroskedastizität ($F(1,89) = 5.5648$, $p = .02051$) an. Die Voraussetzung ist daher verletzt.

Wegen der zwei schwerwiegenden Voraussetzungsverletzungen wurde stattdessen der nicht-parametrische Wilcoxon-Test zur statistischen Überprüfung gewählt. **Voraussetzungsprüfung 2: (Wilcoxon-Test)**

- Die Fidelität der Simulationsmodelle ist metrisch skaliert und lässt sich in eine ordinale Skalierung mit Rangbindungen transformieren. Die Fidelität ist paarweise unabhängig von der gezeigten Stunde. Die erste Voraussetzung ist damit erfüllt.
- Die optische Prüfung auf Vergleichbarkeit der Verteilungsform in den beiden Gruppen auf Grundlage von [Abb. 29](#) (S. 227) ergibt, dass die Verteilungen im wesentlichen ähnlich sind. Die Voraussetzung ist damit erfüllt.

Gewähltes statistisches Verfahren ist die Spearman-Korrelation.

Voraussetzungsprüfung

- Sowohl die Videospielerfahrung (Computer & Konsole) als auch die wahrgenommene Schwierigkeit in der Bewertung der Simulationsmodelle liegen ordinalskaliert vor. Damit sind alle Voraussetzungen erfüllt.

Gewähltes statistisches Verfahren ist der Wilcoxon-Test.

Voraussetzungsprüfung

- Die wahrgenommene Nützlichkeit ist ordinal skaliert und paarweise unabhängig von dem Geschlecht. Die erste Voraussetzung ist damit erfüllt.
- Die optische Prüfung auf Vergleichbarkeit der Verteilungsform in den beiden Gruppen auf Grundlage von [Abb. 27](#) (S. 225) ergibt, dass zwar Unterschiede im Bereich „sehr nützlich“ bestehen, die Verteilungen aber im wesentlichen ähnlich sind. Die Voraussetzung ist damit erfüllt.
- Prüfung auf Gleichheit der Streuungen wird über den Levene-Test berechnet, der die Werte $F(1,101) = .1546$, $p = .695$ lieferte. Homoskedastizität wird angenommen, die Voraussetzung ist damit erfüllt.

7.1.2 Deskriptivstatistische Ergebnisse

Die Fidelität des realen Simulationsmodells betrug $M=2.98$ ($SD=.49$) und entspricht damit einer Wertung von „eher realistisch“. Die Fidelität des Clasivir Simulationsmodell liegt etwas darunter mit $M=2.87$ ($SD=.48$), während das randomisierte Simulationsmodell mit $M=2.55$ ($SD=.57$) am schlechtesten bewertet wurde. Während sich die Fidelität der zwei Unterrichtsausschnitte der Clasivir- und randomisierten Modelle nicht stark unterscheiden, wird das reale Simulationsmodell in dem zweiten Video mit $M=3.14$ ($SD=.55$) etwas höher bewertet als das erste Video mit $M=2.83$ ($SD=.37$). Weitere Details zu den gerateten Fidelitäten sind [Tabelle 8](#) (S. 218) zu entnehmen. Die Boxplots der Simulationsmodell-Fidelitätswerte sind in [Abb. 24](#) (S. 219) angeführt. Dabei ist zu erkennen, dass das Clasivir-Simulationsmodell recht viele Ausreißer hat, aber der Median derselbe ist wie im realen Simulationsmodell. Das randomisierte Simulationsmodell besitzt eine hohe Streuung und liegt im Median niedriger als die anderen beiden Simulationsmodelle.

Tabelle 8: Deskriptive Fidelitätswerte aller Darstellungsformen

		n	mean	sd	median	min	max	skew	kurtosis
Fidelität	aller	91	2.98	.49	3.00	2.00	4.00	.55	-.37
realen Videos									
Fidelität	aller	91	2.87	.48	3.00	1.33	4.00	-.05	1.06
Clasivir Videos									
Fidelität	aller	26	2.55	.57	2.67	1.33	4.00	.29	-.07
randomisierten Videos									
Fidelität	Reale	47	2.83	.37	2.67	2.33	3.67	.31	-.66
Stunde S1									
Fidelität	Reale	44	3.14	.55	3.00	2.00	4.00	.22	-1.01
Stunde S2									
Fidelität	Clasi-	42	2.87	.60	3.00	1.33	4.00	-.26	.10
vir S1									
Fidelität	Clasi-	49	2.86	.36	3.00	2.33	4.00	.74	.94
vir S2									
Fidelität	Rando-	16	2.56	.67	2.33	1.33	4.00	.31	-.65
misiert S1									
Fidelität	Rando-	10	2.53	.39	2.67	2.00	3.00	-.40	-1.55
misiert S2									

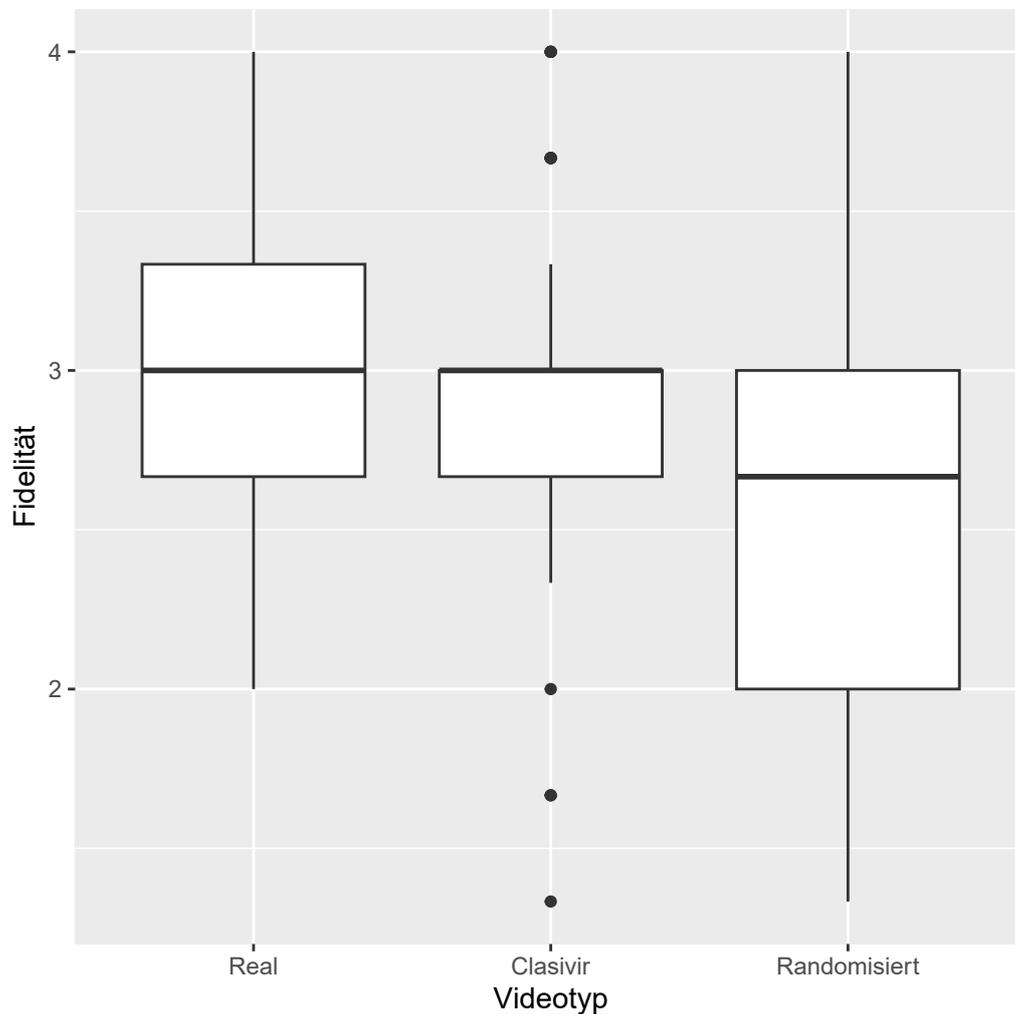


Abbildung 24: Boxplots der Fidelitätswerte aller Videos der verschiedenen Darstellungstypen

Die meisten Teilnehmenden studierten Lehramt im L3 und L1 Studiengang (69%), nur eine Person studierte Erziehungswissenschaft und zwei Personen einen anderen Studiengang. Die restliche Anzahl der Teilnehmenden verteilte sich auf L2 und L5 (siehe [Tabelle 9](#), S. 220).

Tabelle 9: Angegebene Studiengänge

Studiengang	absolut	relativ
L1	31	30%
L2	15	14%
L3	41	39%
L5	15	14%
Erziehungswissenschaft	1	1%
Anderer	2	2%

Die Teilnehmenden verteilen sich recht ausgeglichen auf alle Fachsemester, wobei nur eine Person das erste Fachsemester angab (siehe [Tabelle 10](#), S. 220)

Tabelle 10: Angegebene Fachsemester

	absolut	relativ
1	1	1%
2	18	17%
3	14	13%
4	17	16%
5	22	21%
6	11	10%
7	10	10%
8 oder mehr	12	12%

Partizipantinnen wurden befragt, ob sie bereits Erfahrungen mit einer anderen Simulation in der Lehrkräftebildung hatten. In diesem Item wurde einige Male „Ja“ angegeben, aber lediglich auf die Nutzung von Clasivir 1.0 (siehe [Abschnitt 8.3](#), S. 238) im Rahmen des VR-Seminars rekurriert, was aus der Freitextnennung der konkreten Simulation hervorging. Damit hatten alle Partizipantinnen keine weiteren Erfahrungen mit Simulationen in der Lehrkräftebildung gemacht. Sieben Personen (6.8%) verfügten über eine eigene VR-Brille im Haushalt.

Auffällig war, dass viele Personen angaben, weder auf Computer (48%), noch auf der Konsole (40%) Videospiele zu spielen. Damit liegt der Anteil zwar im bundesdeutschen Durchschnitt, jedoch sehr viel niedriger als das bundesdeutsche Spielverhalten in der Alterskohorte der 16–29 jährigen (88% Spielerinnen) (Tenzer, 2022). Dabei ist zu beachten, dass diese Diskrepanz durch die getrennte Erhebung der Spieleplattformen zu erklären ist, insbesondere das Spielen auf Smartphones wurde häufig angegeben (67%), scheint

aber eher ein Zeitvertreib in Wartezeiten zu sein, da oft angegeben wurde nur seltener als einmal die Woche zu spielen (37 %). Nähere Daten sind in [Tabelle 11](#) (S. 221) zu finden.

Tabelle 11: Angegebene Videospieleerfahrungen

Wie oft spielen Sie Videospiele (auf den folgenden Geräten)?					
	Computer	Konsole	Smartphone	Tablet	
Einmal oder mehrmals pro Tag	15 (14%)	5 (5%)	9 (9%)	3 (3%)	
Einmal oder mehrmals pro Woche	14 (13%)	22 (21%)	22 (21%)	4 (4%)	
seltener	26 (25%)	36 (34%)	39 (37%)	25 (24%)	
gar nicht	50 (48%)	42 (40%)	35 (33%)	73 (70%)	

Die meisten Personen konnten Unterschiede oder große Unterschiede zwischen den beiden dargebotenen Simulationsmodellen erkennen (64%). Bemerkenswert ist, dass bei den Personen, die keinen Unterschied oder kaum Unterschiede zwischen den Simulationsmodellen erkennen konnten, 30 von insgesamt 35 Personen (86%) Gruppen zugeordnet waren, die die reale Stunde mit dem Clavir-Simulationsmodell verglichen haben. Weitere Daten sind in [Tabelle 12](#) (S. 221) zu finden.

Tabelle 12: Unterschiedlichkeitseinschätzung der Videos

Wie sehr unterschieden sich die Videos im Realitätsgrad der Schülerinnen und Schüler?		
	absolut	relativ
sehr unterschiedlich	13	12%
unterschiedlich	57	54%
kaum unterschiedlich	31	30%
kein Unterschied	4	4%

Die Partizipantinnen fanden die Bewertung des Realitätsgrades des von den SuS gezeigten Verhaltens weder besonders schwer, noch besonders einfach. Die genauen absoluten und relativen Häufigkeiten sind [Tabelle 13](#) (S. 222) zu entnehmen. Mit $M=2.48$ ($SD=.67$) für die 2D-Gruppen und $M=2.61$ ($SD=.67$) für die 3D-Gruppen unterscheidet sich die Schwierigkeit der Bewertung der Simulationsmodelle geschlüsselt nach der visuellen Darbietung rein deskriptiv etwas, aber nicht besonders stark. Dies ist auch in [Abb. 25](#) (S. 223) zu erkennen. Die Partizipantinnen schätzen ihre Einschätzkompetenz

als mittelhoch bis hoch ein, wobei sich wiederum die 2D- und 3D-Gruppen wenig unterschieden (2D: $M=2.65$, $SD=.78$, 3D: $M=2.73$, $SD=.57$). Es wurde zusätzlich erfragt, welche der grafischen Aufbereitung zuzuordnenden Eigenschaften die Bewertung der Simulationsmodelle erschwert hat. Dabei ist auffällig, dass die Nicht-Beachtung der konkret gesprochenen SuS-Aussagen in beiden Visualisierungen ein gewisses Problem darstellten. Die meisten der anderen Eigenschaften stellten eine leichte Erschwerung dar, da sie sich im Bereich des arithmetischen Mittels der jeweiligen Items befanden. Nähere Daten dazu sind [Tabelle 14](#) (S. 222) zu entnehmen.

Tabelle 13: Schwierigkeit in der Bewertung der Videos

Wie schwierig fanden Sie es den Realitätsgrad des von den Schülerinnen und Schülern gezeigten Verhalten zu bewerten?		
	absolut	relativ
sehr schwierig	4	4%
eher schwierig	47	45%
eher einfach	48	46%
sehr einfach	6	5%

Tabelle 14: Schwierigkeit nach grafischer Aufbereitung

Wie sehr hat Sie folgende Eigenschaften der grafischen Aufbereitung eine Bewertung der Schülerinnen und Schüler erschwert?							
	sehr erschwert	eher erschwert	eher nicht erschwert	gar nicht erschwert	mean	sd	median
Gruppen mit 3D Visualisierung							
Qualität der 3D-Figuren	4	12	21	10	2.76	.88	3
Qualität der gezeigten Animationen	5	20	17	7	2.53	.87	2
Verzögerung zwischen Grafik und gesprochenem Text	7	19	16	5	2.53	.98	2
Nicht-Beachtung der konkret gesprochenen Schülerinnen- und Schüleraussagen	8	20	16	5	2.37	.88	2
Gruppen mit 2D Visualisierung							
Grafische Aufbereitung als 2D-Darstellung	7	25	15	7	2.41	.88	2
Verzögerung zwischen Grafik und gesprochenem Text	7	20	13	14	2.63	1.01	2.5
Nicht-Beachtung der konkret gesprochenen Schülerinnen- und Schüleraussagen	8	21	17	8	2.46	.93	2

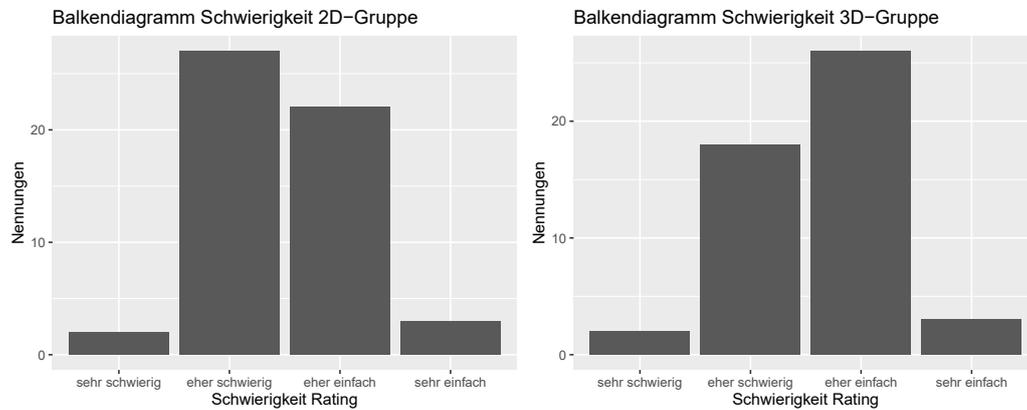


Abbildung 25: Balkendiagramm der wahrgenommenen Schwierigkeit nach visueller Darbietung

Die Subskala FIDEC-Werte der Partizipantinnen betrug (mit Mean-Imputationen einiger weniger fehlender Items) insgesamt $M=.94$ ($SD=.74$) auf einer Skala von -2 (Computer wird als unbeeinflussbare Maschine im Kontext der eigenen Erfahrung und mit dem Computer als Lerninstrument wahrgenommen) bis +2 (Computer wird als nützliches Werkzeug im Kontext der eigenen Erfahrung und mit dem Computer als Lerninstrument wahrgenommen). Die meisten Partizipantinnen haben eher hohe FIDEC-Werte, was bei einer Stichprobe von Studierenden aus Seminaren zu den Themen VR und *digital literacy* nicht verwunderlich ist. Eine Boxplot-Aufschlüsselung in [Abb. 26](#) (S. 224) der Subskala-FIDEC Werte nach Computerspiel- oder Konsolenspiel-Nutzung zeigt bereits deskriptiv, dass recht wenige Unterschiede zu beobachten sind.

FIDEC-Scores in Abhängigkeit zu Videospiele am PC

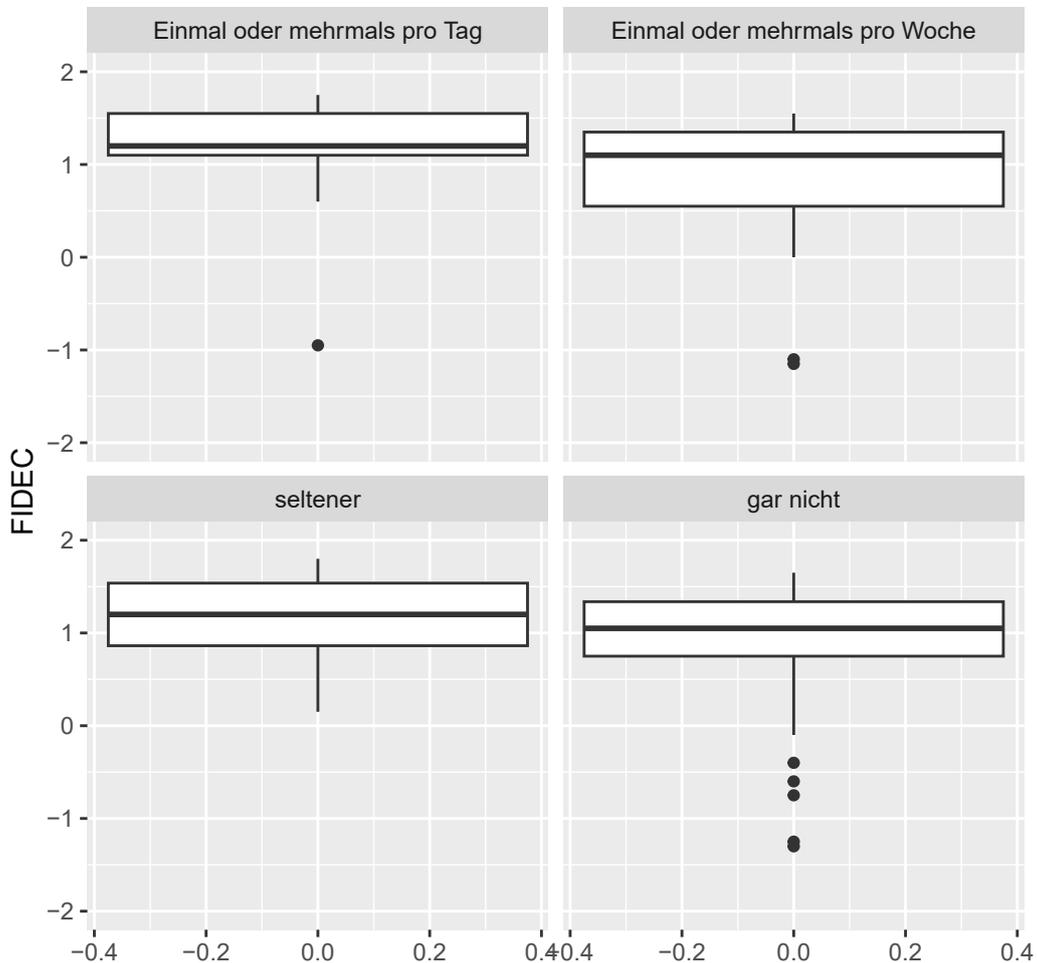


Abbildung 26: Boxplots der FIDEC-Werte geordnet nach Videospieldnutzung Computer

Deskriptiv ist zu erkennen, dass weibliche Nutzerinnen etwas seltener den Einsatz digitaler Klassenraumsimulationen wie Clasivir als „sehr nützlich“ bewerten. Ansonsten unterscheiden sich die Gruppen kaum. Diese Beobachtung wird in [Abb. 27](#) (S. 225) grafisch über ein Balkendiagramm veranschaulicht. Mit einem arithmetischen Mittelwert des wahrgenommenen Nutzens von $M=1.83$ (1=sehr nützlich, 4=nicht nützlich, $SD=.75$) finden die meisten Studierenden den Einsatz von VR-Simulationen in der Lehrkräftebildung eher nützlich bis sehr nützlich.

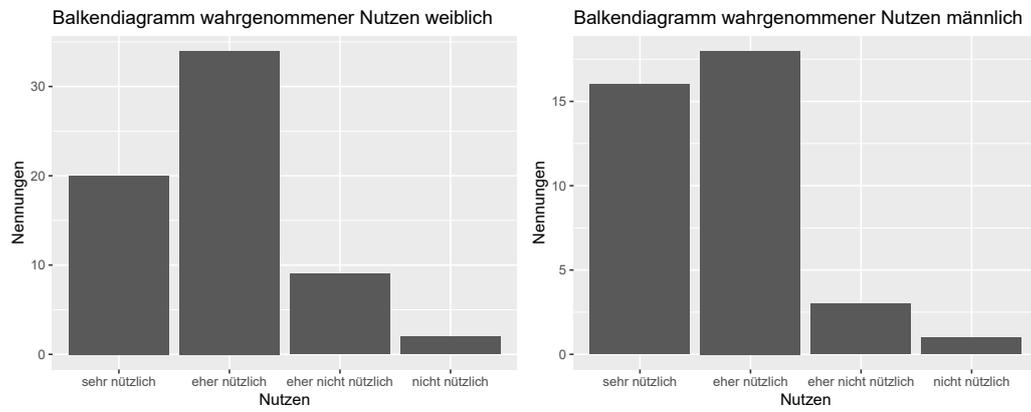


Abbildung 27: Balkendiagramm des wahrgenommenen Nutzens nach Geschlecht

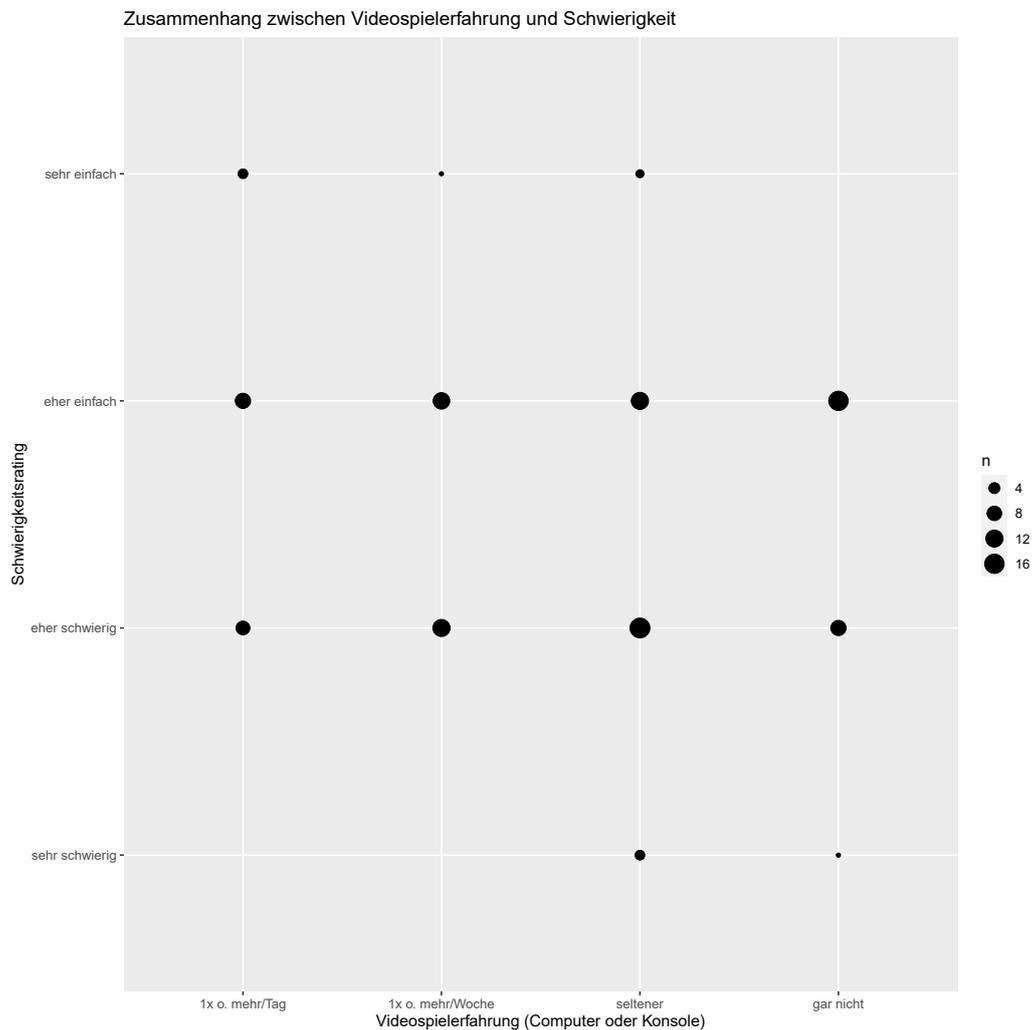


Abbildung 28: Zusammenhang zwischen Videospieleerfahrungen (Computer und Konsole) und wahrgenommener Schwierigkeit

Deskriptiv ist in den Boxplots der Fidelität des ersten und zweiten Unterrichtsausschnittes im Clavivir-Simulationsmodell in [Abb. 29](#) (S. 227) kein Unterschied bezüglich des Medians zu erkennen, wobei die zweite Stunde etwas weniger streut (SD 1. Stunde=.60, SD 2. Stunde=.35). Mit $N=42$ Beobachtungen und einem Mittelwert von $M=2.87$ für die erste Stunde und $N=49$ Beobachtungen und einem Mittelwert von $M=2.86$ für die zweite Stunde ist auch der arithmetische Mittelwert praktisch gleich groß.

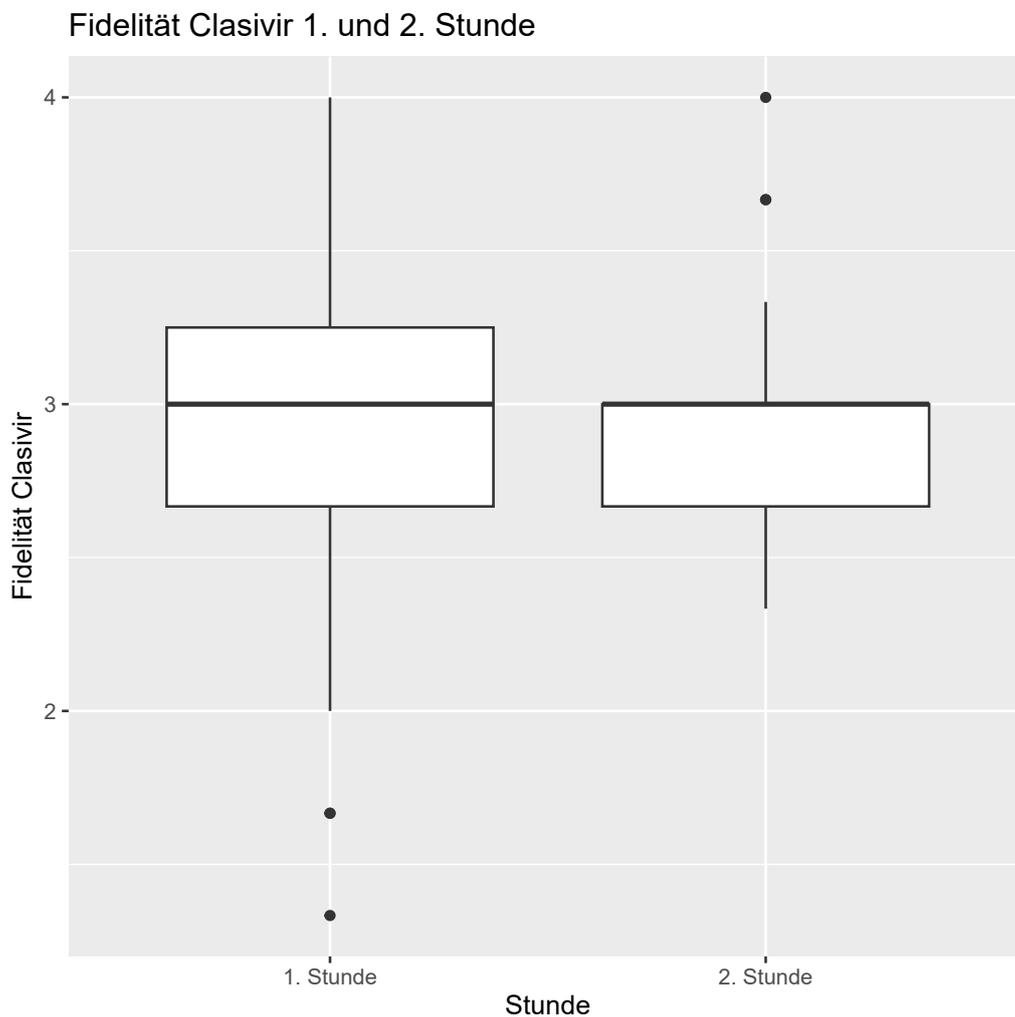


Abbildung 29: Boxplots der Fidelitätsverteilung zwischen dem Clasivir-Modell der präsentierten 1. Stunde und der 2. Stunde

Es ist interessant im Scatterplot in [Abb. 30](#) (S. 228) zu beobachten, dass die Subskala-FIDEC Werte, in keinem einfachen Zusammenhang zum wahrgenommenen Nutzen einer digitalen Klassenraumsimulation wie Clasivir steht. Die Personengruppen mit den niedrigsten FIDEC-Ausprägungen, die den Computer eher als unbeeinflussbare Maschine im Lernkontext und der persönlichen Erfahrung sahen, als auch die Personen mit den höchsten FIDEC-Ausprägungen, die den Computer als nützliches Werkzeug im Lernkontext und der persönlichen Erfahrung sahen, entschieden sich beide dafür den Einsatz von digitalen Klassenraumsimulationen wie Clasivir in der Lehrkräftebildung als „sehr nützlich“ zu klassifizieren. Die Personen, die von einer

eher geringen oder keiner Nützlichkeit ausgingen, befinden sich bis auf einen Ausreißer relativ exakt im arithmetischen Mittel aller FIDEC-Werte.

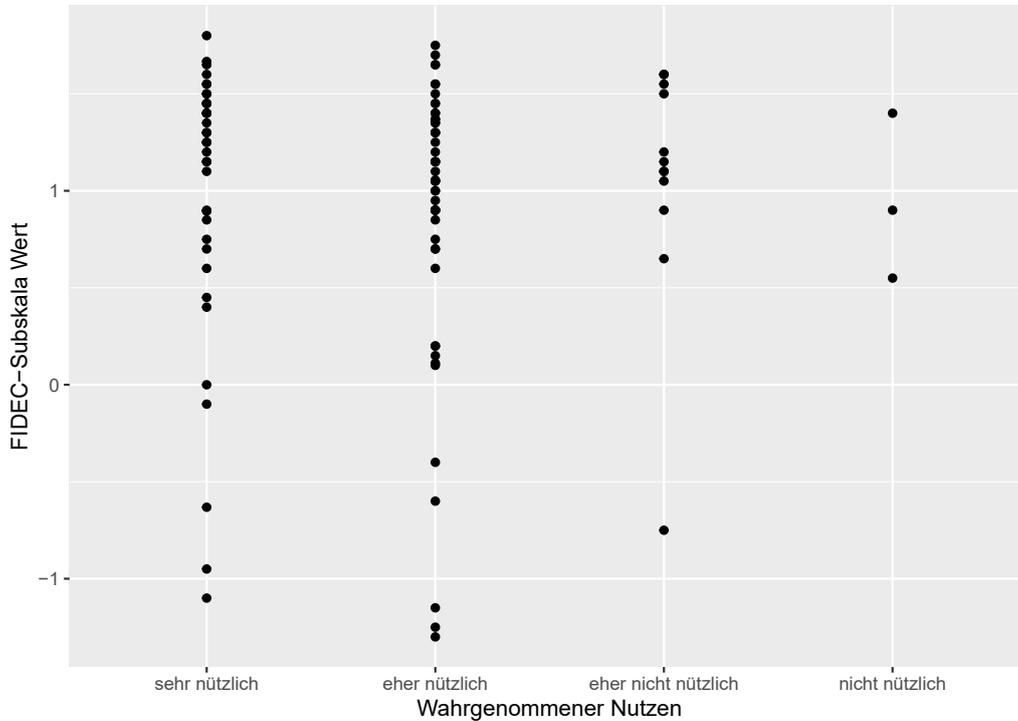


Abbildung 30: Scatterplot zwischen der FIDEC-Subskala und dem wahrgenommenen Nutzen

7.1.3 Inferenzstatistischer Schluss

Die verwendete Skala zur Messung der technischen Kompetenzen im Rahmen des TPACK-Modells kann mit einer tau-äquivalente Reliabilität (Cronbachs Alpha) von .85 (95% CI = [.85,.92]) ebenso wie die Subskala des FIDEC-Instruments zur Erfassung persönlicher Erfahrung mit Computern im Bereich Lernen und Arbeiten mit einer tau-äquivalente Reliabilität von .81 (95% CI = [.75,.86]) als reliabel angesehen werden. Eine Messung der tau-äquivalenten Reliabilität des multidimensionalen Konstruktes „Fidelität“ als Zusammenfassung der Einzelwerte „Störungsverhalten“, „Meldeverhalten“ und „Antwortverhalten“ ist nicht sinnvoll, da auch theoretisch zu erwarten ist, dass die Unterkonstrukte sich unterscheiden und nicht latent dasselbe Konstrukt erfassen.

