

Dissertation
zur Erlangung des Doktorgrades
Dr. phil. nat.
der Fakultät für Physik
an der Universität Duisburg-Essen

Motivierstile von Tutoren physikalischer Experimente

Intervenierbarkeit, Prädiktoren und Konsequenzen

vorgelegt von
Thomas Schlake
geboren in Twistringen

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

Diese Dissertation wird via DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/75826

URN: urn:nbn:de:hbz:465-20220421-103307-9

Alle Rechte vorbehalten.

Diese Arbeit entstand im Rahmen des Projektes ‚Ganz In. Mit Ganzttag mehr Zukunft. Das neue Ganzttagsgymnasium NRW‘. Das Projekt wurde von der Stiftung Mercator gefördert.

Erstgutachter: Prof. Dr. Hendrik Härtig
Zweitgutachter: Prof. Dr. Heiko Krabbe
Prüfer: Prof. Dr. Hans E. Fischer
Vorsitzender des Prüfungsausschusses: Prof. Dr. F.-J. Meyer zu Heringdorf
Tag der mündlichen Prüfung: 16.03.2022

Danksagung

Ich bedanke mich herzlich bei Prof. Dr. Hans Fischer, der mir im Rahmen meiner Stelle als wissenschaftlicher Mitarbeiter seiner AG meine Promotion ermöglicht hat und mich am Tag meiner mündlichen Prüfung als Prüfer begleitete. Er hat mir stets den Rücken gestärkt, mir viele Freiheiten in der Gestaltung meiner Arbeit gelassen und an den richtigen Stellen wertvolle Impulse gegeben. Weiterhin gilt mein tiefer Dank Prof. Dr. Hendrik Härtig, welcher nach dem Weggang von Prof. Dr. Hans Fischer die AG und die Erstbetreuung meiner Promotion herausragend übernahm und mir stets mit viel Unterstützung, Weitsicht und Ruhe zur Seite stand. Weiterhin danke ich Prof. Dr. Heiko Krabbe für das Zweitgutachten meiner Dissertation, die Prüfung und wertvolle Unterstützung sowie die Arbeit am Laborhelferprojekt.

Meine Bürokollegen Joachim Müller und Christine Boubakri habe ich menschlich und fachlich sehr zu schätzen gelernt. Ein besonderer Dank gilt meinem weiteren Bürokollegen und Freund Robert Aleksov, mit dem ich gemeinsam viele Hoch- und Tiefpunkte der Promotionszeit erleben durfte. Ein weiter Dank gilt Anita Stender und Maria Opfermann für viele wertvolle und inhaltlich anregende Gespräche. Darüber hinaus bedanke ich mich herzlich bei Cornelia Geller, Lina Boyer, Nicole Kohnen, Barbara Steffentorweihen, Heike Theyßen, Thomas Kersting, Stefan Kirchner, Rasmus Viefers, Cihan Tomaz, und Nico Schreiber.

Für die logistische Unterstützung bedanke ich mich bei Claudia Evers, Anne Hager und Eike Lange. Außerdem haben studentische Hilfskräfte wertvolle Arbeit für das Projekt geleistet, vor allem: Éva Pásztor, Patrick Wittich, Nils Kuhlmann und Rebecca Lippert. Für die Finanzierung des Teilprojektes im Rahmen des Projekts *Ganz In. Mit Ganztags mehr Zukunft. Das neue Ganztagsgymnasium NRW* bedanke ich mich bei der Stiftung Mercator.

Insbesondere gilt mein Dank meiner Familie und meinen Freunden, besonders meiner Freundin Janine Graser und meiner Ex-Frau Mareike Schlake. Eure Geduld und Hilfe sind unbezahlbar.