Damit kann die Validität des Fidelitätskonstrukts mit den erhobenen Daten quantitativ nicht bestätigt werden. Die Anzahl von drei Bewertungskrite-

rien (Störungsverhalten, Meldeverhalten, Antwortverhalten) ist nicht ausreichend, um eine exploratorische oder konfirmatorische Faktorenanalyse durchzuführen. Zudem lieferte der Shapiro-Wilk-Test signifikante Ergebnisse für die Fidelität des realen und des Clasivir-Simulationsmodells, womit eine Normalverteilung nicht gegeben ist und somit eine Voraussetzungsverletzung der Faktorenanalyse vorläge.

Die Prüfung der Hypothese zeigt, dass sich das Clasivir-Simulationsmodell in Simulationsmodellfidelität statistisch nicht signifikant von dem realen Simulationsmodell unterscheidet ($t = -1.463$, $df = 178.9$, $p = .1452$, 95 % CI = [$-.249$, $.0370$]).

Die Prüfung der Hypothese zeigt, dass das Clasivir-Simulationsmodell in Simulationsfidelität statistisch signifikant besser als das randomisierte Modell bewertet wurde ($t = -2.5231$, $df = 33.581$, $p = .008271$, 95 % CI = [inf , $-.105947$]). Mit einer Effektstärke von $d=.634$ handelt es sich um einen moderaten Effekt.

Die Prüfung der Hypothese zeigt, dass die gefühlte Schwierigkeit die Simulationsmodelle zu bewerten nicht mit steigenden technischen Fähigkeiten (TK) abnimmt ($z = 1.2426$, $p = .107$, $\beta = .09963338$).

Die Prüfung der Hypothese zeigt, dass die gefühlte Schwierigkeit die Simulationsmodelle zu bewerten nicht signifikant einfacher für die 3D-Gruppen gegenüber den 2D-Gruppen war ($W = 1167.5$, $p = .1285$).

Die Prüfung der Hypothese zeigt, dass die Personen, die den Computer für ein nützliches Werkzeug im Bereich Lernen und im persönlichen Umgang damit halten (=höhere FIDEC-Subskala-Werte besitzen), nicht statistisch signifikant höhere wahrgenommene Nützlichkeitswerte für VR-Simulationen in der Lehrkräftebildung angeben ($S = 192957$, $p = .2749$, $\beta = -.05959624$).

Die Prüfung der Hypothese zeigt, dass sich das Clasivir Simulationsmodell in den vergebenen Fidelitätswerten nicht zwischen den zwei dargebotenen Stundenausschnitten statistisch signifikant unterscheidet ($W = 1072.5$, $p = .7226$).

Die Prüfung der Hypothese zeigt, dass Personen, die häufiger Videospiele auf einem Computer oder einer Konsole spielen, die gefühlte Schwierigkeit der Simulationsmodelle nicht statistisch signifikant als leichter bewerten ($S = 195345$, $p = .2327$, $\beta = -.07271322$).

Die Prüfung der Hypothese zeigt, dass männliche Partizipanten den wahrgenommenen Nutzen einer VR-Simulation in der Lehrkräftebildung

nicht statistisch signifikant besser wahrnehmen als weibliche Partizipantinnen ($W = 1065.5$, $p = .102$).

7.2 Qualitative Ergebnisse der Modellvalidierung

Im Rahmen des Fragebogens hatten die Studierenden die Möglichkeit Störverhalten zu nennen, die ihnen in den vorgestellten Simulationsmodellen der jeweiligen Mockup-Videos gefehlt haben. Von den 105 Partizipantinnen nutzten 42 diese Möglichkeit für das erste Video und 19 für das zweite Video. Der überwiegende Anteil an Nennungen, nämlich 38, bezieht sich auf Störverhalten, die zwar im Simulationsmodell theoretisch vorhanden sind, aber entweder nicht als solche identifiziert wurden oder im analysierten Mockup-Video nicht stattfanden.

Die Nutzung von Handys und ‚Herumlaufen im Raum‘ während des Unterrichts beispielsweise fand weder im Clavivir- noch in realen Simulationsmodell statt, allerdings im randomisierten Modell. Ähnliches gilt für das Störverhalten ‚Hereinrufen‘, was zwar selten vorkam, aber theoretisch hätte identifiziert werden können. Dies führt zu einer weiteren Subgruppe dieser Kategorie: Störverhalten, die zwar vorhanden sind, aber deren audiovisuelle Ausgestaltung anders erwartet worden ist. Dies betrifft unter anderem die Unterhaltungen zwischen den SuS, bei denen eine Variation in der Lautstärke („laute Unterhaltungen“ oder „geflüsterte Unterhaltungen“) erwartet wurde. Einschränkend sei erwähnt, dass Flüstern in der sprachlichen Umsetzung der Mockup-Videos stattfand. Sehr laute Unterhaltungen waren allerdings selten. Die häufigen Nennungen sind ein Indiz dafür, dass eine Trennung der recht großen Kategorie „Unterhaltung“ sinnvoll wäre. Dies könnte durch einen Subalgorithmus realisiert werden, welcher bestimmt, ob eine Unterhaltung, ähnlich wie der Meldealgorithmus, als „geflüstert“, „Tischlautstärke“ oder „sehr laut“ zu klassifizieren ist. Ein solcher Algorithmus müsste fein abgestimmt werden. So wäre es unrealistisch, dass eine Schülerin auf eine geflüsterte Unterhaltung mit einer sehr lauten Unterhaltung reagieren würde. In diese Kategorie fällt auch das „Vorsagen von Antworten“. Dies ist in Clavivir implementiert durch das Hereinrufen. Auch hier ist ein Subalgorithmus denkbar, der steuert, ob die SuS die Antwort laut oder geflüstert hereinrufen bzw. vorsagen. Allerdings sollte aus didaktischen Gründen in der Simulation klar erkennbar sein, dass ein solches Verhalten stattfindet, sonst laufen die Partizipantinnen Gefahr es zu übersehen. Von einer Implementierung sollte bei der aktuellen Zielgruppe (Novizinnen im Feld) eher Abstand genommen werden.

Zwei Personen nannten „auf Zettel schreiben“ und implizierten damit ei-

ne Weitergabe von Antworten oder eine private schriftliche Unterhaltung. Da das visuell verwandte Störverhalten „Ins Heft malen (unaufgefordert)“ existiert, ist eine Verwechslungsgefahr in der audiovisuellen Ausgestaltung der beiden Störverhalten relativ groß. Zudem ist die Weitergabe und das Empfangen von Zetteln grafisch recht aufwendig zu animieren. Ungewöhnlich ist die Nennung von motorischer Unruhe („Zappeln“) durch eine Versuchsperson, die in eine Gruppe mit 3D-Visualisierung eingeteilt wurde. In der 3D-Visualisierung fand die Darstellung von motorischer Unruhe durch verschiedene Animationen (Am Kopf kratzen, mit den Haaren spielen, Zappeln) statt. Es scheint so zu sein, dass nicht alle als motorische Unruhe dargestellten Verhaltensweisen als solche identifiziert wurden. Einmal wurde als fehlendes Störverhalten „Reaktion auf Beleidigung“ genannt, welche der Clasivir-Störverhaltensalgorithmus allerdings bereits adressiert. Findet eine Beleidigung statt, so besteht die Chance, dass die so beleidigten SuS verbal oder körperlich auf die Beleidigung reagieren. Im konkreten Fall ist dies allerdings im Clasivir-Simulationsmodell nicht geschehen, da der Schwellenwert zum Auftreten des Reaktionsverhaltens unterschritten wurde.

In sechs Fällen wurden Verhalten genannt deren Klassifikation als Störverhalten fraglich ist. Dazu zählt mit sieben Nennungen „Frage von Schülerin, ob sie auf Toilette gehen darf“. Auch wenn eine solche Frage ggf. die Lehrkraft stört, so ist sie nicht als Störverhalten zu klassifizieren. Zudem ist ein solches Verhalten nicht damit verknüpfbar, wie es um die adaptive Passung der Aufgaben und die daraus resultierende Langeweile der SuS steht - was der modus operandi des Clasivir-Störverhaltensalgorithmus' ist. Eine andere Nennung war „etwas im Rucksack suchen“. Auch dies konstituiert vielleicht, dass sich eine Lehrkraft von den entstehenden Geräuschen gestört fühlt, aber es ist kein Störverhalten im klassischen Sinne. Außerdem kann das Kramen im Rucksack ein aktiv gewolltes Verhalten darstellen, wenn ein Arbeitsheft oder ein Buch herausgeholt werden soll. Dies würde in der konkreten programmier-technischen Ausgestaltung des Simulationsmodells überaus komplex werden (*bloated models*). Deswegen sei in diesem Fall von einer Implementierung abgeraten. Verwandt mit diesem Thema ist die Nennung „fehlende Materialien“, was kein situationsnahes Störverhalten konstituiert – ein Fehlverhalten fand hier zwar statt, ist aber im distalen Schulkontext anzusiedeln. Eine Implementation ist sowohl grafisch schwierig darzustellen als auch konzeptionell nicht durch das Simulationsmodell zu begründen. Daher sollte auch hier auf eine Implementation verzichtet werden. Das einmal genannte „Essen oder Trinken“ kann je nach verhandelten Klassenregeln ein Störverhalten konstituieren, kann aber konzeptionell schlecht in den bestehenden Algorithmus zum Auftreten von Störverhalten integriert werden, weswegen es besser ist dieses Verhalten nicht abzubilden.

Weiterhin wurden Verhalten genannt, die klar zu den Störverhalten zu rechnen sind, aber nicht in Clasivir auftauchen. Ein Beispiel dafür ist „Beschäftigung/Spielen mit anderen Gegenständen“, deren Implementation in späteren Versionen überlegenswert ist. Im Simulationsmodell des geplanten Clasivir 2.0 wurde darauf verzichtet. Sollte für dieses Verhalten eine Animation entwickelt werden, müsste entweder ein Gegenstand spezifiziert werden oder eine Batterie von Animationen mit unterschiedlichen Gegenständen erzeugt werden. Diese Animationen wären zudem durch die nötigen feinen Handbewegungen und externen Gegenstände schwierig zu generieren. Es ist aber denkbar in kommenden Versionen eine Reihe von sinnvollen Verhalten abzubilden, die SuS beim Spielen mit verschiedenen Gegenständen zeigen. Ähnliche Probleme in der Umsetzung der Animationen würde das einmal genannte körperlich-aggressive Störverhalten „Mit Papier werfen“ produzieren, womit vermutlich Papierkugeln oder Papierbälle gemeint sind. Ein solches Verhalten ist theoretisch in das Simulationsmodell integrierbar, allerdings ist es recht nahe an dem bereits implementierten Verhalten „Papier zerreißen“ (als Indikator für Gewalt gegen Gegenstände).

Weiterhin wurden abstrakte Verhalten genannt, die eher die Auswirkungen des Verhaltens beschreiben, z.B. „Störverhalten auf das die ganze Klasse reagiert und es dadurch laut und unruhig wird“. Innerhalb des Level-Videofundus' waren solche Störverhalten, wenn auch sehr selten, zu beobachten. Das Simulationsmodell in Clasivir müsste für eine solche Implementierung aber beträchtlich erweitert und umgeschrieben werden. Auch hier ist eine Integration in das Simulationsmodell zum gegenwärtigen Zeitpunkt nicht anzuraten. Ähnlich abstrakt ist die Nennung von „Diskussionen auf Ermahnung“. Während ein solches Verhalten häufig im Schulalltag zu beobachten ist, stellt sich die Frage nach der konkreten Implementierung. Es ist denkbar mittels einer probabilistischen Funktion zu versuchen solche Diskussionen zu erzeugen, aber durch die Vielzahl an möglichen Störverhalten ist die Umsetzung der Sprache innerhalb der Simulation entweder überaus kompliziert oder so sehr komplexitätsreduziert, dass zweifelhaft ist, ob dies einen didaktischen Nutzen bringen würde.

Zuletzt gibt es noch eine Reihe von Äußerungen (N=5), die sich nicht auf neue Störverhalten beziehen, die Häufigkeit der beobachteten Störverhalten kritisieren. Dies wurde bereits quantitativ gesondert erhoben und ist für eine qualitative Analyse weniger interessant.

Neben der Bewertung der Störverhalten hatten die Partizipantinnen die Gelegenheit generelles Feedback zu den gesehenen Mockup-Videos oder zum Clasivir-Konzept zu geben. Diese Antworten lassen sich etwas schwerer inhaltlich clustern, aber es gibt drei Themenkomplexe. Zum einen wurde auf die konkrete Aufbereitung der Mockup-Videos Bezug genommen: Es gab den

Wunsch von einer Person aus der 2D-Gruppe, die angab, dass es helfen würde, das Video mit animierten dreidimensionalen Figuren zu beurteilen. Dies wurde in der Anlage der Studie bereits berücksichtigt. Zweimal wurde genannt, dass die Sortierung der SuS-Zahlen in den 2D-Gruppen unübersichtlich (siehe [Abb. 20](#), S. 200) war. Die SuS wurden von links nach rechts und von unten (Lehrkraftposition) nach oben benannt, wobei S21 etwas aus der Reihe fällt, da dieser erst später die Klasse betritt. Dennoch ist die Kritik valide, hier wäre eine andere, konsistentere Benennung hilfreich gewesen. Damit zusammenhängend wurde einmal erwähnt, dass die „S?-Schüler*innen absolut keine Interaktionen aufzeigen wie ein Großteil der anderen Schüler“ und dieses Phänomen die Aufmerksamkeit des Studierenden eher gebündelt habe wie das Verhalten der anderen SuS. Das Problem der S?-SuS ist in der Methodenkritik bereits thematisiert worden. In einem zweiten Themenblock führten die Partizipantinnen Überlegungen zum Simulationsmodell an. Eine Person verweist z.B. darauf, dass die Wartezeiten bei SuS beachtet werden müssen. Diese Person war einer Gruppe, die Clasivir und das reale Modell miteinander verglich. Während in Clasivir Wartezeiten fest einprogrammiert sind, werden in der realen Stunde Antworten zum Teil sofort gegeben. Es ist auch möglich, dass damit gemeint war, dass die Wartezeiten im Clasivir-Modell zusätzlich zu knapp kalkuliert waren. Es handelt sich hierbei um eine einzelne Aussage, deren Validität empirisch zu überprüfen wäre. Es wäre eine gute Idee gewesen, auch die Fidelität der Wartezeiten innerhalb des Fragebogens quantitativ zu erfassen, was allerdings nicht antizipiert wurde. Weiterhin diesem Thema zuzuordnen ist die Aussage „Das Klopfen auf den Tisch könnte mit Geräuschen als Störfaktor präsen-ter werden“. Damit ist vermutlich eine Kritik der audiotecnischen Umsetzung verschiedener Störgeräusche, die den Tisch beinhalten, gemeint. Eine Aussage bezieht sich explizit auf das randomisierte Modell und weist darauf hin, dass „es zu unrealistisch [war], wie viel sich beleidigt wurde und wie häufig telefoniert wurde“. Der hier qualitativ wahrgenommene Unterschied lässt sich inferenzstatistisch quantitativ bestätigen.

Ein drittes Thema ist generelles positives oder negatives Feedback zur Idee von Clasivir. Fünf der Aussagen äußern sich positiv zur generellen Idee von Clasivir. Eine Aussage, die mit dem vorherigen Thema verwoben ist, lautet „Ich habe beide Unterrichtsszenarien als sehr realistisch bewertet. Ich bin der Meinung, dass Video 1 realistisch eine eher leistungsstärkere Lerngruppe abbildet und Video 2 eine lernschwächere Lerngruppe mit vermehrten Störungen.“, wobei die Person einer Gruppe zugeteilt war, die Clasivir und das reale Simulationsmodell miteinander verglich. „Viel Potential“ bescheinigt ein anderer Studierender dem Projekt, ein weiterer fordert „Bitte Weiterentwickeln!“. In den drei eher negativ oder verhaltenen Kommentaren spielt der

Vergleich mit authentischen Hospitationen eine große Rolle, z.B. „Ich finde, dass eher mehr Praxisphasen als künstlich simulierte Situationen in der Lehrkräftebildung genutzt werden sollten. Viele Studierende arbeiten ja sowieso schon an Schulen und erleben authentische Situationen.“ Zur Diskussion über den ergänzenden potentiellen Mehrwert von Simulationen gegenüber realen Hospitationen sein an [Abschnitt 2.4](#) (S. 16) verwiesen. Allerdings muss bei der Entwicklung von Clavivir immer mitgedacht und kommuniziert werden, dass es sich um eine Ergänzung realer Unterrichtserfahrungen handelt und nicht um einen Plan diese zu ersetzen. Sehr treffend schreibt eine Person hierzu: „Vor der Umfrage war ich solchen Lernmethoden noch negativ gegenüber eingestellt, jedoch haben mich die beiden Videos überzeugt, dass solche Methoden durchaus sinnvoll sein können. Jedoch sollten Lehramtstudenten diese Möglichkeit nur zstl. haben, die Erfahrung mit ‚echten‘ SuS muss weiterhin bestehen, sie ist hierdurch nicht ersetzbar, zumal sich hiermit keine Bindung zu den SuS aufbauen lässt.“ Andere Personen benennen, dass diese Simulation vor allen bei Themen wie *classroom management* und Aufgabenadaptivität helfen kann, jedoch ihre Grenzen hat und keine echten Hospitationen ersetzen kann. Im Gegensatz dazu fand eine andere Person den Umgang mit VR zwar eine „angenehme Abwechslung zu Texten oder Fallbeispielen“, kann es sich aber für sich persönlich „nicht vorstellen, anhand von VR zu lernen oder es in meinen Lernalltag einzubringen“.

Ein viertes Thema ist der Zweifel an der eigenen Kompetenz bei der Beurteilung der Simulation. Eine Person schreibt: „Es ist als Student*in schwer einzuschätzen welches Unterrichtsverhalten einer Klasse realistisch ist, da sowohl jede Klasse unterschiedlich reagiert, als auch Student*innen kaum Erfahrung haben.“ Hierbei handelt sich um ein Grundproblem in der Methodologie. Studierende sind die Zielgruppe der Simulation, insofern ist es wichtig, dass die Simulation von dieser Gruppe auch als realistisch angesehen wird. Gleichzeitig sind Studierende Novizinnen im Feld und demzufolge nicht in der Lage aus einem großen Fundus an Erfahrungen zu schöpfen um die Fidelität des Simulationsmodells adäquat zu beurteilen. Daher ist es wünschenswert in zukünftigen Forschungen auch erfahrene Lehrkräfte und Fachdidaktikerinnen zu beteiligen. Verbunden mit dem letzten Themenkomplex ist folgende Antwort zu sehen: „Es fiel mir schwer in die Situation einzutauchen. VR-Aspekt fehlte mir etwas. Ich glaube trotzdem, dass hier viel Potential drin steckt.“ Eine weitere Aussage verknüpft das Thema um die Einschätzungskompetenz mit generellen Erwägungen zum Simulationsmodell, die sich so lesen, dass auch ein randomisiertes Modell angemessen sein könnte: „Kommt auch immer auf die Stimmung der Schülerinnen und Schüler drauf an. Nicht jeder Tag ist gleich, mal gibt es wenig Störungen und viel Meldebeteiligung und mal ist es genau umgekehrt.“ Trotz dieser Aussage belegt die inferenzsta-

tistische Prüfung, dass ein randomisiertes Modell als statistisch signifikant schlechter wahrgenommen wird als das Clasivir-Modell, bei welchem auch kein Unterschied zur realen Stunde feststellbar ist.

8 Zusammenfassung und Diskussion

8.1 Einordnung der Ergebnisse

Wie die inferenzstatistische Prüfung zeigt, gibt es keinen statistisch signifikanten Unterschied in der wahrgenommenen Fidelität zwischen dem Clasivir-Simulationsmodell und dem übertragenen Simulationsmodell der realen Stunde. Es kann davon ausgegangen werden, dass das Clasivir-Simulationsmodell die notwendige Fidelität besitzt um in einer VR-Simulation für Lehrkräfte eingesetzt zu werden. Das wesentliche Forschungsziel ist damit erreicht. Die Prüfung von Hypothese zeigt, dass das Clasivir-Simulationsmodell statistisch signifikant besser ist als das randomisierte Modell. Die standardgemäß als moderat angenommene Effektgröße von $d=.634$ ist in diesem Kontext vergleichbar hoch. Dass das randomisierte Simulationsmodell eine recht große Varianz besitzt, lässt sich durch zwei Faktoren erklären. Zum einen methodologisch durch die geringere Anzahl an zugewiesenen Versuchspersonen. Demzufolge ist auch der Unterschied zwischen Clasivir- und randomisiertem Versuchsmodell weniger aussagekräftig als eine wahrgenommene Äquivalenz des realen und des Clasivir-Simulationsmodells. Zum anderen – und dies lässt sich sowohl quantitativ als auch qualitativ belegen – haben Studierende verständlicherweise Probleme die Fidelität der SuS-Aktionen innerhalb der Simulation professionell einzuschätzen.

Es fällt auf, dass die deskriptiven Ergebnisse der Simulationsmodellfidelität des realen Simulationsmodells etwas niedriger als erwartet ausfallen. Mit $M=2.97$ ($SD=.49$) bewerten die Partizipantinnen dieses Modell nur als „eher realistisch“. Diese Ergebnisse sind durch verschiedene Faktoren zu interpretieren. Zum einen sorgt die Übertragung in ein 2D- oder 3D-Format generelle Verluste der wahrgenommenen Fidelität, die von den Partizipantinnen nicht vollständig ausgeblendet werden können. Zum anderen sorgte die Anlage der Untersuchung für ein Framing, das Störverhalten stark in den Vordergrund rückte. In der realen Stunde, auch wenn sie eine der auffälligeren Stunden bezüglich Störverhalten war, fand aber, insbesondere im ersten Block, relativ wenig Störverhalten statt. Die Autorität der Kamera und die Invasivität einer Videoaufzeichnung führten sicherlich dazu, dass generell weniger Störungsverhalten im Level-Videofundus zu beobachten waren. Dies sorgte vermutlich dafür, dass die Werte etwas niedriger als erwartet ausfallen.

Der von einigen anderen Simulationen gewählte Weg, SuS-Verhalten über randomisierte Zuweisungen zu realisieren erweist sich damit als keine gangbare Methode, wenn man versucht eine als realistisch angesehene Simulation zu konstruieren. Von einer solchen Implementierung ist in künftigen Simulationen also abzuraten.

Interessant sind auch die negativen Ergebnisse, unter anderem, dass es zwar einen leichten deskriptiven Trend dazu gibt, die 3D-Visualisierung der Mockup-Videos mit einer geringeren Bewertungsschwierigkeit zu assoziieren, dies bestätigte sich jedoch inferenzstatistisch nicht. Auch qualitativ wurde im Rahmen der Nachbesprechung der Digi_Gap Seminare mehrfach erwähnt, dass es nicht immer leicht fiel, genau zu identifizieren, welches Verhalten mit den entsprechenden 3D-Animationen indiziert werden sollte. Vor dem Hintergrund dessen, dass die Animationen mit äußerst großem Aufwand erstellt und zugewiesen worden sind, lässt sich für kommende Untersuchungen der Simulationsmodellfidelität konstatieren, dass eine Aufbereitung in 3D-Form nicht unbedingt erforderlich ist. Da vor allem die Zuweisung der Animationen, gerade in Verbindung mit interagierbaren Gegenständen sehr viel Zeit verschlungen hat, scheint die sehr viel einfachere Umsetzung als 2D-Darstellung der bessere Weg zu sein.

Weiterhin ist es erfreulich, dass die wahrgenommene Nützlichkeit der Simulation nicht vom Geschlecht der Partizipantinnen und Partizipanten abhängt. Ein befürchteter Gender-Bias ist somit nicht gegeben. Die meisten Personen halten einen Einsatz von VR-Simulationen in der Lehrkräftebildung für nützlich, es gibt also einen Bedarf das Forschungsfeld weiter auszubauen.

Es erscheint theoretisch zunächst plausibel, dass Personen, die häufiger Videospiele auf einem Computer oder eine Konsole spielen es weniger schwierig finden das Simulationsmodell zu bewerten, da diese gewöhnt sind mit Spielmodellen umzugehen, die einen ähnlichen Grad an Komplexität besitzen. Es wurde angenommen, dass diese Gruppe in der Lage ist, zu durchschauen oder zumindest zu extrapolieren was die zugrundeliegenden Mechanismen der Simulationsmodelle sind. Damit sollte diesen Personen die Bewertung der Fidelität einfacher fallen. Dies war nicht der Fall. Eine Erklärungsmöglichkeit dafür wäre, dass viele Videospiele ihre Funktionsmechanismen nicht transparent machen und dieses Wissen demzufolge nicht in vielspielenden Personen vorhanden ist. Vielmehr könnten Personen, die in ihrer Freizeit selbst Spiele programmieren oder vorrangig Spiele spielen, die auf das Lernen von Simulationsmodellen abzielen (wie Globalstrategiespiele), die gesuchte Personengruppe sein, der die Bewertung leichter fällt. Es wäre zudem noch zu prüfen, ob sich Personen in ihrer wahrgenommenen Bewertungskompetenz über die verschiedenen Ausbildungsstufen der Lehrkräftebildung hinweg unterscheiden.

Die qualitativen Untersuchungen zeigen, dass die abgebildeten Störverhalten so gut wie alle in der Realität beobachtbaren Störverhalten abbilden und Kritik sich oftmals nur auf das konkrete Auftreten der Störverhalten bezieht.

8.2 Grenzen der Untersuchung

Es ist wichtig zu berücksichtigen, dass die Gruppe der Lehramtsstudierenden aus zwei Gründen nicht repräsentativ für alle Lehramtsstudierenden ist. Erstens wurde die Gruppe primär aus dem Digi_Gap Seminar (oder anderen *digital learning*-Seminaren bei den incentivierten Partizipantinnen) rekrutiert. Bei den Partizipantinnen ist anzunehmen, dass Personen die solche Seminare wählen, von Anfang an eine höhere Affinität zu VR, Simulationen und digitalen Technologien im Allgemeinen besitzen. Die FIDEC- und TK-Werte der Partizipantinnen sind mit $M=.94$ ($SD=.71$, FIDEC) und $M=.8$ ($SD=.71$, TK) als am oberen Ende der jeweiligen Skalen im Vergleich zur Grundgesamtheit anzusehen. Zweitens ist das gezogene Sample nicht groß genug, um die Ergebnisse zu verallgemeinern.

Wie bereits schon an anderer Stelle beschrieben, kann die gewählte Methodologie nicht alle Aspekte des Simulationsmodells messen. Insbesondere der diversitätssensible Namensvergabealgorithmus konnte nicht validiert werden, ebenso wenig wie der Emotionswechsel-Algorithmus. Es ist bezüglich des Emotionswechsel-Algorithmus fraglich, ob eine quantitative Untersuchung dieser Hintergrundkomponente wirklich gewinnbringende Erkenntnisse liefern würde. Ein Fragebogen-Design scheint aufgrund der anderswo (siehe [Abschnitt 6.4.1](#), S. 207) diskutierten Limitationen der visuellen Repräsentation von Emotionen in der Simulation nicht geboten. Stattdessen sind diese Fragestellungen eher qualitativ im Rahmen von Fokusgruppeninterviews zu adressieren.

Methodologisch arbeitet diese Studie mit einfachen statistischen Verfahren, was durch folgende Überlegungen zu begründen ist: (1) Forschungspragmatisch verschlang bereits die Entwicklung des Simulationsmodells, des Prototypen und der Mockup-Videos einen Großteil der allokierten Zeit. (2) Es existierte bislang noch kein erprobtes Forschungsdesign zur Prüfung der Fidelität eines Simulationsmodells, insbesondere nicht im Kontext der Lehrkräftebildung. Daher ist es ein Beitrag dieser Arbeit, gezeigt zu haben, dass grundsätzlich eine Erfassung der Fidelität des Simulationsmodells möglich ist, bevor der Simulator konkret implementiert worden ist. Außerdem wurden Möglichkeiten aufgezeigt, in welcher Form dies umsetzbar ist. (3) Es wurde angenommen, dass die Partizipantinnen ca. eine Stunde zur Bearbeitung des Fragebogens brauchen, weswegen auf längere Fragebatterien verzich-

tet wurde. Aus den deskriptiven Daten geht hervor, dass die tatsächlichen Bearbeitungszeiten weit niedriger waren. Man hätte retrospektiv die Fidelitätsaspekte statt durch drei ordinal skalierte Ratings differenzierter erfassen können. Dies hätte (4) im Rückschluss erfordert ein – ebenfalls zu validierendes – Instrument zur Messung der Fidelität zu entwickeln, was den Rahmen dieser Arbeit gesprengt hätte.

Die Untersuchung zielte darauf ab, ein Simulationsmodell zu entwickeln und dies zu evaluieren, bevor eine Simulation vorliegt. Dazu wurde ein eigenes Forschungsdesign verwendet, welches auf Mockup-Videos basiert und prüft, ob das Modell als ökologisch valide wahrgenommen wird. Nur qualitativ geprüft wurde, wie das Modell ggf. verbessert werden kann. Während sich dieser Ansatz – wenn auch nicht methodologisch problemfrei – bewährt hat, muss darauf hingewiesen werden, dass Fidelitätskonstrukte von fertiggestellten *educational simulations* noch komplexer zu konzeptionalisieren sind und z.T. davon abhängen welche Granularität gewählt wurde. Im Falle einer Klassenraumsimulation im öffentlichen Unterricht sind viele benutzte Aspekte wie Melde-, Störungs- und Antwortverhalten übertragbar.

8.3 Ausblick

In künftigen Untersuchungen sollten qualitative Fokusgruppeninterviews durchgeführt werden, um weitere Hinweise zur iterativen Verbesserung des Simulationsmodells zu gewinnen. Modellierete kausale Zusammenhänge könnten diskutiert und in Folge nachjustiert werden. Außerdem würde mit einer Zusammensetzung dieser Fokusgruppeninterviews aus Feldexpertinnen wie erfahrenen Lehrkräften, Bildungsforscherinnen und Fachdidaktikerinnen die Simulationsmodellvalidierung erweitert und professionalisiert. Aktuell werden die Fidelitätsunterschiede der Simulationsmodelle ausschließlich durch Ratings von Lehramtsstudierenden untersucht und bilden damit ab, wie die Zielgruppe den Realitätsgrad des Simulationsmodells einschätzt.

Während diese Arbeit verfasst wurde, gab es bereits erste Entwicklungen zur konkreten Umsetzung von Clasivir als VR-Simulation. Im Rahmen des Teilprojektes „VR-Lernsituationen“ des QLB-Projektes „Digi_Gap“ konnte eine erste Version von Clasivir nicht nur konzipiert und erstellt, sondern auch schon konkret an Studierenden erprobt werden. Diese Version von Clasivir, auch Clasivir 1.0 genannt, basiert nicht auf einer komplexen algorithmischen Steuerung, sondern wurde als reiner deterministischer Handlungsbaum (Typ 1) konzipiert, verfügt allerdings über vollständige VR-Kapazitäten. Als technischer Partner für die Umsetzung der Programmierung konnte die Firma NMY gewonnen werden, die alle universitär vorgegebenen Animationen und Pfade in die Unity Engine einpflegte. Ein Screenshot der aktuell verwen-



Abbildung 31: Screenshot von Clasivir 1.0, Quelle: Autor

deten Version von Clasivir 1.0 findet sich in [Abb. 31](#). Alle Teilnehmerinnen der Digi_Gap Seminare hatten die Chance eine *classroom-management*-intensive Lernsituationen mittels HMD-Brille zu erleben. In dieser Situation streiten sich zwei SuS und es kommt zu einer körperlichen Rangelei. Die Lehrkraft muss hier intervenieren und versuchen mit der Situation pädagogisch passend umzugehen. Neben Clasivir 1.0 wurden in der VR-Phase des Seminars auch andere VR-Applikationen benutzt. Außerdem zählte eine dedizierte Vor- und Nachbereitung im Sinne eines BAR und AAR-Feedbacks zur Konzeption des Seminars.

Unity wurde als Simulationsengine mit Bedacht gewählt: Wir erhoffen uns Synergieeffekte mit anderen Projekten aus der allgemeinen Psychologie, die Eye-Tracking Funktionen für die HMDs bereits nutzen. Die Eye-Tracking Funktionen liegen als Codebausteine bereits in Unity vor. In kommenden Iterationen von Clasivir ist geplant die Augenbewegungen der Lehrkräfte zu tracken. Das Innovative an dieser Technik ist, dass wenig Aufbereitung der Eye-Tracking-Daten erforderlich ist. Setzt man Eye-Tracking Brillen in realen Unterrichtssituationen ein, ist eine aufwendige Nachbearbeitung der Daten erforderlich. In der geplanten Version von Clasivir wäre es möglich die Blicke der Partizipantinnen automatisiert zu erfassen. Es könnte in Echtzeit gemessen werden, wie lang der Blick der Partizipantinnen auf spezifischen SuS oder auf Klassenraumobjekten (z.B. Tafel, Fenster) ruhte. Damit wären ganz neue Arten an automatisierten Feedbacksystemen möglich, die insbesondere in der Schulung der professionellen Unterrichtswahrnehmung einen großen Nutzen

versprechen.

Es ist wünschenswert die aus forschungspragmatischen Gründen eher niedrige Zahl an Partizipantinnen auszuweiten. Da inzwischen ein Prototyp von Clasivir 1.0 vorliegt, ist es denkbar bald eine große Anzahl an Studierenden mit dem geplanten Clasivir 2.0 (welches das Simulationsmodell implementiert) zu konfrontieren. An dieser Stelle könnte man dann nicht nur die Fidelität replikativ absichern und um andere Aspekte wie UI/UX oder Grafik erweitern, sondern auch messen, inwiefern sich handlungsbezogene Kompetenzen durch Simulationsnutzung im Vergleich zu einer Kontrollgruppe verbessern. Dies könnten wahrgenommene Verbesserungen im Umgang mit *classroom management* genauso wie Digitalisierungskompetenzen oder die professionelle Unterrichtswahrnehmung sein. Wenn sich bestätigt, dass Simulationsnutzung zu statistisch signifikanten Verbesserungen führt, kann über eine breite Adaption der VR-Klassenraumsimulation nachgedacht werden.

Es lohnt sich, weitere Möglichkeiten der iterativen Verbesserung des Clasivir-Simulationsmodells in den Blick zu nehmen. Die rasante aktuelle Entwicklung von *machine learning* Ansätzen wie GPT4 macht es denkbar, die in den Mockup-Videos verwendeten linguistischen Anteile über GPT4 oder ähnliche *large language models* zu realisieren, während das Simulationsmodell den Prompt dafür liefert. Würde also das Simulationsmodell in Clasivir vorhersagen, dass ein SuS die richtige Antwort auf eine Frage gibt, würde der Prompt lauten: „Versetze dich in die Lage einer Schülerin der \$Klasse, \$Jahrgangsstufe in einem \$Schultyp. Die Schülerin ist eher zurückhaltend, aber nicht neurotisch und sehr verlässlich. Beantworte die folgende Frage: \$Frage. Die Antwort der Schülerin soll dabei korrekt sein.“ Die konkrete Umsetzung und Implementierung der Prompts ist selbstverständlich ein vollständiges eigenes Forschungsprojekt und wurde hier nur zum Zwecke der Illustration eingefügt. Mittels Speech-To-Text und Text-to-Speech würde es damit möglich eine reale Interaktion mit virtuellen SuS zu führen, die sich auf Basis des Clasivir-Modells verhalten und deren konkrete linguistische Ausgestaltung durch ein *large language model* realisiert würde. Diese Art der Implementierung würde das Simulationsmodell stärker zum Typ 4 verschieben, jedoch ohne viele der in der Taxonomie genannten Nachteile, allen voran die Abstraktheit der Simulation. Im Rahmen einer Begleitforschung wäre dabei kritisch zu überprüfen, ob eine solche Verwendung nicht ein Einfallstor für Stereotype und Vorurteile ist.

Eine weitere Möglichkeit zur Verbesserung des Simulationsmodells ist es mit *big data* aus authentischen Unterrichtsvideos zu arbeiten, um die Handlungsalternativen nicht mehr über *behavior trees*, sondern direkt über *machine learning* zu steuern. Das große Fragezeichen an dieser Idee ist es, ob es

in Zukunft möglich sein wird eine ausreichende Anzahl an Daten zu generieren. Hier sind nicht nur datenschutzrechtliche Aspekte zu berücksichtigen, sondern auch die Realisierung der Datenerfassung. Werden alle Daten per Hand kodiert wie in der Studie von Appel (2015), ist eine für *machine learning* Algorithmen sinnvoll nutzbare Datenmenge fast unmöglich zu erzeugen, außer mit enormen Forschungsaufwand. Stattdessen ist zu hoffen, dass die Verhaltensbeobachtung von SuS durch Gesichtserkennungsansätze wie OpenCV automatisierbar wird. Allerdings erscheint es schwierig die Aktionen der Lehrkraft automatisiert zu kodieren. Auf welche kognitive Dimension oder Wissensdimension eine Frage abzielt, ist bislang nur durch manuelles Kodieren möglich. Ohne die Daten der Lehrkraftinteraktion reproduziert ein undifferenziert verwendetes Modell von *machine learning* lediglich Verhalten von SuS ohne Berücksichtigung der Lehrkraftaktionen, was für einen sinnvollen Einsatz sicherlich hinderlich wäre und eher Parallelen mit dem randomisierten Simulationsmodell aufzeigt. Es sind weitere Forschungsarbeiten nötig, um einen Einsatz von *machine learning* zum Ersatz des auf *behavior trees* basierenden Clasivir-Simulationsmodells überhaupt theoretisch zu konzeptionalisieren.

Bei dieser Studie handelt es sich um eine Vorarbeit, auf die weitere Forschungsarbeiten aufbauen sollen. Eine Untersuchung, ob durch Benutzung des fertigen Simulators tatsächlich die handlungsnahen Kompetenzen von Lehrkräften steigen, steht noch aus. Erste Hinweise aus der Benutzung von Clasivir 1.0 zeigen, dass die Nutzung der Simulation mit einer statistisch signifikanten Erhöhung von Leistungsemotionen einhergeht (Glocker et al., 2023). Neben einem klassischen Prä-Post-Testdesign sollten Folgeuntersuchungen im besten Falle longitudinal angelegt sein und untersuchen, wie es um die Langlebigkeit der durch Simulationseinsatz potentiell hervorgerufenen Änderungen der handlungsnahen Kompetenzen bestellt ist. Auch eine Untersuchung darüber, ob SuS von der Simulationsschulung ihrer Lehrkräfte profitieren ist eine Frage, der es sich nachzugehen lohnt. Es ist wichtig nicht nur die handlungsnahen Kompetenzen in den Blick zu nehmen, sondern zu prüfen, ob sich z.B. TPACK-Komponenten durch Simulationseinsatz ändern. Kurz gesagt ist zu erwarten, dass ein Einsatz mit digitalisierungsbezogenen Werkzeugen schon während des Lehramtsstudiums einen Wechsel in der Einstellung und Medienkompetenz der Lehrkräfte erzeugen könnte, so dass im Ergebnis der Unterricht in Deutschland digitalisierungsaffiner wird. Die Ergebnisse von Hattie (2009) haben gezeigt, dass eine solche Veränderung weniger durch Strukturänderungen erzeugt werden kann, sondern die Lehrkraft im Zentrum steht. Daher ist es wünschenswert, VR-Seminare für die zweite und dritte Phase der Lehrkräftebildung zu konzipieren, so dass eine Schulung mit VR-Simulationen die größtmögliche Breitenwirkung erreichen kann.

Anhang

Screenshots und Abbildungen der Simulatoren

Aha! (Process) Classroom Sim



Abbildung 32: Screenshot von Aha! Classroom Sim, Quelle: Screenshot des Autors



Abbildung 33: Screenshot von Aha! Classroom Sim, Quelle: Screenshot des Autors



Abbildung 34: Screenshot von Aha! Classroom Sim, Quelle: Screenshot des Autors

Behavior Breakthroughs



Abbildung 35: Screenshot der Applikation Behavior Breakthroughs, Quelle: <https://medicalxpress.com/news/2018-07-behavioral-problems-iphone-application.html>



Abbildung 36: Screenshot der Applikation Behavior Breakthroughs, Quelle: <https://theappwhisperer.com/2011/02/the-newest-and-best-ios-apps-to-help-children-with-autism/>

Breaking Bad Behaviors

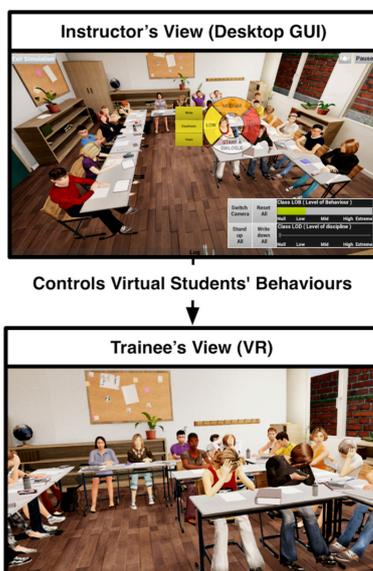


Abbildung 37: Screenshot von Breaking Bad Behaviors, Quelle: Lugin, Oberdörfer et al. (2018, S.625)

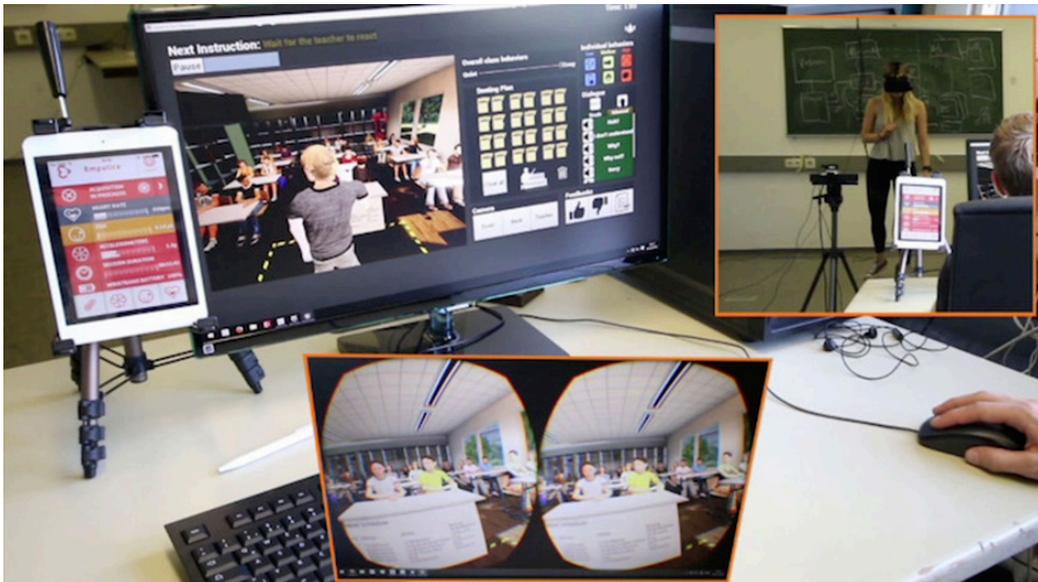


Abbildung 38: Screenshot von Breaking Bad Behaviors, Quelle: <http://hci.uni-wuerzburg.de/projects/breaking-bad-behaviors/>



Abbildung 39: Screenshot von Breaking Bad Behaviors, Quelle: Latoschik et al. (2016, S.317)

Der simulierte Klassenraum (SKR)



Abbildung 44: Screenshot des simulierten Klassenraums (SKR), Quelle: Südkamp und Möller (2009, S.166)

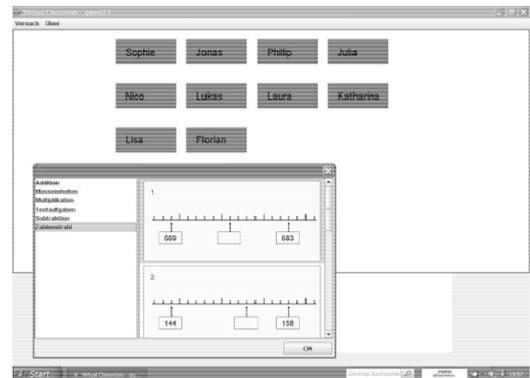


Abbildung 45: Screenshot des simulierten Klassenraums (SKR), Quelle: Südkamp et al. (2008, S.264)

Diversity Simulation



Abbildung 46: Screenshot der Diversity Simulation, Quelle: Manburg et al. (2017, S.131)



Abbildung 47: Screenshot der Diversity Simulation, Quelle: Manburg et al. (2017, S.136)

First Class



Abbildung 48: Screenshot des Simulators First Class, Quelle: <https://beyond.psu.edu/firstclass/>



Abbildung 49: Aufbau des Simulators First Class mit MS Kinect, Tablet und PC-Bildschirm, Quelle: Zidik (2015)

IVT (Interactive Virtual Training) / IVT-T (Interactive Virtual Training for Teachers)



Abbildung 50: Screenshot des Simulators IVT/IVT-T, <http://ascl.cis.fiu.edu/interactive-virtual-training.html>



Abbildung 51: Screenshot des Simulators IVT/IVT-T, Delamarre et al. (2017, S.402)

Kognito Interactive Simulations



Abbildung 52: Screenshot der Simulation At-Risk for Middle School Educators der Firma Kognito Interactive, Quelle: Albright et al. (2013b, S.1)



Abbildung 53: Screenshot der Simulation At-Risk for High School Educators, Quelle: <https://youtu.be/ISkaJIz0iNA>

Lesson Sketch / SimTeach / ThEMaT (Thought Experiments in Mathematics Teaching) I-III

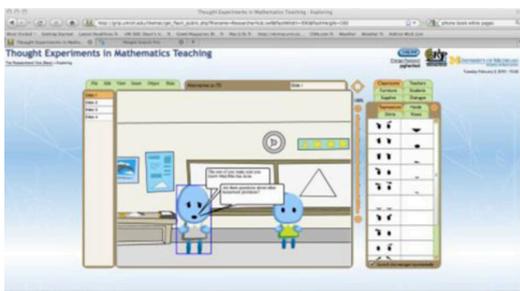


Abbildung 54: Screenshot der Backend-Komponente (LessonSketch) zur Konstruktion von Lerneinheiten, Quelle: Herbst et al. (2011, S.99)

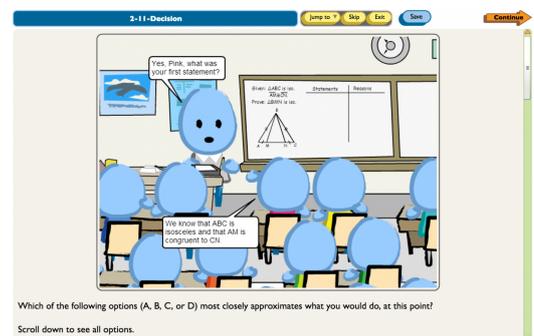


Abbildung 55: Screenshot der Partizipantinnen Perspektive der ThEMaT-Simulation, Quelle: <https://www.lessonsketch.org>

Mr. Land's Sixth Grade

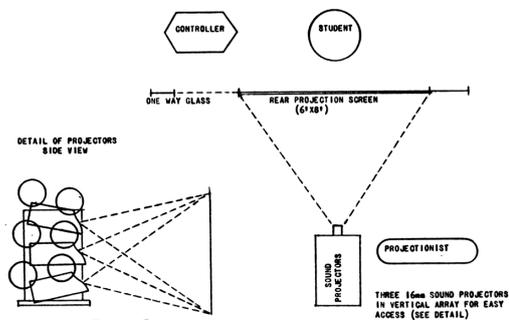


Abbildung 56: Aufbauskitze der Simulation Mr. Land's Sixth Grade, Quelle: Kersh (1963)

SimAula 1.0



Abbildung 57: Screenshot von SimAula 1.0, Quelle: <https://www.youtube.com/watch?v=lnq7IwipW4c>



Abbildung 58: Screenshot von SimAula 1.0, Quelle: <https://web.archive.org/web/20190126044150/http://www.simaulaproject.eu/>

SimAula 2.0

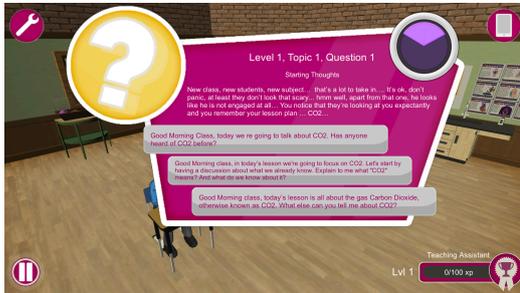


Abbildung 59: Screenshot von SimAula 2.0, Quelle: Autor



Abbildung 60: Screenshot von SimAula 2.0, Quelle: Autor

Flash SimClass

영역	매우 낮음	낮음	보통	높음	모두 0 이하
자기 존중					
자기 유능감	내가 잘 아는 것만 이나도 없어.	난 할 수 없어.	이 정도는 할 만해.	무엇을 해도 난 할 수 있어.	
자기 통제력					

Abbildung 61: Screenshot der flashbasierten Version von SimClass, Quelle: Cheong und Kim (2009, S.68)



Abbildung 62: Schaubild der möglichen graphisch dargestellten Interaktionen in der flashbasierten Version von SimClass, Quelle: Cheong und Kim (2009, S.69)

Web SimClass



Abbildung 63: Screenshot der webbasierten Version von SimClass, Quelle: Baek (2009, S.42)

SimClass II

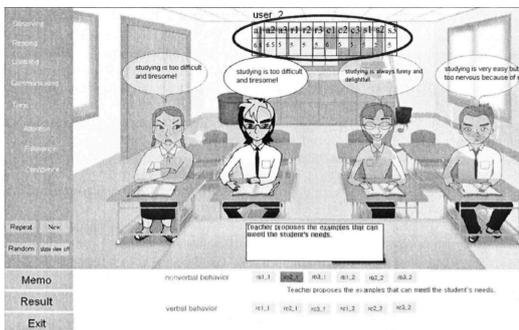


Abbildung 64: Screenshot von SimClass II, Quelle: Cheong und Kim (2009, S.68)

simSchool 1.0



Abbildung 65: Screenshot von simSchool 1.0, Quelle: <https://www.seriousgamemarket.com/2012/02/simschool-serious-games-as-flight.html>

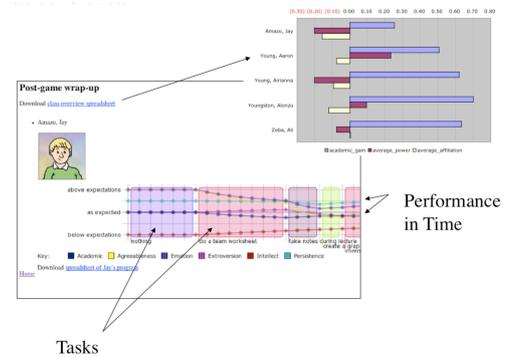


Abbildung 66: Screenshot des Feedbackscreens von simSchool 1.0 mit Veranschaulichung des zugrundeliegenden Algorithmus, Quelle: Gibson (2011, S.8)

simSchool 2.0

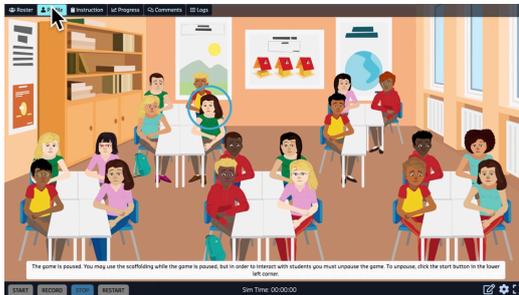


Abbildung 67: Screenshot von simSchool 2.0



Abbildung 68: Screenshot von simSchool 2.0

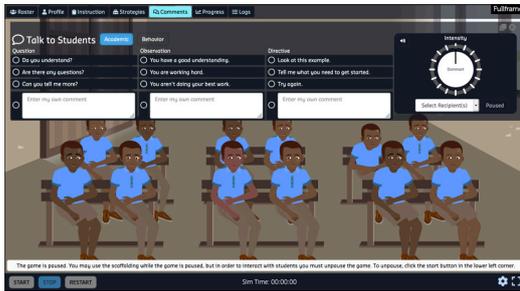


Abbildung 69: Screenshot von sim-School 2.0

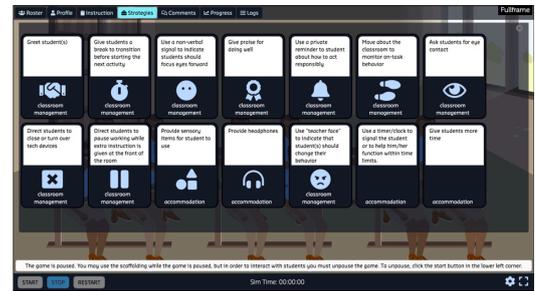


Abbildung 70: Screenshot von sim-School 2.0

Simfta / SimInClass



Abbildung 71: Screenshot von SimIn-Class (englischsprachige Version von Simfta), Quelle: <https://youtu.be/-FmwhD76o5s>

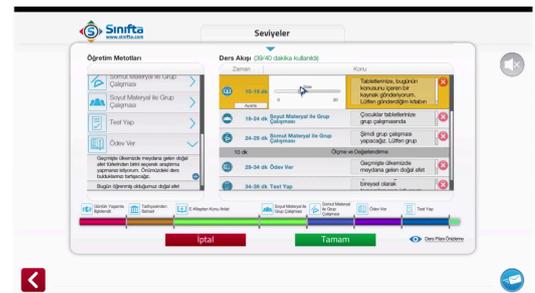


Abbildung 72: Screenshot der Stundenplanung in Simfta, Quelle: <https://youtu.be/-eiFLkSa65M>

STAR Simulator / TeachMe / TeachLivE (TLE) / SimLab / Mursion



Abbildung 73: Foto eines Interaktors, der virtuelle SuS über einen *motion-capturing* Anzug spielt, Quelle: Dieker et al. (2012, S.98)



Abbildung 74: Screenshot von TeachLivE/TeachMe, Quelle: O'Callaghan und Piro (2016, S.41)



Abbildung 75: Foto von TeachLivE-Simulationspartizipantinnen, die mit der Simulation interagieren, Quelle: O'Callaghan und Piro (2016, S.43)



Abbildung 76: Screenshot von Mursion, Quelle: Mursion-Homepage

SimTeacher

This situation is called: How Do I Know?

Elizabeth has been struggling since the first of the year in reading. Her kindergarten teacher noted that Elizabeth was not very interested in letters and did not progress as far as the others in the class. The K teacher felt Elizabeth needed to mature and would catch up. Now it is November and Elizabeth is still reading at a pre-primer level without showing much enthusiasm for improving her skills. You wonder if it's really a maturity problem or something else and decide to find out.



You must now make your first decision and provide justification for it:
Mark the circle next to the decision choice that you would most likely support and then provide your reasoning for choosing that choice in the box below.

You decide to pull Elizabeth aside for one-on-one testing and do a running record of her reading.

You feel that if immaturity is to be ruled out you should evaluate Elizabeth's behavior at other activities. You make a chart of the target behaviors, including on-task rates, ability to follow directions, and sharing and helping behaviors.

Why choose that?

After you take action here (and hit the Proceed button), you will be transported to the next episode of the story. You are advised not to stop until you've reached the end. However, your work will be saved if you do need to stop. If you start the uncompleted situation at a later time, you will be taken to the part of the situation where you stopped.

Abbildung 77: Screenshot eines deterministischen Entscheidungsbaums in SimTeacher, Quelle: Fischler (2007, S.46)

Tigerlake Simulation

Academic Problem: Consultant Scenario 1

This is a new teacher, Ms. Lila, when you are working with it's her third week on the job, and she has been assigned to teach the "low" math group. The fourth grade team, as you quickly assume, groups their students to teach Ms. Lila. Lila was excited to be in the "low" group, but because of a lack of confidence, she has been assigned this math class. Period 3 is your current concern. This classroom of 14 students meets after their lunch period each day, and despite her attempts to motivate them to do better, the students as a group do not seem engaged in the material Ms. Lila presents them. Currently, they are working on basic fraction operations and Common Core. She is just now attempting to use small groups within the class as an approach for teaching. She believes in guided discovery learning, but she has no experience with it. The student discontent, or apathy of each, is really starting to be a behavior management issue. Just yesterday Ms. Lila had to send August down to the vice principal office due to his increasing disruptive and overly defiant behavior toward her and his classmates.

You have begun to suspect that some of the problems may relate to the heavy demand for reading comprehension about many of the fraction problems are easy problems. The poor performance of two students, Matt and Quinn on the last chapter test over percentage story problems, led you to this hypothesis. Although the mathematics is simple, much of it is contained within lengthy word problems that might be difficult to organize for student work in language arts skills. Lila's parents have also contacted Ms. Lila by video-mail asking for an explanation of what is going on in this classroom; she has come home in tears more than once and has been afraid of speaking up because others might criticize her. You want to help her but things seemed here while she still has a chance with this class. What would you suggest as the teacher consultant?

TIPS TO KEEP IN MIND

1. Use "student files" to learn more about individual student patterns and any available details of personal background and/or behavioral history.
2. Be sure to carefully review and discuss the individual test results available on Arnie's. How do you make use of this assessment information to assist this teacher?
3. Use the "teacher's lounge" link to communicate with your other P-21 colleagues; leave messages for each other as you come up with hypotheses or important discoveries.
4. Use the "phone the principal" link to have a message for the school administrator. Make your questions specific, and he will reply within a day or two to your e-mail account.

Review the goals outlined in your blue content packet regarding this Present Theory 10 assignment. Do you see you have made use of all information available in the Tiger Lake School System AND that your write up reflects your command of all relevant topics discussed this scenario?

RESOURCES

- Present Theory 10
- Teacher's Lounge
- Phone the Principal
- Student Files

Abbildung 79: Screenshot der archivierten Tigerlake-Webseite, Quelle: Author

Student Characters

These fictional students can be referenced in your situations. Your real students are considered the teachers (i.e., SimTeachers) in this simulation, so these fictional students essentially become their students within this "virtual school" setting.

11 student crew members.

(Bobby) Robert Guerrero	Ariella Roden	Dammer Manson	Ariel Wei	Miriam Townsend	Aileen Hou
Jose Silveira	Marcia Tessicini	Radoslav Slavic	Todd Slavic	Yona Joo Slavic	

[Jumpstart a classroom of students.]
[Create an original student from scratch.]
[Repurpose another instructor's student.]

Teacher Characters

Fictional teachers could also be referenced in your situations. The fictional teachers displayed below essentially become colleagues of your SimTeachers.

6 teacher crew members.

(Bernie) Bernard Patterson	Peter Rivera-Vinil	Scarlett Worthington	Mahmoud Alkjar	Phyllis Spiceman	(DJ) Joshua D. Filkins
----------------------------	--------------------	----------------------	----------------	------------------	------------------------

[Jumpstart a school of teachers.]
[Create an original teacher from scratch.]
[Repurpose another instructor's teacher.]

Abbildung 78: Screenshot der archivierten SimTeacher Webseite, Quelle: <https://web.archive.org/web/20050118232051/http://simteacher.com/instructor/tour/characters.htm>

Situation 2

Your students worked on projects for three weeks, conducting research, reading various texts, and interviewing experts both face-to-face and over the Internet. You were excited about reading their final written reports, but the results were disappointing. Several of the students whom you consider above-average had spelling and grammatical errors in their papers and so received a "C" on the assignment. You were surprised that **Emma**, who typically excels in English, was one of those, while **Lila**, who rarely does homework, received an "A" for a well-written paper.

After you have handed the papers back, you happen to overhear **August** in the hall disposing his grade with another student. "I can't believe I got a 'C'!", you hear him exclaim. "I worked really hard on this project. I even got answers from Barbara Ringdove! I read lots of sources for my paper, too, and took notes from a lot of Web sites. Then **Lila** goes and gets an "A" without doing hardly anything and making up an interview, just because she writes better. It's just not fair! If we had a multiple choice test I would've got an "A" for sure. It seems like it's not worth it to even try hard!" "I sure know what you mean," comments August's friend. You walk back into your classroom, asking yourself questions about what you've just heard.

Questions

1. How were problems with this assignment? In what ways could the assignment be improved?
2. What good and bad assessment practices were employed? In what general ways could the assessment be made fairer and more authentic for all of your students? Refer to the student files & comment in your brief response.
3. How could/should students' concerns with assessment be addressed?
4. That are some other possible ways of assessing student learning during this project? List at least three and briefly discuss the strong and weak points of using these assessment strategies.

Figure 1. Situation 2 from the Timberlake simulation

Abbildung 80: Abbildung eines Ausdrucks der Tigerlake-Simulation, Quelle: Egbert et al. (2000, S.25)

Transition to Teaching Classroom Management



Abbildung 81: Screenshot des Simulators Transition to Teaching Classroom Management, Quelle: Diefenbach et al. (2019, S.198)



Abbildung 82: Screenshot des Simulators Transition to Teaching Classroom Management, Quelle: Diefenbach et al. (2019, S.198)

VCS – Virtual Classroom Simulator



Abbildung 83: Screenshot des Hauptbildschirms von VCS, Quelle: Skrødal (2010, S.120)

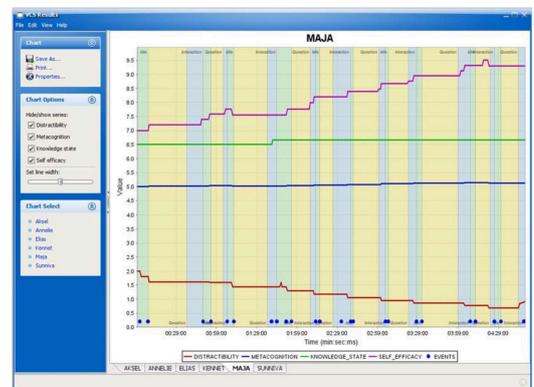


Abbildung 84: Screenshot des Feedbackbildschirms von VCS, Quelle: Skrødal (2010, S.127)



Abbildung 85: Screenshot der Fragen-Komponente in VCS, Quelle: Skrødal (2010, S.124)

Verbal Abuse Simulation



Abbildung 86: Screenshot der Verbal Abuse Simulation, Quelle: Diefenbach et al. (2019, S.199)

Virtual School Class (VSC)



Abbildung 87: Screenshot des Simulators Virtual School Class (VSC), Quelle: Fiedler (2012, S.28)

Virtual Simulated Classroom (V.C.R.)