Zusammenfassung

Da die generell niedrig ausgeprägte intrinsische Motivation für verschiedene physikalische Tätigkeiten im Altersbereich von zehn bis dreizehn Jahren besonders stark abnimmt, kommt der Förderung intrinsischer Motivation eine herausgehobene Bedeutung zu (Flunger et al., 2019; Gaspard et al., 2017; Ryan & Deci, 2017; Wigfield et al., 2015). Die Betreuung von jüngeren Schülerinnen und Schülern (Tutees) durch ältere (Tutoren) kann sowohl bei den Tutoren als auch bei den Tutees Lernzuwächse, die intrinsische Motivation bezüglich der Tätigkeiten während des Tutorings und damit insgesamt schulische Leistungen steigern (P. Cohen et al., 1982; Hänze et al., 2018; Robinson et al., 2005; Topping et al., 2004). Im Rahmen der vorliegenden Studie wurde eine Intervention für angehende Peer-Tutoren außerunterrichtlicher Angebote entwickelt, bei der die Tutoren lernen, ihre Tutees motivationsförderlich bei der Durchführung physikalischer Experimente zu betreuen. Die Intervention wurde in Anlehnung an die etablierte *Selbstbestimmungstheorie* von Ryan und Deci (2017) entwickelt, entsprechend der die intrinsische Motivation von Tutoren und Tutees am besten durch die Förderung einer autonomieunterstützenden Betreuung und die Reduktion einer kontrollierenden Betreuung der Tutees durch die Tutoren gefördert werden kann. Insbesondere wurden allgemeine Empfehlungen für ein motivationsförderliches Verhalten von Lehrkräften auf physikalische Tutoring-Angebote übertragen. Es wurde untersucht, ob die Tutees eine höhere intrinsische Motivation bezüglich der Durchführung des untersuchten physikalischen Experiments angeben, wenn sie sich in ihrem Autonomieerleben unterstützt fühlen. Im Zuge dessen wurde auch untersucht, inwiefern die Teilnehmenden unterschiedlich von der Intervention profitieren, inwiefern sie das Gelernte in die Praxis umsetzen und welche weiteren Faktoren bei der Umsetzung eine Rolle spielen. Zur Untersuchung dieser Forschungsfragen haben 176 Tutoren ($M_{Alter} = 13.73 \pm 0.75$ Jahren) an der entwickelten Intervention teilgenommen. 135 davon haben 135 Tutees ($M_{Alter} = 10.81 \pm 0.48$ Jahren) bei der Durchführung eines Experimentes betreut. Ziel des Experiments war es, die Leitfähigkeit von Kabeln verschiedener Längen, Dicken und Materialien anhand der Leuchtstärke von Glühlampen zu untersuchen. Die Ergebnisse zeigen erwartungskonform, dass die Tutees sich als umso kompetenter erleben und umso stärker intrinsisch motiviert sind, je mehr sie sich in ihrem Autonomieerleben unterstützt fühlen. Die Intervention kann mit einem großen Effekt ($d_{Cohen} > 1.2, p < .001$) die selbstberichtete Tendenz zu Autonomieunterstützung fördern und die Tendenz zu Kontrolle reduzieren. Teilnehmende mit einer hohen präinterventionellen Tendenz zu Autonomieunterstützung profitieren weniger von der Intervention, da sie vermutlich ein geringeres Entwicklungspotential haben. Ein weiterer Prädiktor des Interventionserfolgs ist die Zustimmung zu den Interventionseinhalten ($\beta \in [.26, .42], p < .05$). Alter, Geschlecht sowie weitere Kontrollvariablen scheinen für den Interventionserfolg irrelevant zu sein. Die selbstberichtete Tendenz zu Kontrolle korreliert mit dem beobachteten Tutorverhalten während einer Tutoring-Sitzung ($r = .48, p < .001$). Des Weiteren betreuen Tutoren ihre Tutees in Situationen, in denen keine Hilfe nötig ist, signifikant kontrollierender ($\beta = -.445, p < .001$), wenn sie ihre Tutees als wenig selbstständig wahrnehmen. Sowohl die Tendenz zu Kontrolle als auch die Wahrnehmung der Tutees sagen das Tutorverhalten vorher. Zusammengefasst zeigen die Ergebnisse die Relevanz von Autonomieunterstützung für physikalische Tutoring-Angebote sowie die Nützlichkeit der entwickelten Intervention. Weiterhin konnten durch die detaillierten Analysen zahlreiche Implikationen für die Überarbeitung der Intervention und für die zukünftige Forschung formuliert werden. Somit stellt die überarbeitete Intervention ein gewinnbringendes Angebot für Schulen dar.

Abstract

Due to the especially strong decline of the generally low intrinsic motivation for various physical activities between the ages of ten and thirteen, it is of particular importance to promote intrinsic motivation in this age group (Flunger et al., 2019; Gaspard et al., 2017; Ryan & Deci, 2017; Wigfield et al., 2015). When older students (tutors) support younger students (tutees), they may both gain knowledge as well as intrinsic motivation regarding the activities done during the tutoring (P. Cohen et al., 1982; Hänze et al., 2018; Robinson et al., 2005; Topping et al., 2004). This can improve their overall academic performance (P. Cohen et al., 1982; Hänze et al., 2018; Robinson et al., 2005; Topping et al., 2004). An intervention for prospective peer tutors of extracurricular programs was designed for the study. The intervention teaches tutors to foster tutees' intrinsic motivation while conducting physics experiments. The intervention was developed following the well-established *Self-Determination Theory* (Ryan & Deci, 2017), which suggests that the intrinsic motivation of tutors and tutees can best be fostered by encouraging the tutors to behave in an autonomy-supportive manner and discouraging them from behaving in a controlling manner. In particular, general recommendations for teachers were applied to extra-curricular physics tutoring programs. The study assessed whether tutees would report higher intrinsic motivation regarding the conduct of the physics experiment when they felt that their autonomy was supported. The study also assessed the extent to which participants benefit differently from the intervention, the extent to which they put what they have learned into practice, and what other factors play are relevant to implementation. 176 tutors ($M_{age} = 13.73 \pm 0.75$ years) participated in the designed intervention. 135 of them supported 135 tutees ($M_{age} = 10.81 \pm 0.48$ years) in conducting an experiment. The aim of the experiment was to investigate the conductivity of cables of different lengths, thicknesses and materials based on the luminosity of incandescent lamps.

As expected, the results show that tutees feel more competent and are more intrinsically motivated when their autonomy is supported. The intervention promotes tutors' self-reported tendency toward autonomy support and reduces their tendency toward control with a large effect ($d_{Cohen} > 1.2, p < .001$). Participants with a high preinterventional tendency toward autonomy support benefit less from the intervention, as they are likely to have lower potential for development. Another predictor for the success of the intervention is the participants' overall agreement with the contents of the intervention ($\beta \in [.26, .42], p < .05$). Age, gender, and other control variables appear to be irrelevant to the success of the intervention. The self-reported tendency toward control correlates with the observed tutor behavior during tutoring sessions ($r = .48, p < .001$). In situations in which the tutees do not require or want help, tutors generally behave in a significantly more controlling manner ($\beta = -.445, p < .001$) when they perceive their tutees as unable to work in an autonomous way. Both the tendency to control and the tutors' perception of their tutees predict the tutors' behavior. Altogether, the results demonstrate the relevance of autonomy support for physics tutoring programs as well as the value of the developed intervention. Furthermore, the detailed analyses provided numerous implications for the revision of the intervention as well as for future research. Thus, the revised intervention represents a valuable program for schools.

Inhaltsverzeichnis

Danksagung	v
Zusammenfassung	vi
Abstract	vii
1 Einleitung	1
2 Peer-Tutoring	3
3 Selbstbestimmungstheorie	7
4 Operationalisierung der Motivierstile	11
5 Die Einflüsse der Motivierstile	17
5.1 Wirkungsweise der Motivierstile der Betreuenden auf ihre Betreuten	17
5.2 Einflüsse auf die Motivierstile	19
5.2.1 Der Druck von Innen	19
5.2.2 Der Druck von Unten	21
5.2.3 Der Druck von Oben	22
5.2.4 Das zu untersuchende Wirkmodell der Motivierstile und der individuellen Drücke	23
6 Intervenierbarkeit der Motivierstile	29
6.1 Perspektiven der Erfassung der Motivierstile	29
6.2 Veränderlichkeit der Motivierstile	29
6.3 Gestaltungsmerkmale effektiver Interventionen zur Förderung eines autonomieunterstützenden und Reduktion eines kontrollierenden Motivierstils . . .	30
6.4 Intervenierbarkeit der Perspektiven der Motivierstile von Cross-Age-Tutoren	32
7 Forschungsfragen und Hypothesen	33
7.1 Forschungsfrage 1	33
7.2 Forschungsfrage 2	34
7.2.1 Hypothesen über die Änderungen der Motivierstile durch die Intervention	35
7.2.2 Hypothesen über Prädiktoren der Änderungen der Motivierstile . . .	35
7.2.3 Hypothesen über die Zustimmung zu den Interventionsinhalten als Prädiktor der Änderungen der Motivierstile für Teilnehmende mit unterschiedlichen Kausalitätsorientierungen	36
7.3 Forschungsfrage 3	37
7.4 Forschungsfrage 4	38