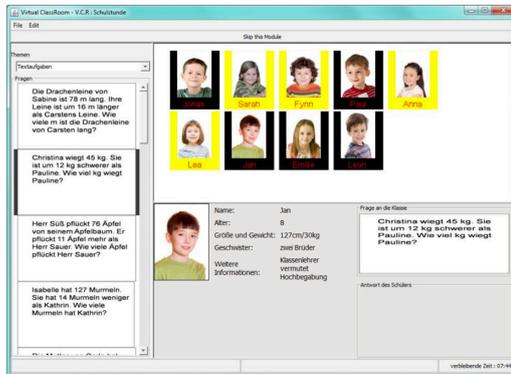


Abbildung 88: Screenshot des Simulators Virtual Simulated Classroom (V.C.R.), Quelle: Fischer et al. (2017)

VR-Klassenraum



Abbildung 89: Screenshot des VR-Klassenraums, Quelle: Wiepke und Paulicke (2018, S.2)

VR-Tool von Bailenson et al. (ohne Bezeichnung)

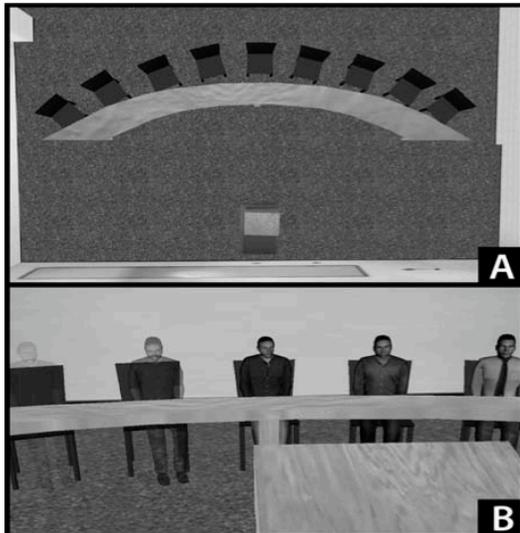


Abbildung 90: Screenshot des VR-Tools von Bailenson et al., Quelle: Bailenson et al. (2008, S.114)



Abbildung 91: Foto des Simulationsaufbaus mittels HMD des VR-Tools von Bailenson et al., Quelle: Bailenson et al. (2008, S.114)

VR-Tool von Yaacov J. Katz (ohne Bezeichnung)



Abbildung 92: Screenshot der VR-Umgebung SuperScope, welche die Simulation nutzt, kein Screenshot der Simulation selbst, Quelle: https://www.youtube.com/watch?v=Xht5KXInD_s

VR-Tool von Stavroulia et al. (ohne Bezeichnung)



Abbildung 93: Screenshot des VR-Tools von Stavroulia et al., Quelle: Stavroulia et al. (2019, S.203)



Abbildung 94: Screenshot des VR-Tools von Stavroulia et al., Quelle: Stavroulia et al. (2018, S.276)

Konzeptionsmodell-Simulationsmodell-Übersetzungsgraphen

261

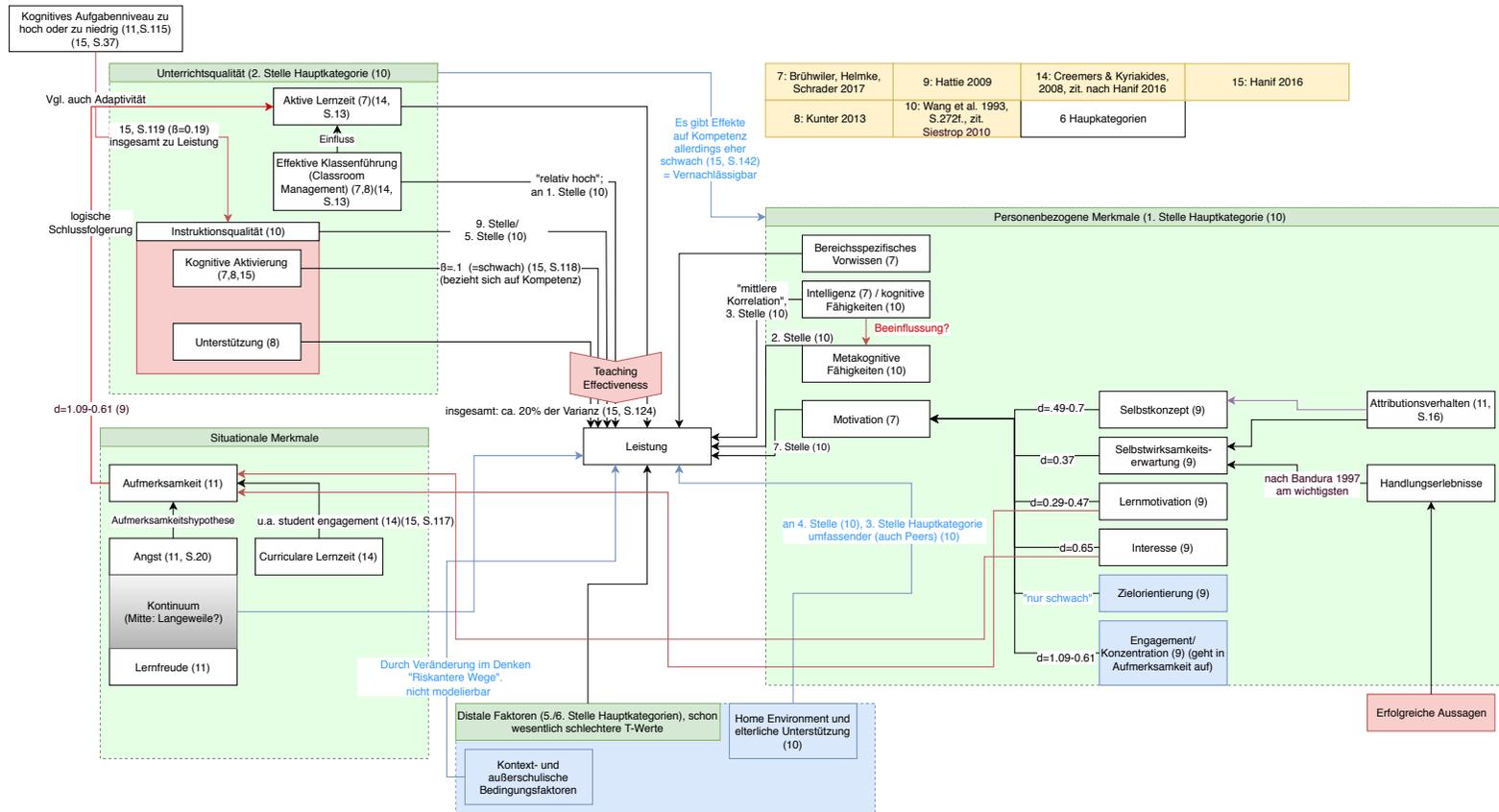
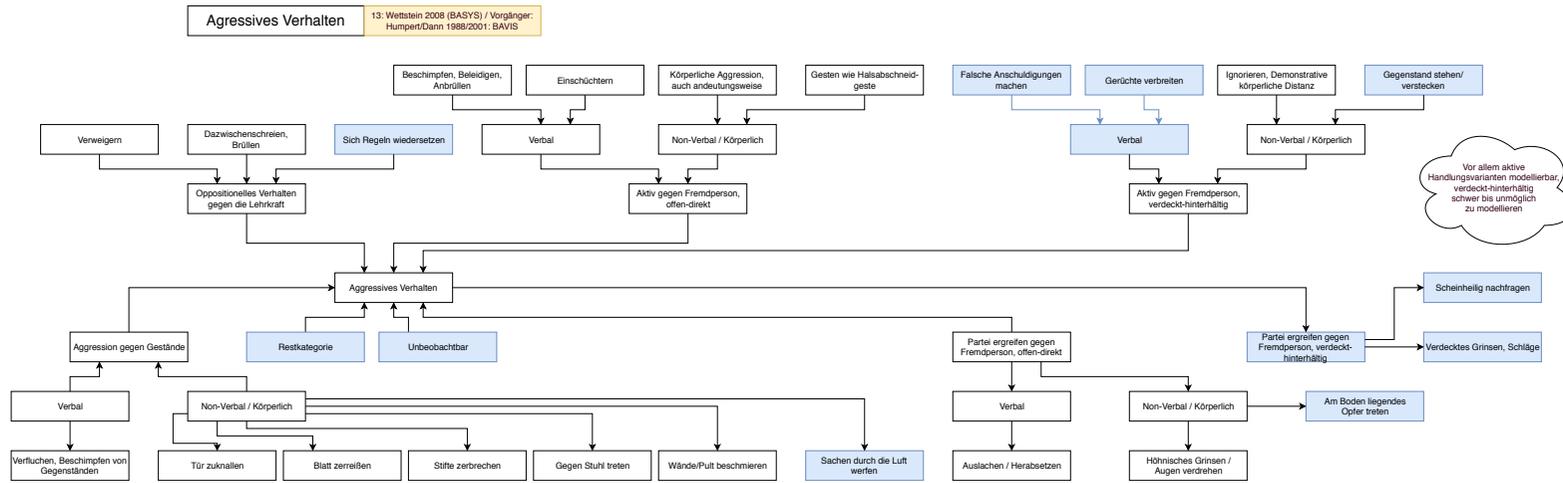


Abbildung 95: Integration verschiedener Konzeptionsmodelle für die Zielgröße akademische Leistung, Quelle: Autor



262

Abbildung 96: Integration verschiedener Konzeptionsmodelle für die Zielgröße aggressives Verhalten, Quelle: Autor

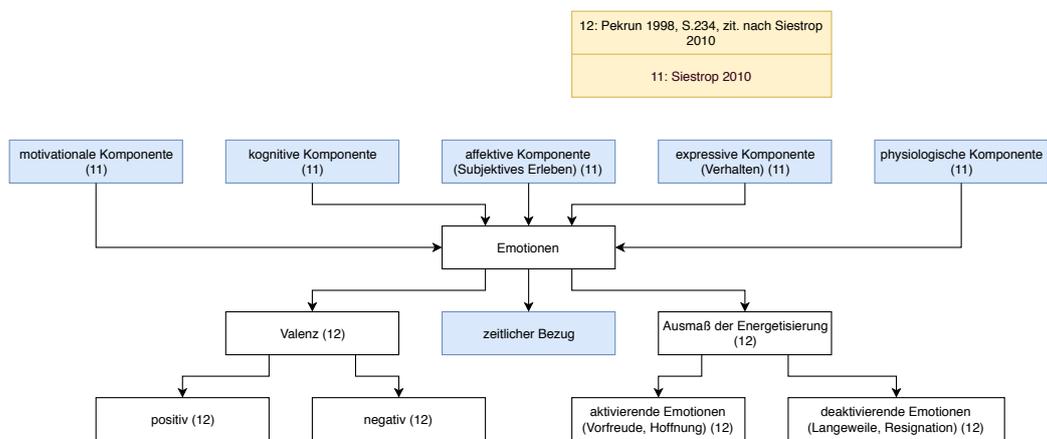


Abbildung 97: Integration verschiedener Konzeptionsmodelle für die Zielgröße Emotionen, Quelle: Autor

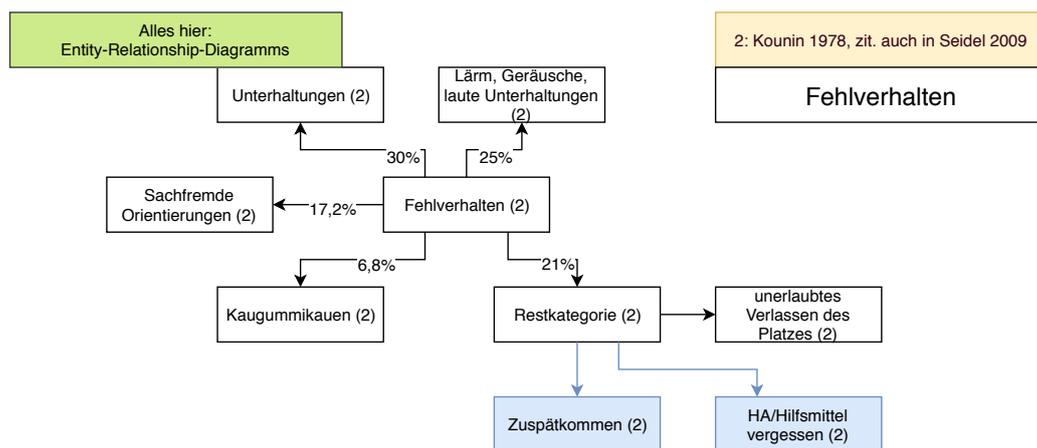


Abbildung 98: Integration verschiedener Konzeptionsmodelle für die Zielgröße Fehlverhalten, Quelle: Autor

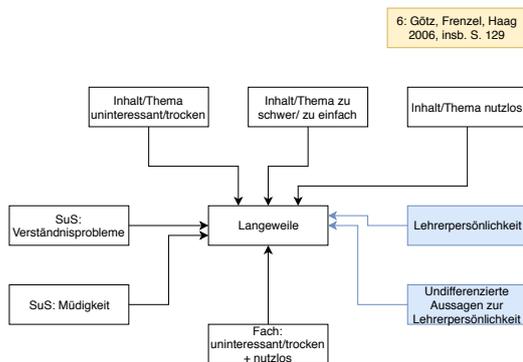


Abbildung 99: Integration verschiedener Konzeptionsmodelle für die Zielgröße Langeweile, Quelle: Autor

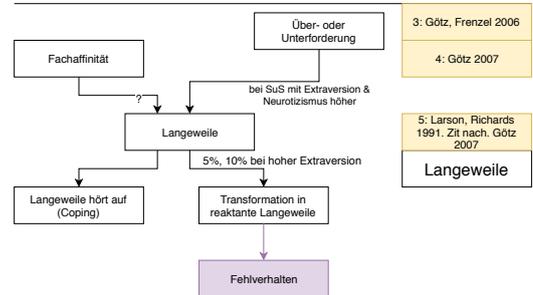


Abbildung 100: Integration verschiedener Konzeptionsmodelle für die Zielgröße Langeweile, Quelle: Autor

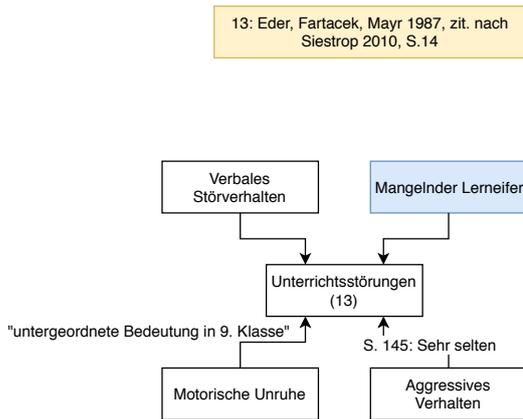


Abbildung 101: Integration verschiedener Konzeptionsmodelle für die Zielgröße Unterrichtsstörungen, Quelle: Autor



Abbildung 102: Integration verschiedener Konzeptionsmodelle für die Zielgröße Unterrichtsstörungen, Quelle: Autor

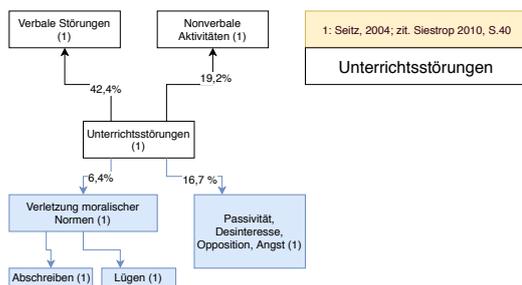


Abbildung 103: Integration verschiedener Konzeptionsmodelle für die Zielgröße Unterrichtsstörungen, Quelle: Autor

Angegebene Literatur in den Konzeptionsmodell-Simulationsmodell-Übersetzungsgraphen:

Brühwiler et al. (2017), Kunter und Trautwein (2013), Hattie (2009), Hanif (2016), Siestrop (2010), Wettstein (2008), Humpert und Dann (1984), Dann und Humpert (2002), Götz et al. (2007), Götz et al. (2006) und Götz und Frenzel (2006)

Tabellen der Clasivir-Variablen

Tabelle 15: Zuweisung der Nation des Migrationshintergrunds

n	Land	Anzahl	Prozent	Edge
1	Türkei	22629	26.20%	0.26204014
2	Italien	6232	7.22%	0.33420568
3	Irak	4874	5.64%	0.39064581
4	Polen	4616	5.35%	0.44409834
5	Russische Föderation	2921	3.38%	0.47792304
6	Montenegro, Serbien	2891	3.35%	0.51140035
7	Syrien	2744	3.18%	0.54317542
8	Marokko	2568	2.97%	0.57291244
9	Bulgarien	2527	2.93%	0.60217469
10	Iran	2118	2.45%	0.62670079
11	Afghanistan	1844	2.14%	0.64805401
12	Griechenland	1570	1.82%	0.66623435
13	Rumänien	1417	1.64%	0.68264298
14	Bosnien und Herzegowina	1361	1.58%	0.69840314
15	Frankreich	1315	1.52%	0.71363063
16	Spanien	1224	1.42%	0.72780435
17	Kosovo	1115	1.29%	0.74071587
18	Tunesien	1078	1.25%	0.75319893
19	Ukraine	997	1.15%	0.76474403
20	Kroatien	954	1.10%	0.77579119
21	Niederlande	883	1.02%	0.78601619
22	Kasachstan	882	1.02%	0.79622961
23	Mazedonien	871	1.01%	0.80631564
24	Portugal	791	0.92%	0.81547529
25	USA	783	0.91%	0.82454231
26	Nigeria	771	0.89%	0.83347036
27	Indien	720	0.83%	0.84180784
28	Großbritannien	686	0.79%	0.84975161
29	Österreich	656	0.76%	0.85734799
30	Albanien	578	0.67%	0.86404113
31	China	533	0.62%	0.87021318
32	Belgien	514	0.60%	0.87616522
33	Brasilien	497	0.58%	0.88192040

Fortsetzung auf nächster Seite

Tabelle 15: Zuweisung der Nation des Migrationshintergrunds (fortgesetzt)

n	Land	Anzahl	Prozent	Edge
34	Ungarn	407	0.47%	0.88663339
35	Ghana	404	0.47%	0.89131165
36	Algerien	373	0.43%	0.89563093
37	Pakistan	369	0.43%	0.89990389
38	Vietnam	322	0.37%	0.90363259
39	Kongo, Demo. Rep.	319	0.37%	0.90732656
40	Eritrea	317	0.37%	0.91099737
41	Togo	305	0.35%	0.91452922
42	Kongo	274	0.32%	0.91770210
43	Schweiz	272	0.31%	0.92085181
44	Aserbajdschan	254	0.29%	0.92379309
45	Libanon	244	0.28%	0.92661857
46	Litauen	234	0.27%	0.92932825
47	Kamerun	233	0.27%	0.93202636
48	Thailand	222	0.26%	0.93459708
49	Äthiopien	212	0.25%	0.93705201
50	Kenia	211	0.24%	0.93949535
51	Angola	200	0.23%	0.94181132
52	Armenien	187	0.22%	0.94397675
53	Tschechien	182	0.21%	0.94608428
54	Ägypten	176	0.20%	0.94812233
55	Guinea	171	0.20%	0.95010248
56	Schweden	169	0.20%	0.95205947
57	Irland	158	0.18%	0.95388909
58	Japan	157	0.18%	0.95570712
59	Slovakei	153	0.18%	0.95747884
60	Kirgisistan	151	0.17%	0.95922739
61	Lettland	150	0.17%	0.96096437
62	Peru	146	0.17%	0.96265503
63	Philippinen	146	0.17%	0.96434568
64	Kuba	140	0.16%	0.96596686
65	Israel	132	0.15%	0.96749540
66	Cote d'Ivoire	130	0.15%	0.96900078
67	Georgien	123	0.14%	0.97042510
68	Mexiko	122	0.14%	0.97183784

Fortsetzung auf nächster Seite

Tabelle 15: Zuweisung der Nation des Migrationshintergrunds (fortgesetzt)

n	Land	Anzahl	Prozent	Edge
69	Kolumbien	120	0.14%	0.97322742
70	Weißrussland	119	0.14%	0.97460542
71	Kanada	119	0.14%	0.97598342
72	Moldau, Republik	118	0.14%	0.97734984
73	Australien	109	0.13%	0.97861204
74	Korea (Süd)	107	0.12%	0.97985108
75	Sri Lanka	102	0.12%	0.98103223
76	Ecuador	98	0.11%	0.98216705
77	Indonesien	97	0.11%	0.98329029
78	Finnland	95	0.11%	0.98439038
79	Senegal	95	0.11%	0.98549046
80	Somalia	95	0.11%	0.98659055
81	Bangladesch	87	0.10%	0.98759799
82	Jordanien	85	0.10%	0.98858228
83	Argentinien	84	0.10%	0.98955499
84	Usbekistan	83	0.10%	0.99051611
85	Dänemark	82	0.09%	0.99146566
86	Dom.Rep.	69	0.08%	0.99226467
87	Chile	63	0.07%	0.99299420
88	Slowenien	62	0.07%	0.99371215
89	Venezuela	62	0.07%	0.99443010
90	Gambia	61	0.07%	0.99513647
91	Libyen	58	0.07%	0.99580810
92	Uganda	58	0.07%	0.99647973
93	Estland	53	0.06%	0.99709346
94	Luxemburg	53	0.06%	0.99770719
95	Mongolei	51	0.06%	0.99829776
96	Südafrika	40	0.05%	0.99876096
97	Bolivien	38	0.04%	0.99920099
98	Norwegen	35	0.04%	0.99960629
99	Malta	15	0.02%	0.99977998
100	Korea (Nord)	10	0.01%	0.99989578
101	Zypern (griechisch)	9	0.01%	1

Tabelle 16: Markov-Modell-Algorithmus zur Modellierung des Wechsels emotionaler Zustände

Emotionsstatus	<i>Neutral</i>	<i>Freudig</i>	<i>Traurig</i>	<i>Ängstlich</i>	<i>Wütend</i>	<i>Überrascht</i>
Neutral	70%	15%	5%	4%	4%	2%
Freudig	20%	70%	2%	1%	1%	6%
Traurig	15%	1%	50%	20%	9%	5%
Ängstlich	20%	1%	20%	40%	14%	5%
Wütend	20%	1%	1%	2%	75%	1%
Überrascht	50%	25%	2%	20%	1%	2%

Tabelle 17: Liste aller in Clasivir implementierten Störverhalten

Störungsverhalten	Clasivir-Variable	Abhängigkeit von OCEAN-Variablen				Literatur
		Basis-Wahrscheinlichkeit	Länge in Intervallen	Benötigt Ziel?		
Kein Störungsverhalten	SV_Kein	-	-	Nein	-	
Nicht-aggressives Störungsverhalten						
Unterhaltung mit Nachbarin	SV_U	E	15%	2*	Ja	Böheim et al. (2020, S.3), Thomas et al. (1968, S.37) und Pfitzner und Schoppek (2000, S.375ff.)
Motorische Unruhe	SV_MOTUNR	N	12.5%	2	Nein	Appel (2015, S.104), Siestrup (2010, S.2010), Thomas et al. (1968, S.37) und Pfitzner und Schoppek (2000, S.375ff.)
Produzieren von Lärm oder Geräuschen	SV_GERAEU	E	10%	1	Nein	Pfitzner und Schoppek (2000, S.375ff.), Kounin (1976, S.36) und Thomas et al. (1968, S.37)
Kaugummikauen	SV_KAUGUM		4%	4	Nein	Kounin (1976, S.36)
Ins Heft malen (unaufgefordert)	SV_ZEICHN		11%	5	Nein	Tadayon und Afhami (2016)
Im Raum herumlaufen	SV_ZEICHN		2%	10	Nein	Pfitzner und Schoppek (2000, S.375ff.)
Blick abwenden / aus dem Fenster schauen	SV_BLICKFEN		11%	3	Nein	Böheim et al. (2020, S.3) und Pfitzner und Schoppek (2000, S.375ff.)
Blick abwenden / auf den Tisch schauen (unaufgefordert)	SV_BLICKTISCH		11%	3	Nein	Böheim et al. (2020, S.3)
Schlafen	SV_SCHLAF		3,5%	100	Nein	-
Mit dem Handy telefonieren	SV_HANDYT		2%	10	Nein	Übernommen aus BBB, Lugrin et al. (2016, S.4)
Mit dem Handy spielen	SV_HANDYS		3%	10	Nein	Übernommen aus BBB, Lugrin et al. (2016, S.4)
Aggressives Störungsverhalten (verbal)						
<i>Fortsetzung auf nächster Seite</i>						

Tabelle 17: Liste aller in Clasivir implementierten Störverhalten (fortgesetzt)

Bedrohen	SV_BEDROH	E,N	3%	1	Ja	Scherzinger et al. (2018, S.196)
Beschimpfen/Beleidigen	SV_BELEID	E,N	4%	2	Ja	Scherzinger et al. (2018, S.196) und Pfitzner und Schoppek (2000, S.375ff.)
Auslachen	SV_AUSLACH	E,N	3%	1	Ja	Scherzinger et al. (2018, S.196) und Pfitzner und Schoppek (2000, S.375ff.)
Herabwürdigen	SV_HERABW	E,N	0.7%	2	Ja	Scherzinger et al. (2018, S.196) und Pfitzner und Schoppek (2000, S.375ff.)
Aggressives Störungsverhalten (physisch gegen Menschen)						
Schwacher Tritt	SV_TRITT_SCHWACH	E,N	1%	1	Ja	Wettstein (2008), Scherzinger et al. (2018, S.196), Pfitzner und Schoppek (2000, S.375ff.) und Dann und Humpert (2002)
Starker Tritt	SV_TRITT_STARK	E,N	0.3%	1	Ja	Wettstein (2008), Scherzinger et al. (2018, S.196), Pfitzner und Schoppek (2000, S.375ff.) und Dann und Humpert (2002)
Schwacher Schlag	SV_SCHLAG_SCHWACH	E,N	1%	1	Ja	Wettstein (2008), Scherzinger et al. (2018, S.196), Pfitzner und Schoppek (2000, S.375ff.) und Dann und Humpert (2002)
Starker Schlag	SV_SCHLAG_STARK	E,N	0.3%	1	Ja	Wettstein (2008), Scherzinger et al. (2018, S.196), Pfitzner und Schoppek (2000, S.375ff.) und Dann und Humpert (2002)
Aggressives Störungsverhalten (physisch gegen Gegenstände)						
Blatt zerreißen	SV_BLATT	E,N	1%	1	Nein	Wettstein (2008) und Dann und Humpert (2002)
Stift zerstören	SV_STIFT	E,N	1%	1	Nein	Wettstein (2008) und Dann und Humpert (2002)
Tisch beschmieren	SV_VANDAL	E,N	1%	1	Nein	Wettstein (2008) und Dann und Humpert (2002)

Tabelle 18: Modifikationen der Störungswahrscheinlichkeiten durch emotionale Stati

	Freude	Traurigkeit	Angst	Überraschung	Wut
Unterhaltung	1.5	0.6	0.5	2.0	0.9
Motorische Unruhe	1.2	1.0	1.1	1.0	1.0
Geräusch / Lärmproduktion	1.3	0.6	0.7	1.5	1.5
Kaugummikauen	1.0	1.0	0.4	0.5	1.0
Ins Heft malen	1.0	1.4	0.8	0.5	0.5
Im Raum herumlaufen	1.1	0.8	0.5	0.5	0.7
Aus dem Fenster blicken	0.7	1.6	1.1	0.5	0.6
Auf den Tisch blicken	0.6	1.8	1.5	0.5	0.8
Schlafen	0.2	0.9	0.5	0.5	0.5
Mit dem Handy telefonieren	1.1	0.5	0.5	0.5	0.8
Mit dem Handy spielen	1.2	1.0	0.5	0.5	0.6
Bedrohung	0.4	1.0	1.1	0.5	2.0
Beschimpfung	0.4	1.0	1.1	0.5	2.0
Auslachen	1.0	0.5	0.5	0.5	2.0
Herabwürdigen	0.5	1.0	1.0	0.5	2.0
Schwacher Tritt	0.5	1.0	1.1	0.5	3.0
Starker Tritt	0.5	1.0	1.1	0.5	3.0
Schwacher Schlag	0.5	1.0	1.1	0.5	3.0
Starker Schlag	0.5	1.0	1.1	0.5	3.0
Blatt zerreißen	0.5	1.5	1.0	0.5	2.0
Stift zerstören	0.5	1.5	1.0	0.5	2.0
Tisch beschmieren	0.5	1.5	1.0	0.5	2.0

Tabelle 19: Berechnung der Aufgabenschwierigkeit

Spezifischer Schwierigkeitsgrad Einfach						
	<i>Erinnern</i>	<i>Verstehen</i>	<i>Anwenden</i>	<i>Analysieren</i>	<i>Evaluiieren</i>	<i>Erstellen</i>
Faktenwissen	.16	.18	.24	.3	0.36	0.42
Konzeptionelles Wissen	.216	.24	.3	.36	.42	.48
Prozedurales Wissen	.288	.3	.36	.42	.48	.54
Metakognitives Wissen	.3	.36	.42	.48	.54	.6
Spezifischer Schwierigkeitsgrad Mittel						
	<i>Erinnern</i>	<i>Verstehen</i>	<i>Anwenden</i>	<i>Analysieren</i>	<i>Evaluiieren</i>	<i>Erstellen</i>
Faktenwissen	.32	.36	.4	.44	0.48	0.56
Konzeptionelles Wissen	.36	.384	.42	.48	.56	.64
Prozedurales Wissen	.48	.5	.528	.56	.64	.72
Metakognitives Wissen	.5	.528	.56	.64	.72	.8
Spezifischer Schwierigkeitsgrad Schwer						
	<i>Erinnern</i>	<i>Verstehen</i>	<i>Anwenden</i>	<i>Analysieren</i>	<i>Evaluiieren</i>	<i>Erstellen</i>
Faktenwissen	.6	.624	.64	.66	0.672	0.7
Konzeptionelles Wissen	.6	.64	.66	.672	.7	.8
Prozedurales Wissen	.64	.68	.672	.728	.8	.9
Metakognitives Wissen	.7	.72	.756	.8	.9	1.0

Tabelle 20: Beispieltabelle der Aufgabenadaptivität für eine Schülerin mit $KF=.5$

Spezifischer Schwierigkeitsgrad Einfach						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluieren	Erstellen
Faktenwissen	Zu einfach	Zu einfach	Zu einfach	Zu einfach	Lösbar	Lösbar
Konzeptionelles Wissen	Zu einfach	Zu einfach	Zu einfach	Lösbar	Lösbar	Lösbar
Prozedurales Wissen	Zu einfach	Zu einfach	Lösbar	Lösbar	Lösbar	Lösbar
Metakognitives Wissen	Zu einfach	Lösbar	Lösbar	Lösbar	Optimal	Optimal
Spezifischer Schwierigkeitsgrad Mittel						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluieren	Erstellen
Faktenwissen	Lösbar	Lösbar	Lösbar	Lösbar	Lösbar	Optimal
Konzeptionelles Wissen	Lösbar	Lösbar	Lösbar	Lösbar	Optimal	Nicht lösbar
Prozedurales Wissen	Lösbar	Lösbar	Lösbar	Optimal	Nicht lösbar	Nicht lösbar
Metakognitives Wissen	Lösbar	Lösbar	Optimal	Nicht lösbar	Nicht lösbar	Nicht lösbar
Spezifischer Schwierigkeitsgrad Schwer						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluieren	Erstellen
Faktenwissen	Optimal	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar
Konzeptionelles Wissen	Optimal	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar
Prozedurales Wissen	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar
Metakognitives Wissen	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar	Nicht lösbar

Tabelle 21: Adaptive Verarbeitungswahrscheinlichkeiten

Dies ist eine Beispieltabelle für eine Schülerin mit $KF=.5$ und den OCEAN-Werten $O=.53257$; $C=.512384$; $E=.283469$; $A=.304634$; $N=.481794$

Spezifischer Schwierigkeitsgrad Einfach						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	83,3%	83,3%	83,3%	83,3%	85,34%	85,34%
Konzeptionelles Wissen	83,3%	83,3%	83,3%	85,34%	85,34%	85,34%
Prozedurales Wissen	83,3%	83,3%	85,34%	85,34%	85,34%	85,34%
Metakognitives Wissen	83,3%	85,34%	85,34%	85,34%	90,14%	90,14%
Spezifischer Schwierigkeitsgrad Mittel						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	85,34%	85,34%	85,34%	85,34%	85,34%	90,14%
Konzeptionelles Wissen	85,34%	85,34%	85,34%	85,34%	90,14%	57,55%
Prozedurales Wissen	85,34%	85,34%	85,34%	90,14%	57,55%	57,55%
Metakognitives Wissen	85,34%	85,34%	90,14%	57,55%	57,55%	57,55%
Spezifischer Schwierigkeitsgrad Schwer						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	90,14%	57,55%	57,55%	57,55%	57,55%	57,55%
Konzeptionelles Wissen	90,14%	57,55%	57,55%	57,55%	57,55%	57,55%
Prozedurales Wissen	57,55%	57,55%	57,55%	57,55%	57,55%	57,55%
Metakognitives Wissen	57,55%	57,55%	57,55%	57,55%	57,55%	57,55%

Diese Beispieltabelle zeigt die Chancen zur Verarbeitung aller möglichen Aufgaben für eine durchschnittlich begabte, eher introvertierte und durchschnittlich offene Schülerin, die dem Unterricht im dargestellten Moment durchschnittlich konzentriert folgt. Diese Tabelle ändert unterscheidet sich für jeden Schüler aufgrund der unterschiedlichen kognitiven Fähigkeit und der unterschiedlichen OCEAN-Werte und für jeden Moment in der Simulation aufgrund der unterschiedlichen Konzentrationswerte.

Tabelle 22: Berechnung der *functional-off-task* Wahrscheinlichkeiten

Chance, dass <i>functional-off-task</i> stattfindet: 6%	
Störverhalten	Chance
Motorische Unruhe	25%
Kaugummikauen	5%
Blick zum Fenster	35%
Blick auf den Tisch	35%

Tabelle 23: Adaptive Wahrscheinlichkeiten für Hereinrufen

Dies ist eine Beispieltabelle für eine Schülerin mit $KF=.5$ und den OCEAN-Werten $O=.53257$;
 $C=.512384$; $E=.283469$; $A=.304634$; $N=.481794$

Spezifischer Schwierigkeitsgrad Einfach						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	4,21%	4,21%	4,21%	4,21%	1,58%	1,58%
Konzeptionelles Wissen	4,21%	4,21%	4,21%	1,58%	1,58%	1,58%
Prozedurales Wissen	4,21%	4,21%	1,58%	1,58%	1,58%	1,58%
Metakognitives Wissen	4,21%	1,58%	1,58%	1,58%	1,05%	1,05%
Spezifischer Schwierigkeitsgrad Mittel						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	1,58%	1,58%	1,58%	1,58%	1,58%	1,05%
Konzeptionelles Wissen	1,58%	1,58%	1,58%	1,58%	1,05%	0,63%
Prozedurales Wissen	1,58%	1,58%	1,58%	1,05%	0,63%	0,63%
Metakognitives Wissen	1,58%	1,58%	1,05%	0,63%	0,63%	0,63%
Spezifischer Schwierigkeitsgrad Schwer						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	1,05%	0,63%	0,63%	0,63%	0,63%	0,63%
Konzeptionelles Wissen	1,05%	0,63%	0,63%	0,63%	0,63%	0,63%
Prozedurales Wissen	0,63%	0,63%	0,63%	0,63%	0,63%	0,63%
Metakognitives Wissen	0,63%	0,63%	0,63%	0,63%	0,63%	0,63%

Tabelle 24: Adaptive Meldewahrscheinlichkeiten nach erfolgter Verarbeitung

Dies ist eine Beispieltabelle für eine Schülerin mit $KF=.5$ und den OCEAN-Werten $O=.53257$;
 $C=.512384$; $E=.283469$; $A=.304634$; $N=.481794$

Spezifischer Schwierigkeitsgrad Einfach						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	16,32%	16,32%	16,32%	16,32%	10,61%	10,61%
Konzeptionelles Wissen	16,32%	16,32%	16,32%	10,61%	10,61%	10,61%
Prozedurales Wissen	16,32%	16,32%	10,61%	10,61%	10,61%	10,61%
Metakognitives Wissen	16,32%	10,61%	10,61%	10,61%	7,07%	7,07%
Spezifischer Schwierigkeitsgrad Mittel						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	10,61%	10,61%	10,61%	10,61%	10,61%	7,07%
Konzeptionelles Wissen	10,61%	10,61%	10,61%	10,61%	7,07%	5,30%
Prozedurales Wissen	10,61%	10,61%	10,61%	7,07%	5,30%	5,30%
Metakognitives Wissen	10,61%	10,61%	7,07%	5,30%	5,30%	5,30%
Spezifischer Schwierigkeitsgrad Schwer						
	Erinnern	Verstehen	Anwenden	Analysieren	Evaluiieren	Erstellen
Faktenwissen	7,07%	5,30%	5,30%	5,30%	5,30%	5,30%
Konzeptionelles Wissen	7,07%	5,30%	5,30%	5,30%	5,30%	5,30%
Prozedurales Wissen	5,30%	5,30%	5,30%	5,30%	5,30%	5,30%
Metakognitives Wissen	5,30%	5,30%	5,30%	5,30%	5,30%	5,30%

Tabelle 25: Mögliche Meldungstypen nach erfolgter Meldung

Dies ist eine Beispieltabelle für eine Schülerin mit $KF=.5$ und den OCEAN-Werten $O=.53257$;
 $C=.512384$; $E=.283469$; $A=.304634$; $N=.481794$

Meldungstyp	Wahrscheinlichkeit
Undeutliche Meldung	25,85%
Aufdringliche Meldung	20,12%
Normale Meldung	54,06%

Tabelle 26: Konzentrationsmodifikationen nach erfolgter *classroom management* Handlung

Handlung	Störverhalten vorhanden?	Konzentrationsmodifikation	
		Einzelziel /Gruppe	Klasse
Loben	Ja	-.1	-.15
Im Verhalten bestärken	Ja	-.1	-.1
Konstruktives Ermahnen	Ja	+.1	+.05
Destruktives Ermahnen	Ja	-.15	-.05
Loben	Nein	+.1	+.02
Im Verhalten bestärken	Nein	+.12	+.03
Konstruktives Ermahnen	Nein	-.2	-.05
Destruktives Ermahnen	Nein	-.25	-.1

Tabelle 27: Konzentrationsmodifikationen nach erfolgter Feedback-Handlung (stets Einzelziel)

Handlung	Frage korrekt beantwortet?	Konzentrations- modifikation
Loben	Ja	+.1
Bejahen (affirm)	Ja	+.15
Verneinen (negate)	Ja	-.2
Kritisieren (criticise)	Ja	-.3
Loben	Nein	-.2
Bejahen (affirm)	Nein	-.15
Verneinen (negate)	Nein	+.05
Kritisieren (criticise)	Nein	-.1

Tabelle 28: Liste aller über die ini-Datei modifizierbaren Simulationsvariablen

Name	Std. Wert	Beschreibung
GV_CHANCE_DISRUPTIVE_BEHAVIOR	0.3	Dies zeigt die Wahrscheinlichkeit an, mit der bei den SuS Störungsverhalten auftritt, wenn die Konzentration unter dem Schwellenwert CONCENTRATION_THRESHOLD_DISRUPTIVE_BEHAVIOR liegt. Dieser Wert wird durch den Neurotizismus der jeweiligen Schülerin modifiziert.
GV_CHANCE_TEACHER_TARGET_SV_BEDROH	0.05	Dies zeigt die Wahrscheinlichkeit an, dass die Lehrkraft – und nicht eine andere Schülerin – beim Störungsverhalten ‚Bedrohung‘ als Ziel ausgewählt wird.
GV_CHANCE_TEACHER_TARGET_SV_BELEID	0.1	Dies zeigt die Wahrscheinlichkeit an, dass die Lehrkraft – und nicht eine Schülerin – beim Störungsverhalten ‚Beleidigung‘ als Ziel ausgewählt wird.
GV_CHANCE_TEACHER_TARGET_SV_AUSLACHEN	0.1	Dies zeigt die Wahrscheinlichkeit an, dass die Lehrkraft – und nicht eine Schülerin – beim Störungsverhalten ‚Auslachen‘ als Ziel ausgewählt wird.
GV_CHANCE_TEACHER_TARGET_SV_HERABW	0.1	Dies zeigt die Wahrscheinlichkeit an, dass die Lehrkraft – und nicht eine Schülerin – beim Störungsverhalten ‚Herabwürdigenden‘ als Ziel ausgewählt wird.
GV_CHANCE_RH_WITHDRAW	0.02	Dies zeigt die Grundwahrscheinlichkeit an, dass eine Schülerin die Meldung zurückzieht.
GV_CHANCE_RH_WITHDRAW_ADD_PER_INTERVAL	0.01	Dieser Wert in jedem folgenden Intervall zur Grundwahrscheinlichkeit des Meldungsrückzugs addiert.
GV_CHANCE_RH_WITHDRAW_MAX	0.1	Dies ist der Maximalwert der Wahrscheinlichkeit eines Melderückzugs. (keine zusätzliche Addierung von GV_CHANCE_RH_WITHDRAW_ADD_PER_INTERVAL zu GV_CHANCE_RH_WITHDRAW über diesen Wert hinaus)
GV_CHANCE_FUNCTIONAL_OFFTASK	0.06	Dies ist die Grundwahrscheinlichkeit dafür, dass ein Schüler nach dem Entschluss eine Aufgabe zu verarbeiten Functional-Off-Task-Verhalten zeigt (
GV_CHANCE_CORRECT_ANSWER_TASKTOOEASY	0.9	Dieser Wert ist die Grundwahrscheinlichkeit für die richtige Lösung einer Aufgabe, die zu einfach für die Schülerin ist.
GV_CHANCE_CORRECT_ANSWER_TASKSOLVEABLE	0.8	Dieser Wert ist die Grundwahrscheinlichkeit für die richtige Lösung einer Aufgabe, die normal für die Schülerin lösbar ist.
GV_CHANCE_CORRECT_ANSWER_TASKOPTIMAL	0.7	Dieser Wert ist die Grundwahrscheinlichkeit für die richtige Lösung einer Aufgabe, die sich im optimalen Lernbereich für die Schülerin befindet.
GV_DIVISOR_RAISED_HAND_TASKTOOEASY	1.3	Dies ist der Divisor, durch den die Grundmotivation einer Schülerin sich zu melden geteilt wird, wenn die Aufgabe zu einfach ist. Die Grundmotivation der Schülerin hängt von ihren OCEAN-Werten ab und befindet sich in der Regel in einem Wertebereich um 0.4–0.5. Der Wert sollte über 1.0 liegen.
GV_DIVISOR_RAISED_HAND_TASKSOLVEABLE	2	Dies ist der Divisor, durch den die Grundmotivation einer Schülerin sich zu melden geteilt wird, wenn die Aufgabe normal lösbar ist. Die Grundmotivation der Schülerin hängt von ihren OCEAN-Werten ab und befindet sich in der Regel in einem Wertebereich um 0.4–0.5. Der Wert sollte über 1.0 liegen.

Fortsetzung auf nächster Seite

GV_DIVISOR_RAISED_HAND_TASKOPTIMAL	3	Dies ist der Divisor, durch den die Grundmotivation einer Schülerin sich zu melden geteilt wird, wenn die Aufgabe optimal lösbar ist. Die Grundmotivation der Schülerin hängt von ihren OCEAN-Werten ab und befindet sich in der Regel in einem Wertebereich um 0.4–0.5. Der Wert sollte über 1.0 liegen.
GV_DIVISOR_RAISED_HAND_TASKNOTSOLVEABLE	4	Dies ist der Divisor, durch den die Grundmotivation einer Schülerin sich zu melden geteilt wird, wenn die Aufgabe nicht lösbar ist. Die Grundmotivation der Schülerin hängt von ihren OCEAN-Werten ab und befindet sich in der Regel in einem Wertebereich um 0.4–0.5. Der Wert sollte über 1.0 liegen.
GV_DIVISOR_CONCENTRATION_DECAY	32	Diese Variable bestimmt, wie stark die Konzentration biexponentiell über die Zeit fällt.
GV_DIVISOR_CHANCE_STUDENT_REPLY	1	Diese Variable bestimmt den Divisor eine spezifische SuS-Chancenberechnung, nämlich ob eine Schülerin einer anderen antwortet, nachdem dieser als Unterhaltungsziel bestimmt wurde. Dies ist abhängig von Extraversion und Conscientiousness der Schülerin.
GV_CONCENTRATION_ADD_FB_PRAISE_CA	0.1	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten addiert, wenn eine Lehrkraft ein Lob auf eine korrekte SuS-Antwort gibt.
GV_CONCENTRATION_ADD_FB_AFFIRM_CA	0.15	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten addiert, wenn eine Lehrkraft eine Bejahung auf eine korrekte SuS-Antwort gibt.
GV_CONCENTRATION_SUBT_FB_NEGATE_CA	0.2	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft eine Verneinung auf eine korrekte SuS-Antwort gibt.
GV_CONCENTRATION_SUBT_FB_CRITICIZE_CA	0.3	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft mit Kritik auf eine korrekte SuS-Antwort reagiert.
GV_CONCENTRATION_SUBT_FB_PRAISE_WA	0.2	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft ein Lob auf eine falsche SuS-Antwort gibt.
GV_CONCENTRATION_SUBT_FB_AFFIRM_WA	0.15	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft eine Bejahung auf eine falsche SuS-Antwort gibt.
GV_CONCENTRATION_ADD_FB_NEGATE_WA	0.05	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten addiert, wenn eine Lehrkraft eine Verneinung auf eine falsche SuS-Antwort gibt.
GV_CONCENTRATION_SUBT_FB_CRITICIZE_WA	0.1	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert (hier ist Kritik negativ zu werten), wenn eine Lehrkraft mit Kritik auf eine falsche SuS-Antwort reagiert.
GV_CONCENTRATION_SUBT_NOVALIDANSWER	0.1	Dieser Wert wird von der Konzentration einer Schülerin abgezogen, wenn die Lehrkraft einer Schülerin Feedback gibt, die gar keine Antwort zum aktuellen Task gegeben hat.
GV_CONCENTRATION_SUBT_CALLIFNOVALIDTASK	0.1	Dieser Wert wird von der Konzentration einer Schülerin abgezogen, wenn die Lehrkraft eine Schülerin aufruft, der kein valider Task zugewiesen ist
GV_CONCENTRATION_THRESHOLD_DISRUPTIVE_BEHAVIOR	0.3	Fällt die Konzentration einer Schülerin in einem Intervall unter diesen Wert, gibt es die Chance, dass Störverhalten auftritt. Die Wahrscheinlichkeit wird in GV_CHANCE_DISRUPTIVE_BEHAVIOR definiert.

Fortsetzung auf nächster Seite

GV_CONCENTRATION _ADD_CONSTRUCTIVE _WHOLECLASS	0.05	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten addiert, wenn eine Lehrkraft eine konstruktive Kritik an der ganzen Klasse äußert.
GV_CONCENTRATION _SUBT_CONSTRUCTIVE _WHOLECLASS	0.05	Dieser Wert wird zur Konzentration einer Schülerin ohne Störungsverhalten subtrahiert, wenn eine Lehrkraft eine konstruktive Kritik an der ganzen Klasse äußert.
GV_CONCENTRATION _ADD_DESTRUCTIVE _WHOLECLASS	-0.05	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten addiert (Achtung, deswegen standardgemäß negativ Wert) wenn eine Lehrkraft eine destruktive Kritik an der ganzen Klasse äußert.
GV_CONCENTRATION _SUBT_DESTRUCTIVE _WHOLECLASS	0.1	Dieser Wert wird zur Konzentration einer Schülerin ohne Störungsverhalten subtrahiert, wenn eine Lehrkraft eine destruktive Kritik an der ganzen Klasse äußert.
GV_CONCENTRATION _ADD_PRAISE_WHOLECLASS	0.02	Dieser Wert wird zur Konzentration einer Schülerin ohne Störungsverhalten addiert, wenn eine Lehrkraft die ganze Klasse lobt.
GV_CONCENTRATION _SUBT_PRAISE_WHOLECLASS	0.15	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft die ganze Klasse lobt.
GV_CONCENTRATION _ADD_AFFIRM_WHOLECLASS	0.03	Dieser Wert wird zur Konzentration einer Schülerin ohne Störungsverhalten addiert, wenn eine Lehrkraft die ganze Klasse im Verhalten bestärkt.
GV_CONCENTRATION _SUBT_AFFIRM_WHOLECLASS	0.1	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft die ganze Klasse im Verhalten bestärkt.
GV_CONCENTRATION _ADD_CONSTRUCTIVE_SINGLE	0.1	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten addiert, wenn eine Lehrkraft eine konstruktive Kritik an einer Einzelschülerin oder einer Gruppe äußert.
GV_CONCENTRATION _SUBT_CONSTRUCTIVE_SINGLE	0.2	Dieser Wert wird zur Konzentration einer Schülerin ohne Störungsverhalten subtrahiert, wenn eine Lehrkraft eine konstruktive Kritik an einer Einzelschülerin oder einer Gruppe äußert.
GV_CONCENTRATION _ADD_DESTRUCTIVE_SINGLE	-0.15	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten addiert (Achtung, deswegen standardgemäß negativer Wert!), wenn eine Lehrkraft eine destruktive Kritik an einer Einzelschülerin oder einer Gruppe äußert.
GV_CONCENTRATION _SUBT_DESTRUCTIVE_SINGLE	0.25	Dieser Wert wird zur Konzentration einer Schülerin ohne Störungsverhalten subtrahiert, wenn eine Lehrkraft eine destruktive Kritik an der ganzen Klasse äußert.
GV_CONCENTRATION _ADD_PRAISE_SINGLE	0.1	Dieser Wert wird zur Konzentration einer Schülerin ohne Störungsverhalten addiert, wenn eine Lehrkraft einer Einzelschülerin oder eine Gruppe lobt.
GV_CONCENTRATION _SUBT_PRAISE_SINGLE	0.1	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft eine Einzelschülerin oder eine Gruppe lobt.
GV_CONCENTRATION _ADD_AFFIRM_SINGLE	0.12	Dieser Wert wird zur Konzentration eines Schülers ohne Störungsverhalten addiert, wenn eine Lehrkraft eine Einzelschülerin oder eine Gruppe im Verhalten bestärkt.

Fortsetzung auf nächster Seite

GV_CONCENTRATION _SUBT_AFFIRM_SINGLE	0.1	Dieser Wert wird zur Konzentration einer Schülerin mit Störungsverhalten subtrahiert, wenn eine Lehrkraft eine Einzelschülerin oder eine Gruppe im Verhalten bestärkt.
GV_CONCENTRATION _SUBT_TOO_MUCH_CM	0.04	Dieser Wert wird zur Konzentration einer Schülerin subtrahiert, wenn die Lehrkraft eine zweite <i>classroom management</i> Aktion innerhalb von weniger Intervalle initiiert. Die Anzahl der Intervalle ist in GV_MAX_INTERVALLS_SINCE_CM_TO_BE_EFFECTIVE definiert.
GV_CONCENTRATION _SUBT_HASALREADYTASK	0.1	Dieser Wert wird zur Konzentration einer Schülerin subtrahiert, wenn die Lehrkraft eine zweite Aufgabe an die betreffende Schülerin vergibt, während die Schülerin noch über die erste Aufgabe nachdenkt.
GV_CONCENTRATION _ADD_TASKTOOEASY	0.1	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet, über einen Task nachzudenken, der eigentlich zu leicht für die Schülerin ist.
GV_CONCENTRATION _ADD_TASKNORMAL	0.15	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet, über einen Task nachzudenken, der normal für die Schülerin lösbar ist.
GV_CONCENTRATION _ADD_TASKOPTIMAL	0.2	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet, über einen Task nachzudenken, der sich auf einem optimalen Schwierigkeitsniveau für die Schülerin befindet.
GV_CONCENTRATION _ADD_TASKNOTSOLVEABLE	0.1	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet über einen Task nachzudenken, der zu schwer für die Schülerin ist. (und damit nicht lösbar)
GV_CONCENTRATION _ADD_INFOTOOEASY	0.03	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet über eine Information nachzudenken, die eigentlich zu leicht für die Schülerin ist.
GV_CONCENTRATION _ADD_INFONORMAL	0.5	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet, über eine Information nachzudenken, die normal passend ist.
GV_CONCENTRATION _ADD_INFOOPTIMAL	0.7	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet, über eine Information nachzudenken, die sich auf einem optimalen Schwierigkeitsniveau für die Schülerin befindet.
GV_CONCENTRATION _ADD_INFOTOOHARD	0.1	Dieser Wert wird zur Konzentration einer Schülerin addiert, wenn diese sich entscheidet, über eine Information nachzudenken, der zu schwer für die Schülerin ist.
GV_MAX_INTERVALLS_SINCE _CM_TO_BE_EFFECTIVE	5	Eine Lehrkraft muss nach Durchführung einer <i>classroom management</i> -Aktion diese Anzahl an Intervallen warten, bis sie eine weitere <i>classroom management</i> -Aktion effektiv durchführen kann. Führt sie während dieser Zeit CM-Aktionen durch, bewirken diese nichts und den Zielen wird GV_CONCENTRATION_SUBT_TOO_MUCH_CM von ihrer Konzentration abgezogen.

Fortsetzung auf nächster Seite

GV_MAX_INTERVALLS_SINCE _INFO_TO_BE_EFFECTIVE	5	Eine Lehrkraft muss nach Durchführung einer Information Aktion diese Anzahl an Intervallen warten, bis sie eine weitere Information-Aktion effektiv durchführen kann. Führt sie während dieser Zeit Information-Aktionen durch, bewirken diese nichts und den Zielen wird ein Teil von ihrer Konzentration abgezogen.
GV_MAX_INTERVALLS_SINCE _DISRUPT_TO_BE_EFFECTIVE	5	Liegt die Störung einer Schülerin mehr Intervalle als dieser Wert zurück, so ist <i>classroom management</i> nicht mehr effektiv.
GV_DURATION_MIN _OF_NEUTRAL	10	Diese Anzahl an Intervallen ist die Emotion ‚Neutral‘ mindestens stabil, außer es finden emotionsverändernde Ereignisse statt.
GV_DURATION_MIN _OF_JOY	10	Diese Anzahl an Intervallen ist die Emotion ‚Freude‘ mindestens stabil, außer es finden emotionsverändernde Ereignisse statt.
GV_DURATION_MIN _OF_SORROW	10	Diese Anzahl an Intervallen ist die Emotion ‚Traurigkeit‘ mindestens stabil, außer es finden emotionsverändernde Ereignisse statt.
GV_DURATION_MIN _OF_FEAR	10	Diese Anzahl an Intervallen ist die Emotion ‚Angst‘ mindestens stabil, außer es finden emotionsverändernde Ereignisse statt.
GV_DURATION_MIN _OF_ANGER	10	Diese Anzahl an Intervallen ist die Emotion ‚Wut‘ mindestens stabil, außer es finden emotionsverändernde Ereignisse statt.
GV_DURATION_MIN _OF_SURPRISE	1	Diese Anzahl an Intervallen ist die Emotion ‚Überraschung‘ mindestens stabil, außer es finden emotionsverändernde Ereignisse statt.

Fragebogen

Zur Beantwortung des Fragebogens und zum Anschauen und ggf. Zurückspulen der Mockup-Videos wurden 60 min veranschlagt.

- (1) **Geschlecht** (Männlich / Weiblich / Divers)
- (2) **Studiengang** (L1,L2,L3,L5, Erziehungswissenschaft, Anderer)
- (3) **Fachsemester des Lehramts- oder Erziehungswissenschaftsstudiengangs** (1,2,3,4,5,6,7,8 oder mehr) [Anmerkung: Wird nun angezeigt, wenn in Studiengang nicht „Anderer“ ausgewählt wurde]
- (4) **TPACK Modell - TK**

Unten finden Sie eine Reihe von Aussagen. Bitte geben Sie an, wie weit Sie den folgenden Aussagen zustimmen bzw. nicht zustimmen. (Likert Skala 5-stufig, stimme nicht zu, stimme eher nicht zu, neutral, stimme eher zu, stimme zu)

 - Ich weiß, wie ich meine eigenen technischen Probleme lösen kann.
 - Ich lerne, schnell mit Technologien umzugehen.
 - Ich halte mich über wichtige neue Technologien auf dem Laufenden.
 - Ich spiele häufig mit Technologien herum.
 - Ich kenne mich mit vielen verschiedenen Technologien aus.
 - Ich besitze die technischen Fähigkeiten, die ich brauche, um Technologien zu nutzen.
 - Ich hatte bereits genügend Möglichkeiten, um mit verschiedenen Technologien zu arbeiten.
- (5) **INCOBI-R (2009) / FIDEC. Nur Subskalen Persönliche Erfahrung/Lernen und Arbeiten/Nützliches Werkzeug (abgekürzt NW) und Persönliche Erfahrung/Lernen und Arbeiten/Unbeeinflussbare Maschine (abgekürzt UM)** Auf dieser Seite sind insgesamt 20 wertende Aussagen aufgeführt. Wir möchten Sie bitten, jeweils anzugeben, in welchem Ausmaß Sie den Aussagen zustimmen. (Likert Skala 5-stufig, stimme nicht zu, stimme eher nicht zu, neutral, stimme eher zu, stimme zu, randomisierte Frageposition) **FIDEC Subskala Unbeeinflussbare Maschine:**
 - [FIDEC013, UM] Um den Computer als Lernmittel zu verwenden, ist er mir zu unzuverlässig.

- [FIDEC014, UM] Ich ärgere mich oft darüber, dass der Computer für normale Menschen einfach nicht verstehbar ist.
- [FIDEC017, UM] Wenn ich am Computer arbeite, habe ich permanent Angst, er könnte „abstürzen“.
- [FIDEC026, UM] Die Arbeit am Computer ist oft frustrierend, weil ich diese Maschine nicht verstehe.
- [FIDEC027, UM] Der Computer macht manchmal Sachen, die ich nicht verstehe und nicht erklären kann.
- [FIDEC050, UM] Die Computerprogramme, die ich beim Lernen und Arbeiten nutze, sind manchmal schwer zu durchschauen.
- [FIDEC056, UM] Wenn ich am Computer arbeite, habe ich manchmal das Gefühl, das Ding macht, was es will.
- [FIDEC058, UM] Wenn mir mein Computer bei der Arbeit Probleme macht, fühle ich mich hilflos.
- [FIDEC062, UM] Manchmal bereitet das Arbeiten am Computer ganz schön viel Ärger.
- [FIDEC068, UM] Ich würde mir wünschen, weniger mit Computern arbeiten zu müssen.

FIDEC Subskala Nützliches Werkzeug:

- [FIDEC008, NW] Es gibt viele Arbeiten, die ich mit dem Computer leichter und schneller verrichten kann als ohne.
- [FIDEC016, NW] Für mich ist der Computer ein nützliches Arbeitsmittel.
- [FIDEC021, NW] Ich kann mir ein Arbeiten ohne den Computer kaum noch vorstellen.
- [FIDEC022, NW] Den Computer und die Neuen Medien empfinde ich als eine wesentliche Bereicherung meiner Ausbildung.
- [FIDEC023, NW] Bei einem großen Teil der arbeits- oder ausbildungsbezogenen Tätigkeiten, die ich zu verrichten habe, ist für mich der Computer ein nützliches Gerät.
- [FIDEC031, NW] Viele Arbeiten, wie zum Beispiel das Verfassen von Texten, gehen mit dem Computer einfach leichter und schneller.
- [FIDEC033, NW] Der Computer ist für mich das wichtigste Arbeitsmittel überhaupt.

- [FIDEC047, NW] Ich würde es begrüßen, wenn Computer und Neue Medien häufiger für Lehr- und Ausbildungszwecke genutzt würden.
 - [FIDEC057, NW] Ich finde es praktisch, für meine Arbeit oder meine Ausbildung einen Computer zur Verfügung zu haben.
 - [FIDEC061, NW] Das Lernen am Computer ist eine sehr effiziente Form des Lernens.
- (6) Haben Sie bereits Erfahrungen mit einem anderen pädagogischen Klassenraumsimulator gemacht? (Nein, Ja. Name des Simulators: (wenn unbekannt, bitte kurz Ort der Nutzung und Einsatzzweck beschreiben)/Freitext)
- (7) Wie oft spielen Sie Videospiele... (Einmal oder mehrmals pro Tag, Einmal oder mehrmals pro Woche, seltener, gar nicht)
- auf dem Computer
 - auf einer Spielekonsole (z.B. Playstation, Nintendo Switch, Xbox)
 - auf einem Smartphone/Mobiltelefon
 - auf einem Tablet (z.B. iPad)
- (8) VR-Vorerfahrungen: Wie oft haben Sie eine VR/AR-Brille (z.B. HTC Vive, Oculus Rift, Valve Index, Google Cardboard, Microsoft HoloLens) verwendet? (sehr oft, oft, selten, noch nie)
- (9) Besitzt der Haushalt, in dem Sie leben, eine VR-Brille? (Ja, Nein)
- (10) Dass im Bereich Lehramtsstudium & Erziehungswissenschaft die Möglichkeiten der VR-Technik beforscht werden, finde ich... (sehr gut, eher gut, weder gut noch schlecht, eher schlecht, sehr schlecht)

Im Fragebogen folgen nun die Einführungs- und Erklärungsvideos sowie die Darbietung der Mockup-Videos. Für eine Übersicht zu den Gruppen und den einzelnen Videos siehe [Tabelle 4](#) (S. 195) und [Tabelle 5](#) (S. 195).

- (11) Wie sehr unterschieden sich die Videos im Realitätsgrad der Schülerinnen und Schüler? (Likert Skala, 4-stufig, sehr unterschiedlich, unterschiedlich, kaum unterschiedlich)
- (12) Wie schwierig fanden Sie es den Realitätsgrad des von den Schülerinnen und Schülern gezeigten Verhalten zu beurteilen? (Likert Skala, 4-stufig, sehr schwierig, eher schwierig, eher einfach, sehr einfach)

- (13) Ich konnte den Realitätsgrad der Schülerinnen und Schüleraktionen...
(schlecht einschätzen, eher schlecht einschätzen, eher gut einschätzen,
gut einschätzen)

Beurteilung des ersten Videos

Einleitungstext: Bitte beurteilen Sie nun ausschließlich das erste Video, das sie gesehen haben.

Bitte nutzen Sie die Möglichkeit, sich die Videos noch einmal anzusehen und zu vergleichen. Nutzen Sie auch die Möglichkeit im Video zu springen.

Sie finden die Videos am Ende dieser Seite.

Versetzen Sie sich in die Position einer beobachtenden Vertretungslehrkraft, die die Hintergrundgeschichte der einzelnen SuS nicht kennt.

Bitte bewerten Sie nur die Aktionen der Schülerinnen und Schüler. Stellen Sie sich vor, wenn Sie durch das ggf. als unrealistisch empfundene Verhalten der Lehrkraft sehr beeinflusst werden, dass es sich um eine Lehrkraft mit wenig Lehrerfahrung handelt.

- (14) Wie realistisch fanden Sie das von den SuS gezeigte... (Likert Skala, 4-stufig, sehr realistisch, eher realistisch, eher nicht realistisch, nicht realistisch)

- Störungsverhalten
- Meldeverhalten
- Antwortverhalten

- (15) Bitte beurteilen Sie den Realitätsgrad der Häufigkeit des gezeigten Störungsverhaltens aller Schülerinnen und Schüler (Likert Skala, 5-stufig, zu häufig, eher zu häufig, genau realistisch, eher zu selten, zu selten)

- (16) Bitte beurteilen Sie den Realitätsgrad der Häufigkeit des gezeigten Meldeverhaltens aller Schülerinnen und Schüler (Likert Skala, 5-stufig, zu häufig, eher zu häufig, genau realistisch, eher zu selten, zu selten)

- (17) Bitte beurteilen Sie den Realitätsgrad der Häufigkeit des gezeigten Antwortverhaltens aller Schülerinnen und Schüler (Likert Skala, 5-stufig, Zu viele richtige Antworten, eher zu viele richtige Antworten, genau realistisch, eher zu viele falsche Antworten, zu viele falsche Antworten)

- (18) Fehlte Ihrer Meinung nach ein wichtiges Störverhalten, das sie erwartet hätten? (Nein, Ja: Um welches Störverhalten handelt es sich?/Freitext)

Beurteilung des zweiten Videos Wiederholung des Einleitungstextes und der Fragen 14-18 bezogen auf das zweite Video. (resultierende Fragen: 19-23)

(24) (a) *Nur 2D-Gruppen:*

Wie sehr hat Sie folgende Eigenschaften der grafischen Aufbereitung eine Bewertung der Schülerinnen und Schüler erschwert? (Likert Skala, 4-stufig, sehr erschwert, eher erschwert, eher nicht erschwert, gar nicht erschwert)

- Grafische Aufbereitung als 2D-Darstellung
- Verzögerung zwischen Grafik und gesprochenem Text
- Nicht-Beachtung der konkret gesprochenen Schülerinnen- und Schüleraussagen

(b) *Nur 3D-Gruppen:*

Wie sehr hat Sie folgende Eigenschaften der grafischen Aufbereitung eine Bewertung der Schülerinnen und Schüler erschwert? (Likert Skala, 4-stufig, sehr erschwert, eher erschwert, eher nicht erschwert, gar nicht erschwert)

- Qualität der 3D-Figuren
- Qualität der gezeigten Animationen
- Verzögerung zwischen Grafik und gesprochenem Text
- Nicht-Beachtung der konkret gesprochenen Schülerinnen- und Schüleraussagen

(25) **Ein Training mit einer digitalen Klassenraumsimulation wie Clasivir (eingebettet in eine Lehrveranstaltung) fände ich...** (Likert Skala, 4-stufig, sehr nützlich, eher nützlich, eher nicht nützlich, nicht nützlich)

(26) Die Entwicklung eines Simulators, der andere Unterrichtsbereiche oder Lehrkrafthandeln thematisiert, (z.B. Gruppenarbeiten, Elterngespräche, etc.) fände ich... (Likert Skala, 4-stufig, sehr sinnvoll, eher sinnvoll, eher nicht sinnvoll, nicht sinnvoll)

(27) **Folgendes möchte ich noch mitteilen:** (Freitext)

Literatur

- Abt, C. C. (1970). *Serious Games*. Viking Press.
- Affective Social Computing Laboratory. (2016). *3D IVT Interactive virtual Training for Early Career Teachers*. <https://web.archive.org/web/20220521212145/http://ascl.cis.fiu.edu/interactive-virtual-training.html>
- Aïmeur, E., Frasson, C., & Alexe, C. (1995). Towards new learning strategies in intelligent tutoring systems. In J. Wainer & A. Carvalho (Hrsg.), *Advances in Artificial Intelligence* (S. 121–130). Springer.
- Akilli, G. K. (2007). Games and Simulation: A New Approach in Education? In D. Gibson, C. Aldrich & M. Prensky (Hrsg.), *Games and Simulation in Online Learning: Research and Development Frameworks* (S. 1–20). Information Science Publishing.
- Albright, G., Adam, C., Serri, D., Bleeker, S., & Goldman, R. (2016). Harnessing the power of conversations with virtual humans to change health behaviors. *mHealth*, 2(11), 1–13. <https://doi.org/10.21037/mhealth.2016.11.02>
- Albright, G., Goldman, R., Shockley, K., & Spiegler, J. (2013a). *Using Simulated Conversations with Virtual Humans to Build Mental Health Skills among Educators, Staff, and Students. A Summary of Five Longitudinal Studies*. https://go.kognito.com/rs/143-HCJ-270/images/HiEd_PK-12_At-Risk_Suite.pdf
- Albright, G., Goldman, R., Shockley, K., & Spiegler, J. (2013b). *Using Simulated Conversations with Virtual Humans to Build Mental Health Skills among Middle School Educators. A Longitudinal Study* [Die ursprüngliche Studie wurde im Verlauf des Schreibens dieser Arbeit aktualisiert, mit weitestgehend identischem Inhalt. Die ursprüngliche Studie war unter <https://kognito.com/wp-content/uploads/At-Risk-Middle-Research.pdf> aufzufinden, ist nun aber nicht mehr aufzufinden.]. https://go.kognito.com/rs/143-HCJ-270/images/PK-12_MiddleSchool_ResearchStudy.pdf
- Aldrich, C. (2009). *The complete guide to simulations and serious games: How the most valuable content will be created in the age beyond Gutenberg to Google*. John Wiley & Sons.
- Alessi, S. M., & Trollip, S. R. (2001). *Multimedia for learning. Methods and development* (3. Edition). Allyn; Bacon.
- Alinier, G. (2010). Developing High-Fidelity Health Care Simulation Scenarios: A Guide for Educators and Professionals. *Simulation & Gaming*, 42(1), 9–26. <https://doi.org/10.1177/1046878109355683>