8	Design und Methoden	41
8.1	Studienablauf	41
8.2	Das Betreuungsexperiment der Tutoring-Sitzung	44
8.3	Konzeption und Entwicklung der Intervention	44
8.4	Konzeption und Entwicklung der Messinstrumente	48
8.4.1	Das Instrument zur Erfassung der selbstberichteten Motivierstile für Cross-Age-Tutoren	48
8.4.2	Das Instrument zur Erfassung der beobachteten Motivierstile von Cross-Age-Tutoren durch Verhaltensanalyse	50
8.4.3	Das Instrument zur Erfassung des wahrgenommenen autonomieunterstützenden Motivierstils von Cross-Age-Tutoren durch Selbstberichte der Tutees	57
8.4.4	Intrinsische Motivation und Kompetenzerleben der Tutees	59
8.4.5	Übrige Instrumente	61
8.4.6	Konvergente Validität der Instrumente anhand einiger Pilotierungserhebungen	63
8.5	Stichprobe	65
9	Deskriptive Ergebnisse	67
9.1	Deskriptive Analyse der Verhaltensdaten	67
9.2	Struktur und Interne Konsistenzen der Instrumente	71
9.2.1	Motivierstile der Tutoren erfasst durch Selbstberichte	71
9.2.2	Motivierstile der Tutoren erfasst durch Verhaltensbeobachtung	72
9.2.3	Motivierstile der Tutoren erfasst durch Berichte der Tutees/ Wahrgenommener autonomieunterstützender Motivierstil	74
9.2.4	Intrinsische Motivation und Kompetenzerleben der Tutees	75
9.2.5	Übrige Instrumente	75
9.2.6	Zusammenfassung der internen Konsistenzen und inhaltliche Einordnung	77
9.3	Untersuchung der Verteilung der Skalen und geplante Auswertungsverfahren	79
9.4	Vergleichbarkeit der Substichproben zur Bildung einer Gesamtstichprobe	81
10	Ergebnisse zur Beantwortung der Forschungsfragen	83
10.1	Ergebnisse der Forschungsfrage 1	83
10.2	Ergebnisse der Forschungsfrage 2	84
10.2.1	Die Mittelwertunterschiede der Motivierstile und Kausalitätsorientierungen vor und nach der Intervention	84
10.2.2	Untersuchung der Prädiktoren der Mittelwertsunterschiede	85
10.2.3	Die Zustimmung zu den Interventionsinhalten als Prädiktor der Änderungen der Motivierstile für Teilnehmende mit unterschiedlichen Kausalitätsorientierungen	89
10.3	Ergebnisse der Forschungsfrage 3	90
10.4	Ergebnisse der Forschungsfrage 4	99
10.4.1	Zusammenhänge der Konstrukte auf der Tuteeebene	99
10.4.2	Relevanz der Geschlechterkomposition	105
11	Diskussion der Ergebnisse	107
11.1	Diskussion der Struktur der Motivierstile (Forschungsfrage 1)	109

11.2 Diskussion der Intervention (Forschungsfrage 2)	109
11.2.1 Diskussion der Prä-Post-Mittelwertsunterschiede	110
11.2.2 Diskussion der Prädiktoren der Prä-Post-Mittelwertsunterschiede . .	110
11.2.3 Diskussion differentieller Effekte	114
11.3 Diskussion der Verhaltensprädiktoren (Forschungsfrage 3)	114
11.4 Diskussion der Tuteeebene (Forschungsfrage 4)	118
11.4.1 Zusammenhänge der Konstrukte auf der Tuteeebene	118
11.4.2 Relevanz der Geschlechterkomposition	119
11.5 Limitationen	120
11.6 Ausblick und Implikationen	123
11.6.1 Implikationen für die Intervention	123
11.6.2 Implikationen für künftige Forschung	127
11.6.3 Implikationen für die Schulpraxis	130
Abbildungsverzeichnis	145
Tabellenverzeichnis	147
Abkürzungsverzeichnis	150
Überblick über die Anhänge	151