- Allen, D. W. (1967). *Micro-Teaching: A Description*. Stanford University. <https://files.eric.ed.gov/fulltext/ED019224.pdf>
- Allen, D. W., & Eve, A. W. (1968). Microteaching. *Theory Into Practice*, 7(5), 181–185. <http://www.jstor.org/stable/1475985>
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2000). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives, Abridged Edition*. Pearson.
- Andreasen, J. B., & Haciomeroglu, E. S. (2009). Teacher Training in Virtual Environments. In S. L. Swars, D. W. Stinson & S. Lemons-Smith (Hrsg.), *Proceedings of the 31st annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (S. 1317–1324). Georgia State University.
- Ansorge, U., & Leder, H. (2017). *Wahrnehmung und Aufmerksamkeit*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-12912-5>
- Appel, J. (2015). *Die Bedeutung der Aufgaben für das Beteiligungsverhalten der Schüler. Eine Videostudie zur Wirksamkeit des Unterrichtsprozesses* (Dissertationsschrift). Johann Wolfgang Goethe-Universität, Fachbereich Erziehungswissenschaft. Frankfurt am Main, Univ.-Bibliothek Frankfurt am Main. <http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/38418>
- Aronin, S. (2011, März). A New Generation in Teacher Preparation: TeachME (Teaching in a Mixed Reality Environment). In M. Koehler & P. Mishra (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2011* (S. 2034–2036). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/36602>
- Artist Made ProductionsPlus. (2011, 21. Februar). *The Cybersmart Connect.Ed Promo*. <https://vimeo.com/20189542>
- Badiee, F., & Kaufman, D. (2015). Design Evaluation of a Simulation for Teacher Education. *SAGE Open*, 5(2), 1–10. <https://doi.org/10.1177/2158244015592454>
- Baek, Y. (2009). Digital Simulations in Teaching and Learning. In D. Gibson & Y. Baek (Hrsg.), *Digital Simulations for Improving Education: Learning Through Artificial Teaching Environments* (S. 25–51). IGI Global / Information Science Reference.
- Bailenson, J. N., Yee, N., Blascovich, J., Beall, A. C., Lundblad, N., & Jin, M. (2008). The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social

- Context. *Journal of the Learning Sciences*, 17(1), 102–141. <https://doi.org/10.1080/10508400701793141>
- Baker, R., Corbett, A. T., Koedinger, K. R., Evenson, S., Roll, I., Wagner, A. Z., Naim, M., Raspat, J., Baker, D. J., & Beck, J. E. (2006). Adapting to When Students Game an Intelligent Tutoring System. In M. Ikeda, K. D. Ashley & T.-W. Chan (Hrsg.), *Intelligent Tutoring Systems* (S. 392–401). Springer Berlin Heidelberg.
- Baker, R., Walonoski, J. A., Heffernan, N. T., Roll, I., Corbett, A. T., & Koedinger, K. R. (2008). Why Students Engage in “Gaming the System” Behavior in Interactive Learning Environments. *Journal of Interactive Learning Research*, 19(2), 185–224.
- Barbour, M. K. (2007). Portrait of Rural Virtual Schooling. *Canadian Journal of Educational Administration and Policy (CJEAP)*, 59, 1–21. <https://eric.ed.gov/?id=EJ806974>
- Barbour, M. K. (2015). Real-Time Virtual Teaching: Lessons Learned From a Case Study in a Rural School. *Online Learning*, 19(5), 54–68. <https://eric.ed.gov/?id=EJ1085762>
- Barmaki, R., & Hughes, C. E. (2015). Providing Real-time Feedback for Student Teachers in a Virtual Rehearsal Environment. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 531–537. <https://doi.org/10.1145/2818346.2830604>
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74(5), 1252–1265. <https://doi.org/10.1037/0022-3514.74.5.1252>
- Baylor, A. L., & Kim, Y. (2005). Simulating Instructional Roles Through Pedagogical Agents. *International Journal of Artificial Intelligence in Education*, 15(2), 95–115. <http://dl.acm.org/citation.cfm?id=1434925.1434927>
- Billingsley, G. M., & Scheuermann, B. K. (2014). Using Virtual Technology to Enhance Field Experiences for Pre-Service Special Education Teachers. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, 37(3), 255–272. <https://doi.org/10.1177/0888406414530413>
- Bligh, D. A. (1972). *What’s the use of lectures?* Jossey-Bass.
- Blomberg, G., Gamoran, M., Alexander, S., Glogger, I., & Seidel, T. (2014). Understanding video as a tool for teacher education : investigating instructional strategies to promote reflection. *Instructional Science*, 42(3), 443–463. <https://doi.org/10.1007/s11251-013-9281-6>

- Blomberg, G. A. (2011). *Der Einsatz von Unterrichtsvideos in der universitären Lehrerbildung* (Dissertationsschrift). Technische Universität München.
- Bloom, B. S., Engelhart, M. D., Furst, E., Hili, W. H., & Krathwohl, D. R. (1976). *Taxonomie von Lernzielen im kognitiven Bereich* (E. Fünér & R. Horn, Übers.; 2. Edition). Beltz.
- Boekaerts, M. (1999). Self-regulated learning: where we are today. *International Journal of Educational Research*, 31(6), 445–457. [https://doi.org/10.1016/s0883-0355\(99\)00014-2](https://doi.org/10.1016/s0883-0355(99)00014-2)
- Böheim, R., Knogler, M., Kosel, C., & Seidel, T. (2020). Exploring student hand-raising across two school subjects using mixed methods: An investigation of an everyday classroom behavior from a motivational perspective. *Learning and Instruction*, 65, 101250. <https://doi.org/https://doi.org/10.1016/j.learninstruc.2019.101250>
- Borich, G. D. (2014, 26. August). *Observation Skills for Effective Teaching* (7, Neuauflage, überarbeitet). Taylor & Francis Ltd. http://www.ebook.de/de/product/21653456/gary_d_borich_observation_skills_for_effective_teaching.html
- Borich, G. D. (2016). *Effective teaching methods. Research-Based Practice* (9. Aufl.). Pearson.
- Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education*, 40(4), 509–513. <https://doi.org/10.1152/advan.00109.2016>
- Bradley, E. G., & Kendall, B. (2014). A Review of Computer Simulations in Teacher Education. *Journal of Educational Technology Systems*, 43(1), 3–12. <https://doi.org/http://dx.doi.org/10.2190/ET.43.1.b>
- Bradley, E. G., & Kendall, B. (2019). Training Teachers to Identify and Refer At-Risk Students Through Computer Simulation. *Journal of Technology in Behavioral Science*, 4(4), 340–345. <https://doi.org/10.1007/s41347-019-00106-w>
- Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. National Academies Press.
- Breitenbach, S., & Appel, J. (2016). Vergleich von Software-Paketen zur Analyse audiovisueller Daten. In U. Rauin, M. Herrle & T. Engartner (Hrsg.), *Videoanalysen in der Unterrichtsforschung. Methodische Vorgehensweisen und Anwendungsbeispiele* (S. 154–173). Beltz Juventa.
- Brophy, J. (1981). Teacher Praise: A Functional Analysis. *Review of Educational Research*, 51(1), 5–32. <https://doi.org/10.3102/00346543051001005>

- Brophy, J. (1999). Toward a model of the value aspects of motivation in education: Developing appreciation for.. *Educational Psychologist*, *34*(2), 75–85. https://doi.org/10.1207/s15326985ep3402_1
- Brophy, J. E., & Good, T. L. (1970). Teachers' communication of differential expectations for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, *61*(5), 365–374. <https://doi.org/https://doi.org/10.1037/h0029908>
- Brown, A. H. (1999). Simulated classrooms and artificial students: The potential effects of new technologies on Teacher Education. *Journal of Research on Computing in Education*, *32*(2), 307–318. <https://doi.org/https://doi.org/10.1080/08886504.1999.10782281>
- Brühwiler, C., Helmke, A., & Schrader, F.-W. (2017). Determinanten der Schulleistung. In M. K. Schweer (Hrsg.), *Lehrer-Schüler-Interaktion: Inhaltsfelder, Forschungsperspektiven und methodische Zugänge* (S. 291–314). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-15083-9_13
- Bulu, S. T. (2012). Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education*, *58*(1), 154–161. <https://doi.org/doi:10.1016/j.compedu.2011.08.024>
- Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How Long Can Students Pay Attention in Class? A Study of Student Attention Decline Using Clickers. *Journal of Chemical Education*, *87*(12), 1438–1443. <https://doi.org/10.1021/ed100409p>
- Burns, D. N. (1973). *Vicarious experience approach to preparing prospective teachers for inner-city teaching* (Dissertationsschrift). Iowa State University, Family, Consumer Sciences Education und Studies. Ames, Iowa, USA, Digital Repository @ Iowa State University. <https://doi.org/10.31274/rtd-180817-4148>
- Candia, C., Jara-Figueroa, C., Rodriguez-Sickert, C., Barabási, A.-L., & Hidalgo, C. A. (2018). The universal decay of collective memory and attention. *Nature Human Behaviour*, *3*(1), 82–91. <https://doi.org/10.1038/s41562-018-0474-5>
- Carmody, R. (2019). *Murdoch University SimLab virtual classrooms transform teacher training*. <https://www.abc.net.au/news/2019-11-03/murdoch-university-virtual-classrooms-used-in-teacher-training/11651650>
- Carrington, L., Kervin, L., & Ferry, B. (2006). Cognitive tools of Class-Sim: Building connections between theory and practice: 2006-11-26 – 2006-11-30. *Australian Association for Research in Education (AARE) Conference 36th Annual Conference*. <https://www.aare.edu.au/data/publications/2006/car06575.pdf>

- Cheong, D. (2010). The effects of practice teaching sessions in second life on the change in pre-service teachers' teaching efficacy. *Computers & Education*, 55(2), 868–880. <https://doi.org/https://doi.org/10.1016/j.compedu.2010.03.018>
- Cheong, D., & Kim, B. (2009). A Simulation for Improving Teachers' Motivational Skills. In D. Gibson & Y. Baek (Hrsg.), *Digital Simulations in Teaching and Learning: Learning Through Artificial Teaching Environments* (S. 227–248). IGI Global / Information Science Reference.
- Cheong, D., & Kim, S. (2009). 예비교사의훈육기술연습을위한시뮬레이터 개발 (Developing a Simulator for Practicing Discipline Skills of Pre-Service Teachers). *컴퓨터교육학회논문지 (The Journal of Korean Association of Computer Education)*, 12(3), 63–74. http://kiss.kstudy.com/journal/thesis_name.asp?tname=kiss2002&key=2780178
- Chieu, V. M., & Herbst, P. (2011). Designing an intelligent teaching simulator for learning to teach by practicing. *ZDM*, 43(1), 105–117. <https://doi.org/10.1007/s11858-010-0291-4>
- Chini, J. J., Straub, C. L., & Thomas, K. H. (2016). Learning from avatars: Learning assistants practice physics pedagogy in a classroom simulator. *Phys. Rev. Phys. Educ. Res.*, 12(1). <https://doi.org/10.1103/PhysRevPhysEducRes.12.010117>
- Cho, Y. H., Yim, S. Y., & Paik, S. (2015). Physical and social presence in 3D virtual role-play for pre-service teachers. *The Internet and Higher Education*, 25, 70–77. <https://doi.org/http://dx.doi.org/10.1016/j.iheduc.2015.01.002>
- Chou, C.-Y., Chan, T.-W., & Lin, C.-J. (2003). Redefining the learning companion: the past, present, and future of educational agents. *Computers & Education*, 40(3), 255–269. [https://doi.org/http://doi.org/10.1016/S0360-1315\(02\)00130-6](https://doi.org/http://doi.org/10.1016/S0360-1315(02)00130-6)
- Christensen, R. R., Knezek, G., Tyler-Wood, T., & Gibson, D. (2011). sim-School: an online dynamic simulator for enhancing teacher preparation. *International Journal of Learning Technology*, 6(2), 201–220. <https://doi.org/10.1504/IJLT.2011.042649>
- Chuang, Y.-W., & Shih, J.-L. (2014). 职前教师模拟低年级教师校园生活之严肃体验游戏设计与开发 (Designing and Developing a Blended Serious Game Simulating the Campus Life of Teachers of Primary Grades for Pre-service Teachers). *Proceedings of the Global Chinese Conference on Computers in Education (GCCCE) 2014*. <http://aic-fe.bnu.edu.cn/gccce2017/programme/proceedings/>
- Clarke, S., Lameris, P., & Arnab, S. (2016). SimAULA: Creating higher-level gamification through adoption of a learning-objective to game-objective mapping approach. *Proceedings of the 10th European Con-*

- ference on Games Based Learning, *ECGBL 2016, 2016-January*, 127–134.
- Clayton, G., & Gizelis, T.-I. (2005). Learning through Simulation or Simulated Learning. *Proceedings from British International Studies Association Conference*, 1–25.
- Collum, D. H. (2015). *Impacting Preservice Teachers' Understanding of the Educational Needs of Diverse Learners Through the Use of Simulation Based Learning* (Dissertationschrift). Missouri Baptist University. Missouri, USA. <https://pqdtopen.proquest.com/doc/1680274303.html?FMT=ABS>
- Collum, D. H., Christensen, R., Delicath, T., & Johnston, V. (2019). Sim-School: SPARCing New Grounds in Research on Simulated Classrooms. In K. Graziano (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2019* (S. 733–739). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/207723>
- Cruickshank, D. R. (1969). The Use of Simulation in Teacher Education: A Developing Phenomenon. *Journal of Teacher Education*, 20(1), 23–26. <https://doi.org/10.1177/002248716902000105>
- Cruickshank, D. R. (1971). Simulation as an Instructional Alternative in Teacher Education. *ATE Research Bulletin*, 8, 1–33. <https://eric.ed.gov/?id=ED053067>
- Cruickshank, D. R. (1972). The Notions of Simulations and Games: A Preliminary Inquiry. *Educational Technology*, 12(7), 17–19. <http://www.jstor.org/stable/44418557>
- Cruickshank, D. R. (1988). The Uses of Simulations in Teacher Preparation. *Simulation & Games*, 19(2), 133–156. <https://doi.org/10.1177/104687818801900202>
- Cruickshank, D. R., & Leonard, J. (1967). *The identification and analysis of perceived problems of teachers in inner-city schools*.
- Cunningham, L. L. (1959). News and Comment. *The Elementary School Journal*, 60(1), 1–13. <http://www.jstor.org/stable/999514>
- Czerwionka, T., Klebl, M., & Schrader, C. (2009). Die Einführung virtueller Klassenzimmer in der Fernlehre. Ein Instrumentarium zur nutzerorientierten Einführung neuer Bildungstechnologien. In N. Apostolopoulos, H. Hoffmann, V. Mansmann & A. Schwill (Hrsg.), *E-Learning 2009. Lernen im digitalen Zeitalter*. (S. 96–105). Waxmann. <http://www.waxmann-verlag.de/fileadmin/media/zusatztexte/2199Volltext.pdf#page=97;%20http://www.pedocs.de/volltexte/2011/3088;%20http://nbn-resolving.de/urn:nbn:de:0111-opus-30889>

- Dann, H.-D., & Humpert, W. (2002). Das Konstanzer Trainingsmodell (KTM) - Grundlagen und neue Entwicklungen. *Zeitschrift für Pädagogik*, 48(2), 215–226.
- Dawson, M. R. (2016). *From TeachLivE™ to the Classroom: Building Preservice Special Educators' Proficiency with Essential Teaching Skills* (Dissertationsschrift). Utah State University, Special Education und Rehabilitation Department. <https://digitalcommons.usu.edu/etd/4930>
- Dawson, M. R., & Lignugaris/Kraft, B. (2017). Meaningful Practice: Generalizing Foundation Teaching Skills From TLE TeachLivE™ to the Classroom. *Teacher Education and Special Education*, 40(1), 26–50. <https://doi.org/10.1177/0888406416664184>
- Deale, D., & Pastore, R. (2014). Evaluation of simSchool: An Instructional Simulation for Pre-Service Teachers. *Computers in the Schools*, 31(3), 197–219. <https://doi.org/10.1080/07380569.2014.932650>
- de Castell, S., Jenson, J., & Thumlert, K. (2014). From Simulation to Imitation. *Simulation & Gaming*, 45(3), 332–355. <https://doi.org/10.1177/1046878114542316>
- de Jong, T., Lane, J., & Sharp, S. (2012). The Efficacy of Simulation as a Pedagogy in Facilitating Pre-Service Teachers' Learning About Emotional Self-Regulation and its Relevance to the Teaching Profession. *Australian Journal of Teacher Education*, 37(3). <https://doi.org/10.14221/ajte.2012v37n3.6>
- Delamarre, A. P., Buche, C., & Lisetti, C. (2019). AIMER: Appraisal Interpersonal Model of Emotion Regulation, Affective Virtual Students to Support Teachers Training. *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, 182–184. <https://doi.org/10.1145/3308532.3329419>
- Delamarre, A. P., Buche, C., Polceanu, M., Lunn, S., Ruiz, G., Bolivar, S., Shernoff, E., & Lisetti, C. L. (2017). An Interactive Virtual Training (IVT) Simulation for Early Career Teachers to Practice in 3D Classrooms with Student Avatars. *30th International FLAIRS Conference. Marco Island, Florida. May 22–24, 2017*, 400–403. <https://aaai.org/ocs/index.php/FLAIRS/FLAIRS17/paper/view/15487>
- Delamarre, A. P., Lunn, S., Buche, C., Shernoff, E., Frazier, S., & Lisetti, C. (2019). Interdisciplinary Collaboration and Establishment of Requirements for a 3D Interactive Virtual Training for Teachers. *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, 185–187. <https://doi.org/10.1145/3308532.3329439>
- Demoulin, Y., Joseph, O., Krikhaar, L., & Reuter, R. A. (2012). *SimSchool 2.0. Simulator Training makes better Teachers, better Teacher Educators and better Learning & Teaching Researchers* (Final Team Project

- Report of "MOOC: Designing a New Learning Environment"). Stanford University.
- Denton, J. (2013, 8. Februar). *Cook School District (Video)*. <https://www.youtube.com/watch?v=dKuVyZEoo4A>
- Dick, W., Carey, L., & Carey, J. O. (2004). *The Systematic Design of Instruction* (6. Edition). Allyn & Bacon.
- Diefenbach, P., Oum, K., & Reisman, F. (2019). Transition to Teaching: 3D classroom simulation and management. *2nd International IEEE Consumer Electronics Society's Games Innovations Conference Proceedings*, 196–200. <https://doi.org/10.1109/ICEGIC.2010.5716910>
- Dieker, L. A., Grillo, K., & Ramlakhan, N. (2012). The use of virtual and simulated teaching and learning environments: Inviting gifted students into science, technology, engineering, and mathematics careers (STEM) through summer partnerships. *Gifted Education International*, 28(1), 96–106. <https://doi.org/10.1177/0261429411427647>
- Dieker, L. A., Hughes, C. E., Hynes, M. C., & Straub, C. (2017). Using Simulated Virtual Environments to Improve Teacher Performance. *School University Partnerships: Special Issue: Technology to Enhance PDS*, 10(3), 62–81.
- Dieker, L. A., Hynes, M., Hughes, C., & Smith, E. (2008). Implications of Mixed Reality and Simulation Technologies on Special Education and Teacher Preparation. *Focus on exceptional children*, 40(6), 1–20.
- Dieker, L. A., Hynes, M., Stapleton, C., & Hughes, C. (2007). Virtual Classrooms: STAR Simulator. Building Virtual Environments for Teacher Training in Effective Classroom Management. *New Learning Technology SALT@*, 4, 1–22.
- Dieker, L. A., Hynes, M. C., Hughes, C. E., Hardin, S., & Becht, K. (2015). TLE TeachLivE™: Using Technology to Provide Quality Professional Development in Rural Schools. *Rural Special Education Quarterly*, 34(3), 11–16. <https://doi.org/10.1177/875687051503400303>
- Dieker, L. A., Kennedy, M. J., Smith, S., III, E. V., Rock, M., & Thomas, C. (2014). *Use of Technology in the Preparation of Pre-Service Teachers: (Document No. IC-11)*. http://cedar.education.ufl.edu/wp-content/uploads/2014/10/IC-11_FINAL_05-26-15.pdf
- Dieker, L. A., Rodriguez, J. A., Lignugaris/Kraft, B., Hynes, M. C., & Hughes, C. E. (2014). The Potential of Simulated Environments in Teacher Education: Current and Future Possibilities. *Teacher Education and Special Education*, 37(1), 21–33. <https://doi.org/https://doi.org/10.1177/0888406413512683>
- Dillenbourg, P., Schneider, D., & Synteta, P. (2002). Virtual Learning Environments. In A. Dimitracopoulou (Hrsg.), *3rd Hellenic Conference*

- "*Information & Communication Technologies in Education*" (S. 3–18). Kastaniotis Editions, Greece. <https://telearn.archives-ouvertes.fr/hal-00190701>
- Dimmel, J., Herbst, P., Gürsel, U., Milewski, A., & Erickson, A. (2016). The Lessonsketch Environment: A Platform for Online Teacher Education and Research on Teaching. *13th International Congress on Mathematical Education*, 1–6. https://www.researchgate.net/publication/317312664_THE_LESSONSKETCH_ENVIRONMENT_A_PLATFORM_FOR_ONLINE_TEACHER_EDUCATION_AND_RESEARCH_ON_TEACHING
- Djaouti, D., Alvarez, J., Jessel, J.-P., & Rampnoux, O. (2011). Origins of Serious Games. In *Serious Games and Edutainment Applications* (S. 25–43). Springer London. https://doi.org/10.1007/978-1-4471-2161-9_3
- Doyle, W. (1986). Classroom Organization and Management. In M. Wittrock (Hrsg.), *Handbook of Research on Teaching* (S. 392–431). Macmillan.
- Drexel University. (2010a). *Promotionsvideo zu Transition to Teaching*. http://replay.drexel.edu/videos/TTT_Trailer_compress.mov
- Drexel University. (2010b). *Promotionsvideo zu Transition to Teaching 2*. http://replay.drexel.edu/videos/TTT_comp.mov
- Eason, G. (2002, 13. November). *Teachers get to make virtual decisions*. http://news.bbc.co.uk/2/hi/uk_news/education/2453907.stm
- Edwards, R. (1865). Normal schools in the United States [Reproduziert in Merle, Borrowman, *Teacher Education in America, A Documentary History* (New York: Teachers College, Columbia University, 1956), S.77–78]. *Lectures and Proceedings of the National Teacher Association (Harrisburg, PA: NEA)*, 277–282.
- Egbert, J. L., Thomas, M. K., & Fischler, R. (2000). Assessing the Tigerlake Public Schools Simulation. *Journal of Computing in Teacher Education*, 16(3), 23–27. <https://doi.org/10.1080/10402454.2000.10784391>
- Enterprise Florida. (2009, 14. Januar). *Simulation Video University of Central Florida*. https://youtu.be/_9VVAFW1Rx8
- Falck, A.-K., Husu, J., Kronlund, T., Kynäslähti, H., Salminen, J., & Salonen, M. (1997). Testing virtual classroom in the school context. *Distance Education*, 18(2), 213–224. <https://doi.org/10.1080/0158791970180203>
- Farrow, D., Chivers, P., Hardingham, C., & Sachse, S. (1998). The effect of video-based perceptual training on the tennis return of serve. *International Journal of Sport Psychology*, 29(3), 231–242.
- Fatimah, S., Setiawan, W., Kusnendar, J., Anggraeni, R., & Junaeti, E. (2015). Teaching Simulator for Development of Professional Teacher.

International Seminar on Mathematics, Science, and Computer Science Education (MSCEIS) 2015 Proceedings, 942–946.

- Fatimah, S., Setiawan, W., Kusnendar, J., Rasim, Junaeti, E., & Anggraeni, R. (2017). Development of the teaching simulator based on animated film to strengthening pedagogical competencies of prospective teachers. *AIP Conference Proceedings*, 1848(1), 1–5. <https://doi.org/10.1063/1.4983987>
- Feinstein, A. H., Mann, S., & Corsun, D. L. (2002). Charting the experiential territory. Clarifying definitions and uses of computer simulation, games, and role play. *Journal of Management Development*, 21(10), 732–744. <https://doi.org/10.1108/02621710210448011>
- Ferry, B., & Kervin, L. (2007). Developing an Online Classroom Simulation to Support a Pre-Service Teacher Education Program. In D. Gibson, C. Aldrich & M. Prensky (Hrsg.), *Games and Simulations in Online Learning: Research and Development Frameworks* (S. 189–205). Information Science Publishing. <https://doi.org/10.4018/978-1-59904-304-3.ch009>
- Ferry, B., Kervin, L., Cambourne, B., Turbill, J., Jonassen, D., Hedberg, J., & Puglisi, S. (2004, 5. Dezember). Online classroom simulation: The next wave for pre-service teacher education? In R. Atkinson, C. McBeath, D. Jonas-Dwyer & R. Phillips (Hrsg.), *Beyond the Comfort Zone: Proceedings of the 21st ASCILITE Conference*. ASCILITE. <http://www.ascilite.org.au/conferences/perth04/procs/ferry.html>
- Ferry, B., Kervin, L., Hedberg, J., Jonassen, D., Cambourne, B., Turbill, J., & Carrington, L. (2005). The development of ClassSim: a simulated learning environment to support the practicum experience for pre-service teachers. In C.-K. Looi, D. Jonassen & M. Ikeda (Hrsg.), *Towards Sustainable and Scalable Educational Innovations Informed by the Learning Sciences* (S. 84–92). IOS Press.
- Fiedler, K. (2012). The Formation of Attitudes and Social Judgements in a Virtual School Class Environment. In J. P. Forgas, K. Fiedler & C. Sedikides (Hrsg.), *Social Thinking and Interpersonal Behavior* (S. 183–200). Psychology Press.
- Fiedler, K., Freytag, P., & Unkelbach, C. (2007). Pseudocontingencies in a simulated classroom. *Journal of Personality and Social Psychology*, 92(4), 665–677. <https://doi.org/10.1037/0022-3514.92.4.665>
- Fiedler, K., Walther, E., Freytag, P., & Plessner, H. (2002). Judgment Biases in a Simulated Classroom - A Cognitive-Environmental Approach. *Organizational Behavior and Human Decision Processes*, 88(1), 527–561. <https://doi.org/10.1006/obhd.2001.2981>

- Fischer, J., Walter, J., Möller, J., & Harms, U. (2017, 3. Juli). ProSim: Prozedurales Professionswissen im Simulierten Klassenraum. Präsentation zur BMBF-Fachtagung "Hochschule im digitalen Zeitalter" (3. und 4. Juli 2017). https://www.e-teaching.org/etresources/pdf/best-practice_2018_fischer-walter-moeller-harms_neues-prosim-prozedurales-professionswissen-im-simulierten-klassenraum.pdf
- Fischler, R. B. (2004). *SimTeacher Blog*. <http://simteacher.blogspot.com/>
- Fischler, R. B. (2006). *SimTeacher. Simulation-based learning in teacher education* (Dissertationsschrift). Indiana University, School of Education. <http://hdl.handle.net/2022/7151>
- Fischler, R. B. (2007). SimTeacher.com: An Online Simulation Tool for Teacher Education. *TechTrends*, 51(1), 44–47. <https://doi.org/10.1007/s11528-007-0011-2>
- Floyd, E. M., Rayfield, A., Eyberg, S. M., & Riley, J. L. (2004). Psychometric Properties of the Sutter-Eyberg Student Behavior Inventory with Rural Middle School and High School Children. *Assessment*, 11(1), 64–72. <https://doi.org/10.1177/1073191103260945>
- Fluck, A. E., & Fox, A. (2011). Engaging training simulations for socially demanding roles. *Proceedings of ASCILITE - Australian Society for Computers in Learning in Tertiary Education Annual Conference 2011*, 398–406. <https://www.learntechlib.org/p/43577>
- Foley, J. A., & McAllister, G. (2005). Making It Real: Sim-School a Backdrop for Contextualizing Teacher Preparation. *AACE Journal*, 13(2), 159–177. <https://www.learntechlib.org/p/5955>
- Fortune, J. C., Cooper, J. M., & Allen, D. W. (1967). The Stanford Summer Micro-Teaching Clinic, 1965. *Journal of Teacher Education*, 18(4), 389–393. <https://doi.org/10.1177/002248716701800402>
- Frick, T. W., & Galindo, E. (2008). SimEd-Math: Modeling Differentiated Instruction in Mathematics. <http://webgrok.com/dossier/papers/NSF%2006%20535%20SimEd%20Math%20Combined.pdf>
- Galarneau, L., & Zibit, M. (2007). Online Games for 21st Century Skills. In D. Gibson, C. Aldrich & M. Prensky (Hrsg.), *Games and Simulation in Online Learning: Research and Development Frameworks* (S. 1–20). Information Science Publishing.
- Garland, K. M. V., Holden, K., & Garland, D. P. (2015). Individualized Clinical Coaching in the TLE TeachLivE Lab. *Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children*, 39(1), 47–59. <https://doi.org/10.1177/0888406415600769>
- Garland, K. M. V., Vasquez, E., & Pearl, C. (2012). Efficacy of Individualized Clinical Coaching in a Virtual Reality Classroom for Increasing

- Teachers' Fidelity of Implementation of Discrete Trial Teaching. *Education and Training in Autism and Developmental Disabilities*, 47(4), 502–515. <http://www.jstor.org/stable/23879642>
- Gibbs, G., Cowan, J., Sharpe, R., & Alexander, W. (2011). *Lecturing 1. Student attention over an hour*. https://epigeum.com/downloads/uct_accessible/uk/01_lecturing1/html/course_files/2_30.html
- Gibson, D. (2007). simSchool and the Conceptual Assessment Framework. In D. Gibson, C. Aldrich & M. Prensky (Hrsg.), *Games and Simulations in Online Learning* (S. 308–322). IGI Global. <https://doi.org/10.4018/978-1-59904-304-3.ch015>
- Gibson, D. (2009a). Designing a Computational Model of Learning. In R. E. Ferdig (Hrsg.), *Handbook of Research on Effective Electronic Gaming in Education* (S. 671–701). IGI Global. <https://doi.org/10.4018/978-1-59904-808-6.ch039>
- Gibson, D. (2009b). Modeling Classroom Behaviors in Software Agents. In D. Gibson & Y. Baek (Hrsg.), *Digital Simulations for Improving Education: Learning Through Artificial Teaching Environments* (S. 119–156). IGI Global / Information Science Reference.
- Gibson, D. (2011). *Modeling Emotions in Simulated Learning*. Lorentz Center International Center for workshops in the Sciences. https://www.researchgate.net/publication/267557376_Modeling_Emotions_in_Simulated_Learning
- Gibson, D. (2012a). *simSchool: Simulation-based teacher education*. <https://web.archive.org/web/20210108063331/http://www2.ed.gov/documents/college-completion/simulation-based-teacher-education.doc>
- Gibson, D. (2012b). Teacher Education with simSchool. In C. D. Maddux & D. Gibson (Hrsg.), *Research Highlights in Technology and Teacher Education 2012* (S. 37–44). Society for Information Technology & Teacher Education.
- Gibson, D. (2013). Psychometric considerations for digital assessments. *10th World Conference on Computers in Education*. http://edu.mat.umk.pl/archiv/wcce2013/publications/v1/V1.27_008-Gibson-fullR-FPR.pdf
- Gibson, D. (2014a). Affective Processes as Network Hubs. In T. Bosse, J. Broekens, J. Dias & J. van der Zwaan (Hrsg.), *Emotion Modeling: Towards Pragmatic Computational Models of Affective Processes* (S. 148–166). Springer International Publishing. https://doi.org/10.1007/978-3-319-12973-0_9
- Gibson, D. (2014b, 13. Mai). *Immersive Technologies, Games and Simulations - David Gibson: Videographisch aufgezeichneter Vortrag von*

- Curtin University's Second Wednesday Event on Immersion, Technologies, Games and Simulations.* <https://www.youtube.com/watch?v=Ia2RkwpypPM>
- Gibson, D., & Halverson, B. (2004). simSchool - Preparing Tomorrow's Teachers to Improve Student Learning. In R. Ferdig, C. Crawford, R. Carlsen, N. Davis, J. Price, R. Weber & D. A. Willis (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2004* (S. 3318–3321). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/13374>
- Gibson, D., & Jakl, P. (2013). *Data challenges of leveraging a simulation to assess learning* [International Association for Development of the Information Society, Paper presented at the International Association for Development of the Information Society (IADIS) International Conference on Cognition and Exploratory Learning in the Digital Age (CELDA), Fort Worth, TX, October 22–24.]. http://www.curveshift.com/images/Gibson_Jakl_data_challenges.pdf
- Gibson, D., Kruse, S., Knezek, G., Tyler-Wood, T., Christensen, R., & Hopper, S. B. (2013). simSchool's Progress: The SITE simSchool Modules Project. In R. McBride & M. Searson (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2013* (S. 1315–1323). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/48306>
- Girod, G. R., Girod, M., & Denton, J. (2007). Lessons Learned Modeling "Connecting Teaching and Learning". In D. Gibson, C. Aldrich & M. Prensky (Hrsg.), *Games and Simulations in Online Learning: Research and Development Frameworks* (S. 206–222). Information Science Publishing.
- Girod, M. (2009). Exploring Teacher Problem Solving Using Simulation. In D. Gibson & Y. Baek (Hrsg.), *Digital Simulations for Improving Education: Learning Through Artificial Teaching Environments* (S. 208–226). IGI Global / Information Science Reference.
- Girod, M., & Girod, G. R. (2006). Exploring the Efficacy of the Cook School District Simulation. *Journal of Teacher Education*, 57(5), 481–497. <https://doi.org/10.1177/0022487106293742>
- Gisbert, M., Esteve-González, V., & Cela-Ranilla, J. M. (2013). Constructing identity through virtual worlds. *Proceedings of EdMedia: World Conference on Educational Media and Technology 2013*, 1991–1998.
- Glocker, L., Breitenbach, S., Hansen, M., Mendzheritskaya, J., & Vö, M. L.-H. (2023). Entwicklung und Einsatz von VR-Lernszenarien für den Lehrkompetenzaufbau. In M. Meier, G. Greefrath, M. Ham-