1 Einleitung

Das Laborhelferprojekt ist aus der Notwendigkeit heraus entstanden, vielfältige außerunterrichtliche Angebote an Schulen zu ermöglichen. Diese Angebote können von entsprechend ausgebildeten Schülerinnen und Schülern geplant, vorbereitet und unter Beaufsichtigung durchgeführt werden. Im Rahmen des Laborhelferprojektes werden Schülerinnen und Schüler ab der achten Klasse zur Betreuung Jüngerer bei MINT-Tätigkeiten ausgebildet und eingesetzt. Beispielsweise betreuen Laborhelferinnen und Laborhelfer Experimente an Tagen der offenen Tür, *Jugend forscht*- oder *freestyle-physics*-Teilnehmende in ihrer Vorbereitung oder verschiedene MINT-AGs. Da die Laborhelferinnen und Laborhelfer (unter Aufsicht) eigenständig Betreuungsaufgaben übernehmen, können Lehrkräfte durch die Laborhelferinnen und Laborhelfer entlastet werden. Die Betreuung von jüngeren Schülerinnen und Schülern durch Laborhelferinnen und Laborhelfer kann als Cross-Age-Tutoring bezeichnet werden (siehe Kapitel 2). Cross-Age-Tutoring kann sowohl bei den Tutoren als auch bei den Tutees Lernzuwächse, die intrinsische Motivation bezüglich der Tätigkeiten während des Tutorings und damit insgesamt schulische Leistungen steigern (Berger et al., 2017; P. Cohen et al., 1982; Hänze et al., 2018; Korner, 2015; Kulik & Kulik, 1982; Robinson et al., 2005; Topping et al., 2004).¹ Damit könnte der niedrigen intrinsischen Motivation für verschiedene naturwissenschaftliche und im Speziellen physikalische Tätigkeiten (Flunger et al., 2019; Gaspard et al., 2017; Ryan & Deci, 2017; Wigfield et al., 2015) entgegen gewirkt werden. Da diese intrinsische Motivation im Alter von zehn bis dreizehn Jahren besonders stark abnimmt (Gottfried et al., 2001; Wigfield et al., 2015), setzt das Laborhelferprojekt präventiv bei jüngeren Lernenden im Bereich der fünften bis neunten Klasse an. Um die intrinsische Motivation aller Teilnehmenden bezüglich physikalischer Tätigkeiten möglichst effektiv zu fördern, soll das Projekt, aufbauend auf etablierten psychologischen Motivationstheorien, weiterentwickelt werden. Eine etablierte Motivationstheorie ist die *Selbstbestimmungstheorie* von Ryan und Deci (2017). Diese formuliert Empfehlungen, wie die Motivationen von Individuen bezüglich vielfältiger Handlungen gefördert werden können. In Einklang mit dieser Theorie konnte für Cross-Age-Tutoren gezeigt werden, dass eine autonomieunterstützende Betreuung der Tutees im Vergleich zu einer kontrollierenden Betreuung zu einer höheren intrinsischen Motivation bezüglich der Beschäftigung mit physikalischen Inhalten bei den Tutoren führt (Berger et al., 2017). Außerdem berichten die Tutees der autonomieunterstützenden Versuchsbedingung eine höhere intrinsische Motivation bezüglich der Tätigkeiten während des Tutorings (Hänze et al., 2018). Deshalb wurde die Ausbildung der Laborhelferinnen und Laborhelfer um eine Intervention zur Förderung einer autonomieunterstützenden Betreuung erweitert. In dieser Arbeit soll die Wirksamkeit der Intervention untersucht werden, um sie anschließend entsprechend wei-

¹Vorbemerkung zur verwendeten Sprachform: Im Folgenden sollen sich alle Geschlechter ausdrücklich gleichermaßen repräsentiert fühlen. Deshalb werden im Regelfall mehrere Geschlechter angesprochen (z.B. Schülerinnen und Schüler). Bei englischen Lehnwörtern (z.B. Tutee) wird darauf verzichtet, da eine solche Unterscheidung in der englischen Sprache nicht existiert. Beim Begriff Tutor wird aus Gründen der Parallelität zum Begriff Tutee ebenfalls von einer differenzierten Nennung verschiedener Geschlechter abgesehen.

terzuentwickeln. Die folgenden Fragen sind in diesem Zusammenhang relevant: Profitieren alle Teilnehmenden im gleichen Maße? Inwiefern wird das Gelernte in der Praxis umgesetzt und welche weiteren Faktoren spielen bei der Umsetzung eine Rolle? Um dem oben dargestellten Motivationsproblem für physikalische Tätigkeiten Rechnung zu tragen, wird weiterhin untersucht, ob Schülerinnen und Schüler eine höhere intrinsische Motivation bezüglich der Durchführung des untersuchten physikalischen Experiments angeben, wenn sie sich in ihrem Autonomieerleben unterstützt fühlen. Die Befunde liefern einen Beitrag zur fachdidaktischen und psychologischen Forschung, da untersucht wird, welche individuellen und externen Einflüsse Betreuende zu einem bestimmten Verhalten bewegen, wie diese Einflüsse geändert werden können und inwiefern sich die Wahrnehmung der Tutoren auf die oben genannte Motivation auswirkt.

Die vorliegende Arbeit stellt die oben umrissene Forschung im Detail vor. Zu Beginn werden das Lernsetting Peer-Tutoring im Allgemeinen und Cross-Age-Tutoring im Speziellen vorgestellt (siehe Kapitel 2). Anschließend werden die relevanten Teile der Selbstbestimmungstheorie vorgestellt (siehe Kapitel 3). Autonomieerleben wird dann als psychologisches Grundbedürfnis charakterisiert und es werden die individuellen Orientierungen von Betreuenden zu Autonomieunterstützung bzw. Kontrolle als Motivierstile („*motivating styles*“, Reeve et al., 2014, S. 94) definiert und operationalisiert (siehe Kapitel 4). Im fünften Kapitel wird dargestellt, inwiefern Betreute von einer autonomieunterstützenden Betreuung profitieren können bzw. durch eine kontrollierende Betreuung frustriert werden. Weiterhin werden die möglichen Prädiktoren der Motivierstile, Wirkbeziehungen zu anderen relevanten Konstrukten und deren Operationalisierungen vorgestellt. Danach wird aufbauend auf den vorhergehenden Kapiteln diskutiert, ob der somit als erstrebenswert dargestellte autonomieunterstützende Motivierstil gefördert und der kontrollierende Motivierstil reduziert werden kann (siehe Kapitel 6). Abschließend werden aus der vorgestellten Theorie und den Forschungslücken die Forschungsfragen und Hypothesen abgeleitet (siehe Kapitel 7). Die Untersuchung dieser Hypothesen wird im empirischen Teil der Arbeit vorgestellt (II). Nach der Präsentation der Forschungsmethodik (siehe Kapitel 8), also des Studienablaufs, der entwickelten Intervention sowie der entwickelten und eingesetzten Instrumente, werden zuerst die deskriptiven Ergebnisse (siehe Kapitel 9) und anschließend die Ergebnisse in Bezug auf die aufgestellten Hypothesen dargestellt (siehe Kapitel 10). Abschließend werden die Ergebnisse bezüglich wissenschaftlicher und praktischer Implikationen diskutiert (siehe Kapitel 11). Die Diskussion umfasst u. a. die Darlegung von Limitationen und das Aufzeigen offener Fragen, die in folgenden Forschungsarbeiten untersucht werden können, sowie konkrete Vorschläge zur Weiterentwicklung der Intervention.

2 Peer-Tutoring

Peer tutoring [...] is characterised by specific role-taking as tutor or tutee [...] in which participants receive [...] training. Some peer tutoring methods scaffold the interaction with structured materials, while others prescribe structured interactive behaviours that can be effectively applied to any materials of interest.