- mann, R. Wodzinski & K. Ziepprecht (Hrsg.), *Lehr-Lern-Labore und Digitalisierung* (S. 211–224). Springer VS. https://doi.org/https://doi.org/10.1007/978-3-658-40109-2_22
- Göbel, K. (2003). Critical Incidents – aus schwierigen Situationen lernen. Vortrag im Rahmen der Fachtagung Lernnetzwerk Bürgerkompetenz, 17./18. Dezember 2003 in Bad Honnef. <https://www.dipf.de/de/forschung/aktuelle-projekte/pdf/biqua/critical-incidents-aus-schwierigen-situationen-lernen>
- Goldenberg, D., Andrusyszyn, M. A., & Iwasiw, C. (2005). The effect of classroom simulation on nursing students' self-efficacy related to health teaching. *J Nurs Educ*, 44(7), 310–314.
- Goldhammer, F., & Moosbrugger, H. (2006). Aufmerksamkeit. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 16–33). Springer.
- Götz, T., & Frenzel, A. C. (2006). Phänomenologie schulischer Langeweile. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 38(4), 149–153. <https://doi.org/10.1026/0049-8637.38.4.149>
- Götz, T., Frenzel, A. C., & Haag, L. (2006). Ursachen von Langeweile im Unterricht. *Empirische Pädagogik*, 20(2), 113–134.
- Götz, T., Frenzel, A. C., & Pekrun, R. (2007). Regulation von Langeweile im Unterricht. Was Schülerinnen und Schüler bei der "Windstille der Seele"(nicht) tun. *Unterrichtswissenschaft*, 35(4), 312–333. http://www.pedocs.de/volltexte/2012/5499;%20http://nbn-resolving.de/urn:nbn:de:0111-opus-54998;%20http://www.digizeitschriften.de/dms/img/?PPN=PPN513613439_0035&DMDID=DMDLOG_0037
- Gregory, S., Dalgarno, B., Crisp, G., Reiners, T., Masters, Y., Dreher, H., & Knox, V. (2012). *VirtualPREX: Innovative assessment using a 3D virtual world with pre-service teachers* (Final Report). University of New England (Lead Institution), u.a. https://ltr.edu.au/resources/PP10_1775_Gregory_report_2013.pdf
- Gundel, E., Piro, J. S., Straub, C., & Smith, K. (2019). Self-Efficacy in Mixed Reality Simulations: Implications for Preservice Teacher Education. *The Teacher Educator*, 54(3), 244–269. <https://doi.org/10.1080/08878730.2019.1591560>
- Hanif, S. (2016). *Distal and proximal indicators of teacher effectiveness as predictors of students' mathematics competence in Germany (Grade 5 –7)* (Dissertationsschrift). University of Bamberg Press, Universitätsbibliothek Bamberg.
- Hattie, J. (2009). *Visible learning*. Routledge.
- Hayes, A. T. (2015). *The Experience of Physical and Social Presence in a Virtual Learning Environment as Impacted by the Affordance of Mo-*

- vement Enabled by Motion Tracking* (Dissertationsschrift). University of Central Florida. <http://stars.library.ucf.edu/etd/5014>
- Hayes, A. T., Hardin, S. E., & Hughes, C. E. (2013). Perceived Presence's Role on Learning Outcomes in a Mixed Reality Classroom of Simulated Students. *VAMR 2013: Virtual, Augmented and Mixed Reality. Systems and Applications*, 142–151. https://doi.org/10.1007/978-3-642-39420-1_16
- Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight Simulator Training Effectiveness: A Meta-Analysis. *Military Psychology*, 4(2), 63–74. https://doi.org/10.1207/s15327876mp0402_1
- HCI Group Würzburg. (2017, 1. August). *Breaking Bad Behaviors 4.x Overview (with use in Seminar)*. <https://youtu.be/hmXeocHy8cQ>
- Heinich, R., Molenda, M., & Russell, J. D. (1992). *Instructional Media and the New Technologies of Instruction*. Macmillan Coll Div.
- Helmke, A. (2015). *Unterrichtsqualität und Lehrerprofessionalität : Diagnose, Evaluation und Verbesserung des Unterrichts* (6. Edition). http://scans.hebis.de/HEBCGI/show.pl?36580858_kla.pdf
- Helmke, A., Schneider, W., & Weinert, F. E. (1986). Quality of instruction and classroom learning outcomes: The German contribution to the IEA classroom environment study. *Teaching and Teacher Education*, 2(1), 1–18. [https://doi.org/10.1016/0742-051x\(86\)90002-8](https://doi.org/10.1016/0742-051x(86)90002-8)
- Helsgaun, K. (2000). Discrete Event Simulation in Java. *Datalogiske Skrifter (Writings on Computer Science)*, 89, 1–62. <http://akira.ruc.dk/~keld/research/JAVASIMULATION/JAVASIMULATION-1.0/docs/Report.pdf>
- Herbst, P., Chazan, D., Chen, C.-L., Chieu, V.-M., & Weiss, M. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM*, 43(1), 91–103. <https://doi.org/10.1007/s11858-010-0290-5>
- Hershey, G. L., Shepard, L. V., & Krumboltz, J. D. (1965). Effectiveness of Classroom Observation and Simulated Teaching in an Introductory Educational Psychology Course. *The Journal of Educational Research*, 58(5), 233–236. <http://www.jstor.org/stable/27531560>
- Herz, J. (2002). Gaming the system: What higher education can learn from multiplayer online worlds. In M. Devlin, R. Larson & J. Meyerson (Hrsg.), *The internet and the university: Forum 2001* (S. 169–191). Educause. <http://www.educause.edu/ir/library/pdf/ffpiu019.pdf>
- Hetmanek, A., & van Gog, T. (2017). Förderung von diagnostischer Kompetenz: Potential von Ansätzen aus der medizinischen Ausbildung. In A. Südkamp & A.-K. Praetorius (Hrsg.), *Diagnostische Kompetenz*

- von Lehrkräften. *Theoretische und methodische Weiterentwicklungen* (S. 209–215). Waxmann.
- Hettler, L., Gibson, D., Christensen, R., & Zibit, M. (2008, 26. März). *simMentoring. Guiding Development from Virtual to Real Teaching!* https://web.archive.org/web/20130327031438/http://iittl.unt.edu/IITTL/fipse/simMentoring_web/simHandbook.pdf
- Hixon, E., & So, H.-J. (2009). Technology's Role in Field Experiences for Preservice Teacher Training. *Educational Technology & Society*, 12(4), 294–304.
- Höntzsch, S., Katzky, U., Bredl, K., Kappe, F., & Krause, D. (2011). Simulationen und simulierte Welten. Lernen in immersiven Lernumgebungen. In M. Ebner & S. Schön (Hrsg.), *Lehrbuch für Lernen und Lehren mit Technologien (L3T)*. (S. 1–7). BIMS e.V. <http://l3t.eu/homepage/das-buch/ebook/kapitel/o/id/32/name/simulationen-und-simulierte-welten>
- Hopper, S. B. (2014). *The Relationship of Personality Traits to Teacher Candidate Perceptions of Teaching Confidence and Teaching Experience in a Simulated Classroom Environment* (Dissertationsschrift). University of North Texas. <https://digital.library.unt.edu/ark:/67531/metadc500089/>
- Hopper, S. B. (2018). The Heuristic Sandbox: Developing Teacher Know-How Through Play in simSchool. *Journal of Interactive Learning Research*, 29(1), 77–112. <https://files.eric.ed.gov/fulltext/EJ1170969.pdf>
- Hopper, S. B., Knezek, G., & Christensen, R. (2013). Assessing Alignment of Pedagogical Experience and Confidence in a Simulated Classroom Environment. In R. McBride & M. Searson (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2013* (S. 2870–2876). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/48527>
- Horz, H., & Schulze-Vorberg, L. (2017). Digitalisierung in der Hochschullehre (Konrad Adenauer Stiftung, Hrsg.). *Analysen & Argumente: Digitale Gesellschaft*, 282, 1–12. https://www.kas.de/c/document_library/get_file?uuid=51d9958b-ae68-e2f4-d98b-7508d9bf4123&groupId=252038
- Horz, H., & Ulrich, I. (2013). Strategische Entwicklung neuer Lerndesigns. *Wirtschaft und Beruf*, 65, 9–14. https://www.researchgate.net/publication/332028562_Strategische_Entwicklung_neuer_Lerndesigns
- Howard, P. J., & Howard, J. M. (2000). *The Owner's Manual for Personality at Work: How the Big Five Personality Traits Affect Your Per-*

- formance, Communication, Teamwork, Leadership, and Sales*. Bard Press.
- Hsu, C.-M., Yeh, Y.-C., & Yen, J. (2009). Development of design criteria and evaluation scale for web-based learning platforms. *International Journal of Industrial Ergonomics*, 39(1), 90–95. <https://doi.org/10.1016/j.ergon.2008.08.006>
- Humpert, W., & Dann, H.-D. (1984). *Das Beobachtungssystem BAVIS: Ein handlungstheoretisch orientiertes Beobachtungsverfahren zur Analyse von aggressionsbezogenen Interaktionen im Schulunterricht*. Universität Konstanz.
- Humphreys, T. (1995). *A Different Kind of Teacher*. Cassell.
- Indra. (2010). *SIMAULA - Tomorrow's Teachers training*. <https://www.indracompany.com/en/indra/simaula-tomorrows-teachers-training>
- International ICT Literacy Panel. (2002). *Digital Transformation: A Framework for ICT Literacy*. Educational Testing Service.
- IsaBella, C., Vézina, N., & Fournier, H. (2006). Un environnement 3D qui favorise le sentiment d'appartenance en situation de formation à distance. *Canadian Journal of Learning and Technology / La revue canadienne de l'apprentissage et de la technologie*, 32(2). <https://doi.org/10.21432/T2K309>
- Issenberg, S. B., Mcgaghie, W. C., Petrusa, E. R., Gordon, D. L., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Medical Teacher*, 27(1), 10–28. <https://doi.org/10.1080/01421590500046924>
- Jewett, M. E., & Kronauer, R. E. (1999). Interactive Mathematical Models of Subjective Alertness and Cognitive Throughput in Humans. *Journal of Biological Rhythms*, 14(6), 588–597. <https://doi.org/10.1177/074873099129000920>
- Kaiser, J., & Südkamp, A. (2017). Get the picture - Eine Response-Surface-Analyse des Zusammenhangs zwischen Urteilssicherheit und Urteilsakkuratheit. In J. Retelsdorf, F. Zimmermann, A. Südkamp & O. Köller (Hrsg.), *Im Blickpunkt pädagogisch-psychologischer Forschung. Selbstbezogene Kognitionen, sprachliche Kompetenzen und Professionalisierung von Lehrkräften. Festschrift für Jens Möller* (S. 255–270). Waxmann.
- Kallonis, P., & Sampson, D. G. (2010). Implementing a 3D Virtual Classroom Simulation for Teachers' Continuing Professional Development. *Workshop Proceedings of the 18th International Conference on Computers in Education*, 36–44.
- Katz, Y. J. (1999). Kindergarten Teacher Training Through Virtual Reality: Three-Dimensional Simulation Methodology. *Educational Me-*

- dia International*, 36(2), 151–156. <https://doi.org/10.1080/0952398990360211>
- Kaufman, D., & Ireland, A. (2016). Enhancing Teacher Education with Simulations. *TechTrends*, 60(3), 260–267. <https://doi.org/10.1007/s11528-016-0049-0>
- Keller, J., & Kopp, T. (1987). An application of the ARCS model of motivational design. In C. M. Reigeluth (Hrsg.), *Instructional theories in action. Lessons illustrating selected theories and models* (S. 289–320). Hillsdale.
- Keller, J. M. (1983). Motivational design of instruction. In C. Reigeluth (Hrsg.), *Instructional-design theories and models: An overview of their current status* (S. 383–436). Lawrence Erlbaum Associates.
- Kelley, M. J., & Wenzel, T. (2018). A parent-teacher reading conference project: Using a virtual environment (TeachLivE™) to improve elementary pre-service teachers' conferencing skills. *The Reading Professor*, 41(1), 48–59.
- Kemp, J. W., Livingstone, D., & Bloomfield, P. R. (2009). SLOODLE: Connecting VLE tools with emergent teaching practice in Second Life. *British Journal of Educational Technology*, 40(3), 551–555. <https://doi.org/10.1111/j.1467-8535.2009.00938.x>
- Kersh, B. Y. (1962). The Classroom Simulator. *Journal of Teacher Education*, 13(1), 109–110. <https://doi.org/10.1177/002248716201300128>
- Kersh, B. Y. (1963, 30. Juni). *Classroom Simulation: A New Dimension in Teacher Education: Final report, National defense education act of 1958* (Techn. Ber.) [Project number 886]. Oregon State System of Higher Education. Monmouth.
- Kervin, L., Cambourne, B., Turbill, J., Ferry, B., Hedberg, J., Jonassen, D., & Puglisi, S. (2004). From Classroom Reality to Virtual Classroom: The Role of Teacher-Created Scripts in the Development of Classroom Simulation Technology. *Australian Association of Research in Education (AARE) Conference Paper*. <https://www.aare.edu.au/data/publications/2004/ker04657.pdf>
- Kervin, L., Ferry, B., & Carrington, L. (2006). ClassSim: Preparing Tomorrow's Teachers for Classroom Reality. In C. Crawford, C. Carlsen, K. McFerrin, J. Price, R. Weber & D. Willis (Hrsg.), *Society for Information Technology & Teacher Education Conference* (S. 3204–3211). Association for the Advancement of Computing in Education.
- Kim, B., & Cheong, D. (2009). simClass: Simulate Your Class Before You Teach. In D. Gibson & Y. Baek (Hrsg.), *Digital Simulations in Teaching and Learning: Learning Through Artificial Teaching Environments* (S. 289–307). IGI Global / Information Science Reference.

- Kinder, D., & Carnine, D. (1991). Direct instruction: What it is and what it is becoming. *Journal of Behavioral Education, 1*(2), 193–213. <https://doi.org/10.1007/bf00957004>
- Klinzing, H. G. (2002). Wie effektiv ist Microteaching? Ein Überblick über fünfunddreißig Jahre Forschung. *Zeitschrift für Pädagogik, 48*(2), 194–214. <https://doi.org/10.25656/01:3829>
- Knezek, G., Hopper, S. B., Christensen, R., Tyler-Wood, T., & Gibson, D. C. (2015). Assessing Pedagogical Balance in a Simulated Classroom Environment. *Journal of Digital Learning in Teacher Education, 31*(4), 148–159. <https://doi.org/10.1080/21532974.2015.1055011>
- Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary issues in technology and teacher education, 9*(1), 60–70.
- Köknar, C. (2015). *Modelling Student Behaviours in a Virtual cClass-room with Incorporation of Social Learning Theory into Belief-Desire-Intention Model* (Dissertationsschrift). Middle East Technical University. Department of Modeling und Simulation.
- Kounin, J. S. (1976). *Techniken der Klassenführung* (Reprint 2006). Waxmann.
- Kruse, S., & Gibson, D. (2011). *Next Generation Learning Challenge*. <http://er.educause.edu/articles/2011/12/next-generation-learning-challenge-simulating-teaching>
- Kunter, M., & Trautwein, U. (2013). *Psychologie des Unterrichts*. Ferdinand Schöningh.
- Kwon, C. (2019). Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR technologies. *Virtual Reality, 23*(1), 101–118. <https://doi.org/10.1007/s10055-018-0364-1>
- Latoschik, M. E., Lugin, J.-L., Habel, M., Roth, D., Seufert, C., & Grafe, S. (2016). Breaking Bad Behavior: Immersive Training of Class Room Management. *Proceedings of the 22Nd ACM Conference on Virtual Reality Software and Technology, 317–318*.
- Ledger, S., & Fischetti, J. (2020). Micro-teaching 2.0: Technology as the classroom. *Australasian Journal of Educational Technology, 36*(1), 37–54. <https://doi.org/10.14742/ajet.4561>
- Lee, K., & Ashton, M. C. (2012). *The H factor of personality*. Wilfrid Laurier University Press.
- Lipowsky, F., Rakoczy, K., Pauli, C., Reusser, K., & Klieme, E. (2007). Gleicher Unterricht – gleiche Chancen für alle? Die Verteilung von Schülerbeiträgen im Klassenunterricht. *Unterrichtswissenschaft, 35*(2), 125–147. <https://doi.org/10.25656/01:5489>

- Lorenzi, E. C. (2013). *Are You Paying Attention? Understanding Student Attention Allocation using Bayesian Models* (Master's Thesis). Carnegie Mellon University, Department of Statistics. <https://doi.org/10.1184/R1/6683960.v1>
- Lowdermilk, J., Martinez, D., Pecina, J., Beccera, L., & Lowdermilk, C. (2012). Behavior Breakthroughs™: Future Teachers Reflect on a Focused Game Designed to Teach ABA Techniques. *TechTrends*, 56(3), 29–35. <https://doi.org/https://doi.org/10.1007/s11528-012-0573-5>
- Lugrin, J.-L., Charles, F., Habel, M., Dudaczy, H., Oberdörfer, S., Matthews, J., Porteous, J., Wittmann, A., Seufert, C., Grafe, S., & Latoschik, M. E. (2018). Benchmark Framework for Virtual Students' Behaviours. *Proceedings of the 17th Conference on Autonomous Agents and Multi Agent Systems*, 2236–2238. <https://downloads.hci.informatik.uni-wuerzburg.de/2018-acm-aamas-lugrin-preprint.pdf>
- Lugrin, J.-L., Latoschik, M. E., Habel, M., Roth, D., Seufert, C., & Grafe, S. (2016). Breaking Bad Behaviours: A New Tool for Learning Classroom Management using Virtual Reality. *Frontiers in ICT*, 3, 1–21. <https://doi.org/https://doi.org/10.3389/fict.2016.00026>
- Lugrin, J.-L., Oberdörfer, S., Latoschik, M. E., Wittmann, A., Seufert, C., & Grafe, S. (2018). VR-Assisted vs Video-Assisted Teacher Training. *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 625–626. <https://downloads.hci.informatik.uni-wuerzburg.de/2018-ieeevr-lugrin-vr-teacher-training-poster-preprint.pdf>
- Lund, A. (2001). Measuring Usability with the USE Questionnaire. *Usability and User Experience Newsletter of the STC Usability SIG*, 8. https://www.researchgate.net/publication/230786746_Measuring_Usability_with_the_USE_Questionnaire
- MacLellan, C. J. (2017). *Computational Models of Human Learning: Applications for Tutor Development, Behavior Prediction, and Theory Testing* (Dissertationsschrift). Carnegie Mellon University, Human-Computer Interaction Institute, School of Computer Science. https://kilthub.cmu.edu/articles/Computational_Models_of_Human_Learning_Applications_for_Tutor_Development_Behavior_Prediction_and_Theory_Testing/6715271/files/12248333.pdf
- Manburg, J., Moore, R., Griffin, D. K., & Seperson, M. (2017). Building Reflective Practice through an Online Diversity Simulation in an Undergraduate Teacher Education Program. *Contemporary Issues in Technology and Teacher Education*, 17(1), 128–153.
- Mayring, P., & Fenzl, T. (2019). Qualitative Inhaltsanalyse. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung*

- (S. 633–648). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-21308-4_42
- McBrien, J. L., Cheng, R., & Jones, P. (2009). Virtual Spaces: Employing a Synchronous Online Classroom to Facilitate Student Engagement in Online Learning. *The International Review of Research in Open and Distributed Learning*, 10(3). <https://doi.org/10.19173/irrodl.v10i3.605>
- McCrae, R. R., & Costa Jr., P. T. (1996). Toward a new generation of personality theories: The Five-Factor Model of Personality. In *The five-factor model of personality: Theoretical perspectives* (S. 51–87). Guilford Press.
- McPherson, R., Tyler-Wood, T., Ellison, A. M., & Peak, P. (2011). Using a Computerized Classroom Simulation to Prepare Pre-Service Teachers. *Journal of Technology and Teacher Education*, 19(1), 93–110. <https://www.learntechlib.org/p/31438>
- Megawati, F., Mukminatien, N., Anugerahwati, M., Indrayani, N., & Unsiyah, F. (2020). Critical Incidents: Exploring EFL Prospective Teachers' Teaching Experiences. *European Journal of Educational Research*, 9(3), 943–954. <https://doi.org/10.12973/eu-jer.9.3.943>
- Mehan, H. (1979). *Learning Lessons. Social Organization in the Classroom*. Harvard University Press.
- Mengel, S. (2011). *Didaktische Szenarien für Live-E-Learning im virtuellen Klassenzimmer*. https://lucylection.files.wordpress.com/2012/05/vc_didaktik_mengel_201111.pdf
- Merchant, Z., Goetz, E. T., Keeney-Kennicutt, W., Kwok, O.-m., Cifuentes, L., & Davis, T. J. (2012). The learner characteristics, features of desktop 3D virtual reality environments, and college chemistry instruction: A structural equation modeling analysis. *Computers & Education*, 59(2), 551–568. <https://doi.org/https://doi.org/10.1016/j.compedu.2012.02.004>
- Mikropoulos, T. A. (2006). Presence: a unique characteristic in educational virtual environments. *Virtual Reality*, 10(3-4), 197–206. <https://doi.org/10.1007/s10055-006-0039-1>
- Mohanam, R., Stringfellow, C., & Gupta, D. (2017). An emotionally intelligent tutoring system. *2017 Computing Conference*, 1099–1107. <https://doi.org/10.1109/SAI.2017.8252228>
- Monsour, D. (2012, 10. Mai). *simSchool May 2012 News*. <http://www.aace.org/review/simschool-may-2012-news>
- Moosbrugger, H., & Goldhammer, F. (2006). Aufmerksamkeits- und Konzentrationsdiagnostik. In K. Schweizer (Hrsg.), *Leistung und Leistungsdiagnostik* (S. 83–102). Springer.

- Moreno, R., & Mayer, R. E. (2004). Personalized Messages That Promote Science Learning in Virtual Environments. *Journal of Educational Psychology, 96*(1), 165–173. <https://doi.org/10.1037/0022-0663.96.1.165>
- Mukhtar, M. A., Hasim, Z., & Yunus, M. M. (2018). The efficacy of simulated teaching in preparing pre-service teachers for practicum. *Journal of Nusantara Studies (JONUS), 3*(1), 64–74. <https://doi.org/10.24200/jonus.vol3iss1pp64-74>
- Nagendran, A., Pillat, R., Kavanaugh, A., Welch, G., & Hughes, C. (2013). AMITIES: Avatar-Mediated Interactive Training and Individualized Experience System. *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology - VRST '13*. <https://doi.org/10.1145/2503713.2503731>
- Naumann, J., Richter, T., & Groeben, N. (2001). Validierung des INCOBI anhand eines Vergleichs von Anwendungsexperten und Anwendungsnovizen. *Zeitschrift für Pädagogische Psychologie, 15*(3/4), 219–232. <https://doi.org/10.1024//1010-0652.15.34.219>
- o.A. (2002, 11. November). *Virtual school becomes reality for teachers*. <https://www.theguardian.com/education/2002/nov/11/schools.uk>
- o.A. (2019). *Sinifta Homepage*. <https://www.sinifta.com/>
- o.A. (2023). *The teaching screening process at Murdoch*. <https://www.murdoch.edu.au/study/courses/study-areas/teaching/simlab-screening>
- Oberdick, J. (2016, 9. August). *Penn State education professor developing artificial intelligence classroom for pre-service teacher practice*. <https://web.archive.org/web/20170702234406/http://tlt.psu.edu/2016/08/09/penn-state-education-professor-developing-artificial-intelligence-classroom-for-pre-service-teacher-practice/>
- O'Callaghan, C., & Piro, J. S. (2016). Shifting Towards Practice-Based Teacher Preparation: Using Mixed Reality Simulations to Impact Pre-service Teachers' Situated Cognition. In T. Bousfield, L. A. Dieker, C. Hughes & M. Hynes (Hrsg.), *Proceedings of the 4th Annual TeachLivE Conference, June 1–3, 2016* (S. 36–50).
- Ortony, A., Clore, G. L., & Collins, A. (1988, Juli). *The Cognitive Structure of Emotions*. Cambridge University Press. <https://doi.org/10.1017/cbo9780511571299>
- Overbaugh, R. C. (1995). The Efficacy of Interactive Video for Teaching Basic Classroom Management Skills to Pre-Service Teachers. *Computers in Human Behavior, 11*(3), 511–527.
- Page, R. L. (2000). Brief History of Flight Simulation. *SimTecT 2000 Proceedings*, 11–17.

- Paloff, R. M., & Pratt, K. (2013). *Lessons from the Virtual Classroom : The Realities of Online Teaching* (2. Edition). Wiley.
- Park, K.-M., Ku, J., Choi, S.-H., Jang, H.-J., Park, J.-Y., Kim, S. I., & Kim, J.-J. (2011). A virtual reality application in role-plays of social skills training for schizophrenia: A randomized, controlled trial. *Psychiatry Research*, *189*(2), 166–172. <https://doi.org/10.1016/j.psychres.2011.04.003>
- Payne, R. K. (2006). *Working with Students: Discipline Strategies for the Classroom*; aha! Process, Inc.
- Persky, S., Kaphingst, K. A., McCall, C., Lachance, C., Beall, A. C., & Blascovich, J. (2009). Presence Relates to Distinct Outcomes in Two Virtual Environments Employing Different Learning Modalities. *CyberPsychology & Behavior*, *12*(3), 263–268. <https://doi.org/10.1089/cpb.2008.0262>
- Pfeifer, W. (1999). simulieren. In *Etymologisches Wörterbuch des Deutschen*. Dt. Taschenbuch-Verlag.
- Pfitzner, M., & Schoppek, W. (2000). Gemeinsamkeiten und Diskrepanzen in der Bewertung von Unterrichtsstörungen durch Lehrer und Schüler – eine empirische Untersuchung. *Unterrichtswissenschaft*, *28*(4), 350–378.
- Plutchik, R. (1997). The circumplex as a general model of the structure of emotions and personality. In R. Plutchik & H. R. Conte (Hrsg.), *Circumplex models of personality and emotions*. (S. 17–45). American Psychological Association. <https://doi.org/10.1037/10261-001>
- Praetorius, S., Al-Kabbani, D., Bohndick, C., Hilkenmeier, J., König, S., Müsche, H., Sommer, S., & Klingsieck, K. (2016). Spielend Lehrer/in werden: problembasiertes Lernen mit virtuellen Schülerinnen/Schülern. *Zeitschrift für Hochschulentwicklung*. <https://doi.org/10.3217/zfhe-11-03/14>
- Pritchard, J. K. (2013). SwRI's Behavior Breakthroughs™. *Behavior Analysis in Practice*, *6*(2), 84–85. <https://doi.org/10.1007/BF03391810>
- Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science*, *23*(5), 405–431. <https://doi.org/10.1007/BF00896880>
- Ravipati, S. (2016, 15. August). *Penn State Explores AI in Teaching*. <https://campustechnology.com/articles/2016/08/15/penn-state-explores-ai-in-teaching.aspx>
- Renkl, A. (1996). Träges Wissen: Wenn Erlerntes nicht genutzt wird. *Psychologische Rundschau*, *47*, 78–92.
- Richards, D., & Taylor, M. (2015). A Comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment

- to find the right representation involving the Marginal Value Theorem. *Computers & Education*, 86, 157–171. <https://doi.org/https://doi.org/10.1016/j.compedu.2015.03.009>
- Richter, T., Naumann, J., & Groeben, N. (2000). Attitudes toward the computer: construct validation of an instrument with scales differentiated by content. *Computers in Human Behavior*, 16(5), 473–491. [https://doi.org/10.1016/s0747-5632\(00\)00025-x](https://doi.org/10.1016/s0747-5632(00)00025-x)
- Richter, T., Naumann, J., & Horz, H. (2010). Eine revidierte Fassung des Inventars zur Computerbildung (INCOBI-R). *Zeitschrift für Pädagogische Psychologie*, 24(1), 23–37. <https://doi.org/https://doi.org/10.1024/1010-0652/a000002>
- Rolfe, J. M., & Staples, J. (1986). *Flight Simulation*. Cambridge University Press.
- Rubin, J. A., & Crockett, R. S. (2015, 28. Dezember). *U.S. Patent Nr. 20160139666 A1: "Whole-body human-computer interface"* (20160139666 A1).
- Sacher, W. (1995). *Meldungen und Aufrufe im Unterrichtsgespräch. Theoretische Grundlagen, Forschungsergebnisse, Trainingselemente und Diagnoseverfahren*. Wissner.
- Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., & Müller, K.-R., Hrsg. *Explainable AI: Interpreting, explaining and visualizing deep learning* (1. Edition). 1. Edition. Lecture notes in computer science. Cham, Switzerland: Springer Nature, 2019, August.
- Sauvé, L., Renaud, L., Kaufman, D., & Marquis, J.-S. (2007). Distinguishing between games and simulations: A systematic review. *Journal of Educational Technology & Society*, 10(3), 247–256. <https://www.jstor.org/stable/jeductechsoci.10.3.247>
- Sawchuk, S. (2011). Simulations Helping Novices Hone Skills. *Education Week*. <https://www.edweek.org/teaching-learning/simulations-helping-novices-hone-skills/2011/01>
- Sawyer, B., & Rejeski, D. (2002). *Serious Games: Improving Public Policy Through Game-based Learning and Simulation*. (Whitepaper). <https://www.wilsoncenter.org/publication/executive-summary-serious-games-improving-public-policy-through-game-based-learning-and>
- Scherzinger, M., Wettstein, A., & Wyler, S. (2018). Aggressive und nicht aggressive Unterrichtsstörungen durch Schülerinnen und Schüler sowie durch Klassen- und Fachlehrpersonen. Eine Videostudie. *Empirische Sonderpädagogik*, 10(4), 388–407. <https://doi.org/https://doi.org/10.25656/01:16782>
- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological Pedagogical Content Knowledge

- (TPACK). *Journal of Research on Technology in Education*, 42(2), 123–149. <https://doi.org/10.1080/15391523.2009.10782544>
- Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martínez, V. (2007). The Geographic Distribution of Big Five Personality Traits. *Journal of Cross-Cultural Psychology*, 38(2), 173–212. <https://doi.org/10.1177/0022022106297299>
- Seidel, T. (2006). The role of student characteristics in studying micro teaching–learning environments. *Learning Environments Research*, 9(3), 253–271. <https://doi.org/10.1007/s10984-006-9012-x>
- Seidel, T. (2009). Klassenführung. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (S. 135–148).
- Seidel, Tina. (2014). Angebots-Nutzungs-Modelle in der Unterrichtspsychologie. Integration von Struktur- und Prozessparadigma. *Zeitschrift für Pädagogik*, 60, 850–866. <https://doi.org/10.25656/01:14686>
- Serious Games Showcase and Challenge. (2018). *SimInClass by Simsoft Technologies*. <http://sgschallenge.com/2018siminclass/>
- SGSChallenge. (2018, 24. September). *SimInClass by Simsoft Technologies*. <https://youtu.be/LsZF1g7RY24>
- Shannon, R. E. (1998). Introduction to the art and science of simulation. *1998 Winter Simulation Conference Proceedings*, 1, 7–14. <https://doi.org/10.1109/WSC.1998.744892>
- Siestrup, C. G. (2010). *Unterrichtsstörungen aus der Sicht von Lehrenden und Lernenden. Ursachenzuschreibungen, emotionales Erleben und Konzepte zur Vermeidung*. Lang.
- simSchool Educator Training. (2019a, 11. März). *Overview of simSchool2*. <https://www.youtube.com/watch?v=477ihRCU04U>
- simSchool Educator Training. (2019b, 11. März). *simSchool Teacher Talkaloud*. <https://youtu.be/bRIDW-ELbTM>
- simSchool Educator Training. (2019c, 11. März). *simSchool Understanding the AI*. <https://youtu.be/oVfptG-6L-Q>
- simSchool Educator Training. (2019d, 14. März). *simSchool Teach Aloud*. https://youtu.be/_xVbgzeDiuU
- Simsoft Computer Technologies. (2018, 17. August). *Sınıfta – Öğretmen Eğitimi için Simülasyon*. <https://youtu.be/-eiFLkSa65M>
- simulare. (2018). In *Lagenscheidt Online-Wörterbuch*. <https://woerterbuch.langenscheidt.de>
- Sınıfta. (2018, 5. September). *SimInClass – Trailer*. <https://youtu.be/-FmwhD76o5s>
- Skrødal, S. (2010). *Virtual Classroom Simulation. Design and Trial in a Pre-service Teacher Education Program* (Dissertationsschrift). University of Adelaide, School of Education. <http://hdl.handle.net/2440/65251>

- Southwest Research Institute. (2011, 11. April). *SwRI-developed training program for dealing with behavioral problems is available for home computers*. <https://www.swri.org/press-release/training-program-behavioral-problems-home-computers>
- Sripada, C., Kessler, D., & Jonides, J. (2014). Methylphenidate Blocks Effort-Induced Depletion of Regulatory Control in Healthy Volunteers. *Psychological Science*, 25(6), 1227–1234. <https://doi.org/10.1177/0956797614526415>
- Stadt Köln. (2018, 23. März). *Neue Kölner Statistik. Thema: Bevölkerung im Jahr 2017*. https://www.stadt-koeln.de/mediaasset/content/pdf15/statistik-einwohner-und-haushalte/1084795_k%C3%B6lnerinnen_und_k%C3%B6lner_in_2017_ew_nks_1_2017.pdf
- Stavroulia, K. E. (2015, 9. Juli). Simulation-based Training for greek preservice teachers Developing a 3D Classroom Environment for professional development via practical experience. In L. Morris & C. Tsolakidis (Hrsg.), *International Conference on Information Communication Technologies in Education (ICICTE) 2015 Proceedings* (S. 390–401).
- Stavroulia, K. E. (2016). Emotional experiences in simulated classroom training environments (G. Anderson, Hrsg.). *The International Journal of Information and Learning Technology*, 33(3), 172–185. <https://doi.org/10.1108/IJILT-10-2015-0030>
- Stavroulia, K. E., Baka, E., Lanitis, A., & Magnenat-Thalmann, N. (2017). Virtual Reality-based Learning Environments in Teacher Training: New Opportunities and Challenges. *Proceedings of the 10th annual International Conference of Education, Research and Innovation (ICERI) 2017*, 3562–3571. <https://doi.org/10.21125/iceri.2017.0967>
- Stavroulia, K. E., Baka, E., Lanitis, A., & Magnenat-Thalmann, N. (2018). Designing a Virtual Environment for Teacher Training: Enhancing Presence and Empathy. *Proceedings of Computer Graphics International (CGI) 2018*, 273–282. <https://doi.org/10.1145/3208159.3208177>
- Stavroulia, K. E., Christofi, M., Baka, E., Michael-Grigoriou, D., Magnenat-Thalmann, N., & Lanitis, A. (2019). Assessing the emotional impact of virtual reality-based teacher training. *The International Journal of Information and Learning Technology*, 36(3), 192–2017. <https://doi.org/https://doi.org/10.1108/IJILT-11-2018-0127>
- Stavroulia, K. E., & Lanitis, A. (2018). A systematic virtual reality-based approach to support the professional development of teachers. *International Conference on Information Communication Technologies in Education (ICICTE) 2018 Proceedings*, 29–38. http://www.iciete.org/assets/1.4_stavroulia_lanitis.pdf

- Stavroulia, K. E., Makri-Botsari, E., Psycharis, S., & Kekkeris, G. (2015a). Using Simulations as a tool to enhance classroom management practice. *International Conference on Information Communication Technologies in Education (ICICTE) 2014 Proceedings*, 205–217. <http://www.icitte.org/Proceedings2014/Papers%202014/6.2%20Stavroulia.pdf>
- Stavroulia, K. E., Makri-Botsari, E., Psycharis, S., & Kekkeris, G. (2015b, 9. Juli). Emotional Experiences in Simulated Classroom Training Environments. In L. Morris & C. Tsolakidis (Hrsg.), *International Conference on Information Communication Technologies in Education (ICICTE) 2015 Proceedings* (S. 390–401). [http://icitte.org/ICICTE2015Proceedings\(Papers\)/11.1%20Stavroulia.pdf](http://icitte.org/ICICTE2015Proceedings(Papers)/11.1%20Stavroulia.pdf)
- Strang, H. R. (1996). The TD-Simulation: An Interactive Vehicle for Mapping Teaching Decisions. In B. Robin, J. D. Price, J. Willis & D. A. Willis (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education (SITE) International Conference 1996* (S. 616–619). Association for the Advancement of Computing in Education (AACE).
- Strang, H. R. (1997). The Use of Curry Teaching Simulations in Professional Training. *Computers in the Schools*, 13(3), 135. <https://www.learntechlib.org/p/84024>
- Strang, H. R., Landrum, M. S., & Lynch, K. A. (1989). Talking with the computer: A simulation for training basic teaching skills. *Teaching and Teacher Education*, 5(2), 143–153. [https://doi.org/http://dx.doi.org/10.1016/0742-051X\(89\)90012-7](https://doi.org/http://dx.doi.org/10.1016/0742-051X(89)90012-7)
- Strang, H. R., & Moore, S. D. (1994). Exploring Teaching Techniques via a Microcomputer Simulation. *Proceedings of Society for Information Technology & Teacher Education (SITE) International Conference 1996*, 620–622. <https://doi.org/10.1080/0962029940030103>
- Strang, H. R., Sullivan, A. K., & Yeh, Y.-C. (1996). Responding to Student Learning Styles Within a Visual Basic Simulation. In B. Robin, J. D. Price, J. Willis & D. A. Willis (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education (SITE) International Conference 1996* (S. 652–654). Association for the Advancement of Computing in Education (AACE).
- Straub, C., Dieker, L. A., Hynes, M., & Hughes, C. (2014). *Using virtual rehearsal in TLE TeachLivETM mixed reality classroom simulator to determine the effects on the performance of mathematics teachers* (Techn. Ber.). TeachLivE National Research Project. https://web.archive.org/web/20190716083250/http://teachlive.org/wp-content/uploads/2014/10/2014_GR_Technical_Report_10_20_FINAL.pdf

- Südkamp, A., & Möller, J. (2009). Referenzgruppeneffekte im Simulierten Klassenraum: Direkte und Indirekte Einschätzungen von Schülerleistungen. *Zeitschrift für Pädagogische Psychologie*, 23(3-4), 161–174. <https://doi.org/10.1024/1010-0652.23.34.161>
- Südkamp, A., Möller, J., & Pohlmann, B. (2008). Der Simulierte Klassenraum. *Zeitschrift für Pädagogische Psychologie*, 22(34), 261–276. <https://doi.org/10.1024/1010-0652.22.34.261>
- Suev, S. (2013, 11. Februar). *Simaula complete movie title*. <https://youtu.be/lnq7IwipW4c>
- Sullivan, A. K., Yeh, Y.-c., & Strang, H. R. (1998). Using a Computer Simulation To Explore Teacher Lesson Planning. In S. McNeil, J. D. Price, S. Boger-Mehall, B. Robin & J. Willis (Hrsg.), *Proceedings of Society for Information Technology & Teacher Education International Conference 1998* (S. 998–1001). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/p/47912>
- Sweeney, J., Milewski, A., & Amidon, J. (2018). On-Ramps to Professional Practice: Selecting and Implementing Digital Technologies for Virtual Field Experiences. *Contemporary Issues in Technology and Teacher Education*, 18(4), 670–691. <https://www.learntechlib.org/p/182990>
- Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/https://doi.org/10.1207/s15516709cog1202_4
- Sweller, J. (2011). Cognitive Load Theory. In *Psychology of Learning and Motivation* (S. 37–76). Elsevier. <https://doi.org/10.1016/b978-0-12-387691-1.00002-8>
- Tadayon, M., & Afhami, R. (2016). Doodling Effects on Junior High School Students' Learning. *International Journal of Art and Design Education*, 36(1), 118–125. <https://doi.org/10.1111/jade.12081>
- Tansey, P. (1970). Simulation Techniques in the Training of Teachers. *Simulation & Games*, 1(3), 281–303. <https://doi.org/10.1177/104687817000100303>
- Tenzer, F. (2022). *Statistiken zum Thema Gaming*. <https://de.statista.com/themen/1095/gaming/#topicOverview>
- Thomas, D. R., Becker, W. C., & Armstrong, M. (1968). Production and elimination of disruptive classroom behavior by systematically varying teacher's behavior. *Journal of Applied Behavior Analysis*, 1(1), 35–45. <https://doi.org/10.1901/jaba.1968.1-35>
- Thompson, L. M. A. (1896, 5. Oktober). *U.S. Patent Nr. US665763: "Pleasure-Railway"*. <https://patents.google.com/patent/US665763A/en?q=US665763>

- Tripp, D. (2011). *Critical Incidents in Teaching (Classic Edition) : Developing Professional Judgement*. Taylor & Francis Group.
- Turan, S. (2015). Simulation-based learning: how can be used to prepare teacher? *Educação, Ciência e Cultura*, 20(1), 27–43. <https://doi.org/10.18316/2236-6377.15.2>
- Twelker, P. A. (1967). Classroom Simulation and Teacher Preparation. *The School Review*, 75(2), 197–204. <http://www.jstor.org/stable/1084325>
- Tyler-Wood, T., Estes, M., Christensen, R., Knezek, G., & Gibson, D. (2015). SimSchool: An Opportunity for Using Serious Gaming for Training Teachers in Rural Areas. *Rural Special Education Quarterly*, 34(3), 17–20. <https://doi.org/10.1177/875687051503400304>
- Ur, S., & VanLehn, K. (1995). Steps: A simulated, tutorable physics student. *Journal of Artificial Intelligence in Education*, 6(4), 405–437. https://www.public.asu.edu/~kvanlehn/Stringent/PDF/95AIED_CC_KVL.pdf
- VanLehn, K. (1996). Cognitive Skill acquisition. *Annual Review of Psychology*, 47(1), 513–539. <https://doi.org/10.1146/annurev.psych.47.1.513>
- VanLehn, K., Ohlsson, S., & Nason, R. (1994). Applications of Simulated Students: An Exploration. *Journal of Artificial Intelligence in Education*, 5(2), 135–175. <http://dl.acm.org/citation.cfm?id=191871.191873>
- Vygotskij, L. S. (1987a). Das Problem der Altersstufen [Original publiziert ca. 1934] (R. Kossert, Übers.). In J. Lompscher (Hrsg.), *Lew Vygotskji: Ausgewählte Schriften* [Band II] (S. 53–90). Lehmanns Media.
- Vygotskij, L. S. (1987b). Unterricht und geistige Entwicklung im Schulalter (R. Kossert, Übers.). In J. Lompscher (Hrsg.), *Lew Vygotskji: Ausgewählte Schriften* [Band II] (S. 287–306). Lehmanns Media. Original publiziert ca. 1934.
- Wang, F., & Burton, J. K. (2012). Second Life in education. A review of publications from its launch to 2011. *British Journal of Educational Technology*, 44(3), 357–371. <https://doi.org/10.1111/j.1467-8535.2012.01334.x>
- Wang, X., Jabri, A., & Efros, A. A. (2019). Learning Correspondence from the Cycle-Consistency of Time. <https://arxiv.org/abs/1903.07593>
- Ward, P., Williams, A. M., & Hancock, P. A. (2006). Simulation for Performance and Training. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Hrsg.), *The Cambridge Handbook of Expertise and Expert Performance* (S. 243–262). Cambridge University Press. <https://doi.org/10.1017/cbo9780511816796.014>
- Werni, J. (2016, 8. Oktober). Meet Ann Clements, the Penn State Professor Pioneering a New Way to Train Teachers. <https://studybreaks.com/>

college/meet-ann-clements-penn-state-professor-pioneering-new-way-train-teachers/

- Weston, T., Kosko, K., Amador, J., & Estapa, A. (2018). Preservice Teachers Questioning: Comparing Platforms for Practice-Based Teacher Education. *Journal of Technology and Teacher Education*, 26(1), 149–172. <https://www.learntechlib.org/p/181137>
- Wettstein, A. (2008). *Beobachtungssystem zur Analyse aggressiven Verhaltens in schulischen Settings (BASYS)*. Huber.
- Whitehead, A. N. (2012). *Die Ziele von Erziehung und Bildung* (C. Kann & D. Sölch, Hrsg.; 1. Edition) [Original publiziert 1927]. Suhrkamp.
- Wiepke, A., & Paulicke, P. (2018). Aufbau von Handlungswissen zum Umgang mit Unterrichtsstörungen im VR-Klassenraum [Wettbewerbsbeitrag zur "Gelungene VR/AR-Lernszenarien" des Stifterverband der deutschen Wissenschaft, GI-Fachgruppen E-Learning und VR/AR 2018]. https://ak-vrarl.gi.de/fileadmin/AK/VRAR/PDF/VRARL_Wettbewerb2018-Klassenraum-20180522-final.pdf
- Wilson, K., & Korn, J. H. (2007). Attention During Lectures: Beyond Ten Minutes. *Teaching of Psychology*, 34(2), 85–89. <https://doi.org/10.1080/00986280701291291>
- Woods, P. (1993). Critical Events in Education. *British Journal of Sociology of Education*, 14(4), 355–371. Verfügbar 5. Juni 2023 unter <http://www.jstor.org/stable/1392882>
- Wooldridge, M. (2002). Intelligent Agents: The Key Concepts. In V. Mařík, O. Štěpánková, H. Krautwurmová & M. Luck (Hrsg.), *Multi-Agent Systems and Applications II* (S. 3–43). Springer Berlin Heidelberg.
- Yeh, Y.-C. (2004). Nurturing reflective teaching during critical-thinking instruction in a computer simulation program. *Computers & Education*, 42(2), 181–194. [https://doi.org/10.1016/s0360-1315\(03\)00071-x](https://doi.org/10.1016/s0360-1315(03)00071-x)
- Yeh, Y.-C. (2006). The interactive effects of personal traits and guided practices on preservice teachers changes in personal teaching efficacy. *British Journal of Educational Technology*, 37(4), 513–526. <https://doi.org/10.1111/j.1467-8535.2006.00550.x>
- Yeh, Y.-c. (2002). Preservice teachers thinking styles, dispositions, and changes in their teacher behaviors. *International Conference on Computers in Education, 2002. Proceedings*. <https://doi.org/10.1109/cie.2002.1186019>
- Yıldırım, G. (2016, 18. Februar). *Öğretmenler yapay zeka öğrencilere karşı*. <https://www.aa.com.tr/tr/egitim/ogretmenler-yapay-zeka-ogrencilere-karsi/523281>
- Zazula, A., Myszor, D., Antemijczuk, O., & Cyran, K. A. (2013). Flight Simulators - From electromechanical analogue computers to modern

- laboratory of flying (Review Article). *Advances in Science and Technology Research Journal*, 7, 51–55.
- Zibit, M., & Gibson, D. (2005). simSchool: The Game of Teaching [Article 4]. *Innovate: Journal of Online Education*, 1(6). <http://nsuworks.nova.edu/innovate/voll/iss6/4>
- Zidik, Z. (2015). *Blog/Entwicklungstagebuch von Zac Zidik, Artikelsammlung 'Tech to Teach the Teacher'*. <https://sites.psu.edu/zaczidik/author/zez1/>
- Zifreund, W. (1966). *Konzept für ein Training des Lehrverhaltens mit Fernseh-Aufzeichnungen in Kleingruppenseminaren*. Cornelsen.

Abbildungsverzeichnis

1	Screenshot eines <i>intelligent tutoring systems (ITS)</i> , Quelle: MacLellan (2017, S.35)	37
2	Hypothetischer Zusammenhang zwischen Simulationsfidelität und Lernoutcome, Quelle: Alessi und Trollip (2001, S.234)	40
3	Beispielhafte Veranschaulichung eines deterministischen Entscheidungsbaums, Quelle: Autor, adaptiert von Deale und Pastore (2014, S.199)	56
4	Beispielhafte Veranschaulichung eines randomisierten Entscheidungsbaums, Quelle: Autor	57
5	Beispielhafte Veranschaulichung eines komplexen Entscheidungsbaums mit algorithmischer Steuerung, Quelle: Author	58
6	Interaktionen in einer Simulation mit komplexer algorithmischer Steuerung, Quelle: Deale und Pastore (2014, S.199)	59
7	Aufgaben-SuS-Wechselwirkung in simSchool, Quelle: Hettler et al. (2008, S.25)	116
8	OCEAN-Änderungen nach Aufgaben in simSchool, Quelle: Hettler et al. (2008, S.25)	117
9	Das DOIT-Modell, Quelle: Horz und Ulrich (2013)	143
10	Schaubild eines Beispiels für Handlungsinitiierungen durch die Simulationspartizipantinnen, Quelle: Autor	154
11	Schematische Darstellung der Partizipantinnenperspektive in Clasivir, Quelle: Autor	155
12	Zuordnung wichtiger Clasivir-Variablen zum verwendeten Framework, Quelle: Kunter und Trautwein (2013) und Autor	157
13	Screenshot des programmierten Prototyps, Quelle: Autor	159
14	Graphische Darstellung der Konzentrationsverfall-Funktion	169
15	Beispiel der Konzentrationsverfall-Funktion nach Lehrkraftinteraktionen	170
16	Screenshot des Clasivir Prototyps, Veranschaulichung der Wahrscheinlichkeiten für Störungsverhalten eines Schülers	176
17	Graphische Darstellung der Anderson/Krathwohl'schen Taxonomie	178
18	Exponentielles Abklingen im Lösungsalgorithmus	182
19	Balkendiagramm der Gruppeneinteilungen	197
20	2D-Mockup Video der Simulationsmodelle	200
21	3D-Mockup Video der Simulationsmodelle	201
22	Nutzung des Kinect-Sensors zur Erstellung von motion capturing-Daten, Quelle: Autor	205

23	Darstellung des manuellen Animationsverfahrens in Blender, Quelle: Autor	206
24	Boxplots der Fidelitätswerte aller Videos der verschiedenen Darstellungstypen	219
25	Balkendiagramm der wahrgenommenen Schwierigkeit nach visueller Darbietung	223
26	Boxplots der FIDEC-Werte geordnet nach Videospieldnutzung Computer	224
27	Balkendiagramm des wahrgenommenen Nutzens nach Geschlecht	225
28	Zusammenhang zwischen Videospielderfahrungen (Computer und Konsole) und wahrgenommener Schwierigkeit	226
29	Boxplots der Fidelitätsverteilung zwischen dem Clasivir-Modell der präsentierten 1. Stunde und der 2. Stunde	227
30	Scatterplot zwischen der FIDEC-Subskala und dem wahrgenommenen Nutzen	228
31	Screenshot von Clasivir 1.0, Quelle: Autor	239
32	Aha! Classroom Sim, Screenshot 1	242
33	Aha! Classroom Sim, Screenshot 2	242
34	Aha! Classroom Sim, Screenshot 3	242
35	Behavior Breakthroughs, Screenshot 1	243
36	Behavior Breakthroughs, Screenshot 2	243
37	Breaking Bad Behaviors, Screenshot 1	243
38	Breaking Bad Behaviors, Screenshot 2	244
39	Breaking Bad Behaviors, Screenshot 3	244
40	ClassSim, Screenshot 1	245
41	ClassSim, Screenshot 2	245
42	Cook School District Simulation, Screenshot 1	245
43	Cook School District Simulation, Screenshot 2	245
44	Der simulierte Klassenraum (SKR), Screenshot 1	246
45	Der simulierte Klassenraum (SKR), Screenshot 2	246
46	Diversity Simulation, Screenshot 1	246
47	Diversity Simulation, Screenshot 2	246
48	First Class, Screenshot 1	247
49	First Class, Screenshot 1	247
50	IVT/IVT-T, Screenshot 1	247
51	IVT/IVT-T, Screenshot 2	247
52	At-Risk for Middle School Educators (Kognito Interactive), Screenshot	248
53	At-Risk for High School Educators (Kognito Interactive), Screenshot	248

54	LessonSketch, Screenshot	248
55	ThEMaT, Screenshot	248
56	Aufbauskizze der Simulation Mr. Land's Sixth Grade	249
57	SimAula 1.0, Screenshot 1	249
58	SimAula 1.0, Screenshot 2	249
59	SimAula 2.0, Screenshot 1	250
60	SimAula 2.0, Screenshot 2	250
61	SimClass (Flash), Screenshot 1	250
62	SimClass (Flash), Screenshot 2	250
63	SimClass (web), Screenshot 1	251
64	SimClass II, Screenshot	251
65	simSchool 1.0, Screenshot 1	252
66	simSchool 1.0, Screenshot 2	252
67	simSchool 2.0, Screenshot 1	252
68	simSchool 2.0, Screenshot 2	252
69	simSchool 2.0, Screenshot 3	253
70	simSchool 2.0, Screenshot 4	253
71	SimInClass (englischsprachige Version von Simfta), Screenshot	253
72	Simfta, Screenshot	253
73	Foto eines Interaktors, der virtuelle SuS über einen <i>motion-</i> <i>capturing</i> Anzug spielt	254
74	Screenshot von TeachLivE/TeachMe	254
75	Foto von TeachLivE-Simulationspartizipantinnen, die mit der Simulation interagieren	254
76	Screenshot von Mursion	254
77	SimTeacher, Screenshot 1	255
78	SimTeacher, Screenshot 2	255
79	Tigerlake Simulation, Screenshot 1	255
80	Tigerlake Simulation, Ausdruck	255
81	Transition to Teaching Classroom Management, Screenshot 1 .	256
82	Transition to Teaching Classroom Management, Screenshot 1 .	256
83	Virtual Classroom Simulator (VCS), Screenshot 1	256
84	Virtual Classroom Simulator (VCS), Screenshot 2	256
85	Virtual Classroom Simulator (VCS), Screenshot 3	257
86	Verbal Abuse Simulation, Screenshot	257
87	Virtual School Class (VSC), Screenshot	257
88	Virtual Simulated ClassRoom (V.C.R.), Screenshot	258
89	VR-Klassenraum, Screenshot	258
90	VR-Tool von Bailenson et al., Screenshot 1	259
91	VR-Tool von Bailenson et al., Screenshot 2	259
92	VR-Tool von Yaacov J. Katz, Screenshot	259

93	VR-Tool von Stavroulia et al., Screenshot 1	260
94	VR-Tool von Stavroulia et al., Screenshot 2	260
95	Integration verschiedener Konzeptionsmodelle für die Zielgröße akademische Leistung, Quelle: Autor	261
96	Integration verschiedener Konzeptionsmodelle für die Zielgröße aggressives Verhalten, Quelle: Autor	262
97	Integration verschiedener Konzeptionsmodelle für die Zielgröße Emotionen, Quelle: Autor	263
98	Integration verschiedener Konzeptionsmodelle für die Zielgröße Fehlverhalten, Quelle: Autor	263
99	Integration verschiedener Konzeptionsmodelle für die Zielgröße Langeweile, Quelle: Autor	264
100	Integration verschiedener Konzeptionsmodelle für die Zielgröße Langeweile, Quelle: Autor	264
101	Integration verschiedener Konzeptionsmodelle für die Zielgröße Unterrichtsstörungen, Quelle: Autor	264
102	Integration verschiedener Konzeptionsmodelle für die Zielgröße Unterrichtsstörungen, Quelle: Autor	264
103	Integration verschiedener Konzeptionsmodelle für die Zielgröße Unterrichtsstörungen, Quelle: Autor	265

Tabellenverzeichnis

1	Störungsverhalten in Breaking Bad Behaviors, Quelle: Lugin et al. (2016, S.4)	76
2	Verwendete Variablen in der Cook School District Simulation .	84
3	<i>Traits</i> in simSchool 2.0	123
4	Übersicht über die Einteilung der Gruppen	195
5	Übersicht über die verwendeten Mockup-Videos	195
6	Geschlechtsangaben der Partizipantinnen und Partizipanten .	196
7	Partizipantinnen nach Fachsemester	197
8	Deskriptive Fidelitätswerte aller Darstellungsformen	218
9	Angegebene Studiengänge	220
10	Angegebene Fachsemester	220
11	Angegebene Videospielerfahrungen	221
12	Unterschiedlichkeitseinschätzung der Videos	221
13	Schwierigkeit in der Bewertung der Videos	222
14	Schwierigkeit nach grafischer Aufbereitung	222
15	Zuweisung der Nation des Migrationshintergrunds	266
16	Markov-Modell-Algorithmus zur Modellierung des Wechsels emotionaler Zustände	269
17	Liste aller in Clasivir implementierten Störverhalten	270
18	Modifikationen der Störungswahrscheinlichkeiten durch emotionale Stati	272
19	Berechnung der Aufgabenschwierigkeit	273
20	Beispieltabelle der Aufgabenadaptivität für eine Schülerin mit KF=.5	274
21	Adaptive Verarbeitungswahrscheinlichkeiten	275
22	Berechnung der <i>functional-off-task</i> Wahrscheinlichkeiten . . .	275
23	Adaptive Wahrscheinlichkeiten für Hereinrufen	276
24	Adaptive Meldewahrscheinlichkeiten nach erfolgter Verarbeitung	277
25	Mögliche Meldungstypen nach erfolgter Meldung	277
26	Konzentrationsmodifikationen nach erfolgter <i>classroom management</i> Handlung	278
27	Konzentrationsmodifikationen nach erfolgter Feedback-Handlung (stets Einzelziel)	278
28	Liste aller über die ini-Datei modifizierbaren Simulationsvariablen	279

Abstract

Die vorliegende Arbeit beschreibt die Entwicklung eines interaktionalen Simulationsmodells zum späteren Einsatz in der VR-Simulation Clasivir 2.0 (Classroom Simulator in Virtual Reality), welche in der Lehrkräftebildung eingesetzt werden soll. Das Clasivir-Simulationsmodell wurde im Rahmen eines Prototyps implementiert und zwei anderen Simulationsmodellen in einem Fragebogen entgegengestellt. Ein Simulationsmodell beschreibt im Kontext einer digitalen Schulunterrichtssimulation, wie sich SuS in der Simulation verhalten.

Die drei Simulationsmodelle wurden über zwei unterschiedliche Typen von Video-Visualisierungen, genannt Mockup-Videos, dargestellt: Zum einen über eine 2D-Darstellung aus Vogelperspektive, zum anderen über eine 3D-Darstellung, in welcher 3D-Modelle von SuS animiert wurden. Bei dem realen Simulationsmodell handelt es sich um eine Übertragung einer authentischen Videoaufzeichnung von Unterricht einer hessischen Realschule in 2D/3D-Visualisierungen. Im randomisierten Simulationsmodell führen SuS ihre Verhalten zufällig aus. Alle Modelle basieren auf zweisekündigen Intervallen. Im Falle des realen Simulationsmodells wurde dies durch Analyse aller beobachtbaren einundzwanzig SuS gewonnen, im Falle des Clasivir-Simulationsmodells wurden die Vorhersagen des Simulationsmodells übertragen. Das Simulationsmodell von Clasivir basiert auf *behavior trees*, stellt eine Art von künstlicher Intelligenz dar und modelliert das SuS-Verhalten größtenteils in Abhängigkeit von Lehrkrafthandlungen. Die Entwicklung des interaktionalen Simulationsmodells von Clasivir ist eine Kernkomponente dieser Arbeit. Das Simulationsmodell basiert auf empirischen Ergebnissen aus den Bereichen der Psychometrie, der pädagogischen Psychologie, der Pädagogik und Ergebnissen der Simulations-/KI-Forschung. Ziel war die Entwicklung eines Modells, das nicht nur auf normativen Vorhersagen basiert, sondern empirisch und theoretisch valide ist. Nur wenige Simulationsmodelle in Unterrichtssimulationen werden mit dieser Art von Transparenz beschrieben, was eines der Alleinstellungsmerkmale dieser Arbeit ist. Es wurden Anstrengungen unternommen die vorliegenden empirischen Ergebnisse in einen kausalen Zusammenhang zu bringen, der mathematisch modelliert wurde. Im Zentrum steht die Konzentration von SuS, welche Einfluss auf Stör-, Melde- und Antwortverhalten hat. Diese Variable wird durch andere situative und personenbezogene Variablen (im Sinne von *traits*) ergänzt. Wo keine direkten empirischen Ergebnisse vorlagen wurde versucht plausibles Verhalten anhand der Übertragung von Konzeptionsmodellen zu gewinnen.

Da die bisherige Verwendung der angrenzenden Begriffe rund um die Simulationsentwicklung bislang sehr inkonsistent war, wurde es notwendig die-

se Termini zu definieren. Hervorzuheben ist die Entwicklung einer Taxonomie digitaler Unterrichtssimulationen, die so bislang nicht existierte. Anhand dieser Taxonomie und der erarbeiteten Fachtermini wurden Simulationen in der Lehrkräftebildung auf ihre Modellierung des Simulationsmodells hin untersucht. Die Untersuchung der Simulationen simSchool und VCS war, da sie einen verwandten Ansatz zu Clasvir verfolgen, besonders ergiebig.

Nach der Generierung der Mockup-Videos wurden $N=105$ Studierende, $N=102$ davon Lehramtsstudierende, aufgefordert, in einem Online-Fragebogen zwei der Simulationsmodelle miteinander zu vergleichen. Lehramtsstudierende wurden ausgewählt, da sie die Zielgruppe der Simulation sind. Welche Modelle die Partizipantinnen verglichen, war abhängig von der Gruppe der sie zugeteilt wurden. Hierbei wurde neben den Simulationsmodellen auch die visuelle Darstellung variiert. Insbesondere wurden die Partizipantinnen darum gebeten, den Fidelitätsgrad des Simulationsmodells, also den Maßstab, wie realistisch die Partizipantinnen das Verhalten der SuS in der Simulation fanden, zu bewerten. Inferenzstatistisch bestätigte sich, dass Partizipantinnen keinen Unterschied zwischen dem realen Simulationsmodell und dem Clasvir-Simulationsmodell erkennen konnten ($t=1.463$, $df=178.9$, $p=.1452$), aber das randomisierte Simulationsmodell mit einer moderaten Effektstärke von $d=.634$ als signifikant schlechter einschätzten ($t=-2.5231$, $df=33.581$, $p=.008271$). Die Art der Darbietung (2D oder 3D) hatte keinen statistisch signifikanten Einfluss auf die wahrgenommene Schwierigkeit der Bewertung ($z=1.2426$, $p=.107$, $\eta^2=.099633$). Damit kann festgestellt werden, dass eine komplexe und zeitintensive 3D-Visualisierung eines Simulationsmodells bei noch nicht vorliegender Simulation nicht erforderlich ist. Das Clasvir-Simulationsmodell wird als realistisch wahrgenommen. Es kann damit empfohlen werden, es in der VR-Simulation zu verwenden.

Im Ausblick werden bereits während des Schreibens der Arbeit gemachte Entwicklungen beschrieben und Konzepte zum weiteren Einsatz der Ergebnisse entwickelt. Es wird darauf verwiesen, dass eine erste Version eines VR-Simulators entwickelt wurde (Clasvir 1.0), der jedoch rein deterministisch funktioniert und noch nicht das in dieser Arbeit entwickelte Simulationsmodell inkludiert.

Detaillierte Stellungnahme zu anderen Studien des Autors in der Dissertationsschrift

Der Autor verfasste zwei andere Schriften, die im Rahmen dieser monographischen Dissertationsarbeit zitiert wurden. Im Folgenden soll Stellung zum Anteil des Autors an Entwicklung der Fragestellung, Design, Durchführung und Auswertung der Studien genommen werden. Alle hier beschriebenen Studien und Textanteile des Autors haben die gute wissenschaftliche Praxis beachtet und sind eigenständig ohne Verwendung weiterer Hilfsmittel erstellt worden. Alle Entlehnungen aus anderen Schriften wurden mit Angabe der betreffenden Schriften gekennzeichnet.

(1) Breitenbach und Appel (2016)

Die Schrift erschien in dem Sammelband „Videoanalysen in der Unterrichtsforschung. Methodische Vorgehensweisen und Anwendungsbeispiele“ und fußt auf den Video-Untersuchungen zum Thema Heterogenität unter Prof. Dr. Udo Rauin/Dr. Johannes Appel, welche dem QLB-Projekt Level vorgelagert waren. Der Autor war zu diesem Zeitpunkt studentische Hilfskraft und unterstützte Dr. Johannes Appel in technischen Verfahrensfragen und bei der Kodierung. Dabei gewann der Autor Expertise über verschiedene technische Verfahren und Software-Pakete zur Kodierung von videografischem Material, die über die eigentliche Fragestellung hinaus gehen. Neben dem von Dr. Appel verwendeten Videopaket Interact wurden auch die Applikationen Videograph, ANVIL und ELAN für eine Anwendung im Kontext der Unterrichtsbeobachtung untersucht. Dies erfolgte vollständig durch den Autor. Johannes Appel steuerte einen einleitenden Text bei (Kapitel 1) und ergänzte vereinzelt Literaturangaben in Kapitel 2. Zudem teilten sich der Autor und Dr. Appel das Fazit arbeitsteilig etwa zur Hälfte. Den größten Teil der Arbeit (Kapitel 3 „Vergleich der Softwarepakete“, S.158–171) wurden ausschließlich vom Autor dieser Dissertation geschrieben.

(2) Glocker et al. (2023)

Diese Schrift beschäftigt sich mit der Messung von Leistungsemotionen im Rahmen der Implementierung der VR-Simulation Clasivir 1.0. Der Autor arbeitete im Digi_Gap Projekt in Leitungsfunktion und kümmerte sich vorrangig um die Konzeptionalisierung der Simulation (Deterministische Entscheidungsbäume, Generierung des ersten *critical incidents* über Flußdiagramme) und die technische Umsetzung (Erstellung von Animation, Ansprechpartner für den externen Dienstleister NMY). Im Artikel gehen vor allem die Teile, die die technische und technisch-

konzeptionelle Funktionsweise des Simulators beschreiben auf den Autor zurück. Der Autor unterstützte Frau Glocker in der Transformation der gewonnenen Simulationsdaten in R, war aber an der statistischen Prüfung nicht beteiligt.