(Topping, 2005, S. 632)

Entsprechend der Definition von Topping (2005) stellen Laborhelferinnen und Laborhelfer trainierte Tutoren dar, welche Tutees z. B. bei der Durchführung physikalischer Experimente betreuen. Die Experimentiermaterialien und Arbeitsaufträge geben die in der Definition angegebene Struktur. Nach Topping (2005) ist Peer-Tutoring ein Überbegriff und Tutoring-Angebote können sich bezüglich einiger Kategorien unterscheiden. Im Folgenden werden die fünf Kategorien (1) der Typ des Tutorings, (2) die thematisierten Inhalte, (3) die Ausbildung der Tutoren, (4) die Schulform der Tutoren und Tutees sowie (5) die Geschlechterkomposition und ihr Einfluss auf den Wissenszuwachs der Tutees anhand der Metaanalyse von (Leung, 2015) vorgestellt.

Zum einen können die Rollen des Tutors und des Tutees ausgetauscht werden (Reciprocal oder Non-Reciprocal). Beim Reciprocal-Tutoring können die Schülerinnen und Schüler jeweils Expertinnen oder Experte bezüglich einer Kompetenz oder eines Wissensgebietes sein und sich abwechselnd betreuen. Außerdem können die Tutoren und Tutees ein ähnliches oder unterschiedliches Alter haben (Same-Age oder Cross-Age). Same-Age-Tutoring kann sowohl in festen Rollen, also als Same-Age-Non-Reciprocal-Tutoring, sowie mit wechselnden Rollen auftreten, also als Same-Age-Reciprocal-Tutoring. Cross-Age-Tutoren sind im Regelfall älter als ihre Tutees und wissen mehr über den Lerngegenstand. Somit ist Cross-Age-Tutoring üblicherweise durch feste Rollen gekennzeichnet. Deshalb wird das Non-Reciprocal-Cross-Age-Tutoring im Folgenden nur als Cross-Age-Tutoring bezeichnet. Diese drei Tutoring-Typen führen bei den Tutees im gleichen Maße zu Wissenszuwächsen bezüglich der thematisierten Inhalte (Leung, 2015).

Neben dem Tutoring-Typen können sich Tutoring-Angebote auch hinsichtlich der thematisierten Inhalte unterscheiden. Tutoring bezüglich sportlicher und künstlerischer Inhalte hat mit $d_{Cohen} = 0.90, p < .001$ und $d_{Cohen} = 0.82, p < .001$ die größten Effekte (J. Cohen, 1988). Tutoring bezüglich naturwissenschaftlicher beziehungsweise mathematischer Inhalte oder die Förderung von Lesefähigkeiten weisen mit $d_{Cohen} = 0.45, p < .001$, $d_{Cohen} = 0.34, p < .001$ und $d_{Cohen} = 0.15, p < .01$ kleinere Effektstärken auf. Tutoren und Tutees profitieren bezüglich dieser Themenbereiche von einer Teilnahme am Tutoring im Vergleich zu nicht Teilnehmenden. Den Befunden entsprechend könnten motorische oder prozedurale Fähigkeiten, die bei künstlerischen und sportlichen Tätigkeiten relevant

sind, durch ein Tutoring möglicherweise besser trainiert werden als kognitive Fähigkeiten, die für Problemlöseaufgaben in Naturwissenschaften, Mathematik und zum Trainieren der Lesefähigkeiten nötig sind (Leung, 2015). Dass in den Naturwissenschaften neben kognitiven Fähigkeiten auch motorische oder prozedurale Experimentierfähigkeiten relevant sind, könnte die höheren Effektstärken erklären. Eine dritte Kategorie, bezüglich der sich Tutoring-Angebote unterscheiden können, ist die Ausbildung der Tutoren. Einige Tutoring-Angebote setzen unausgebildete Tutoren ein, andere bereiten die Tutoren primär auf die zu betreuenden Inhalte vor und wieder andere bringen den Tutoren ein bestimmtes pädagogisches oder psychologisches Betreuungskonzept bei (Leung, 2015; Topping, 2005). Die Teilnahme der Tutoren an einem inhaltlichen Training ist irrelevant für die Wissenszuwächse der Tutees bezüglich der thematisierten Inhalte (Leung, 2015). Jedoch scheint ein solches Training erwartungskonform zu einem Wissenszuwachs bei den Tutoren zu führen (Berger et al., 2017; Korner, 2015).

Die vierte Unterscheidungskategorie von Tutoring-Angeboten bezieht sich auf die betrachteten Schulstufen der Tutoren und Tutees. Das Tutoring ist für Tutees aus der Sekundarstufe I und II mit $d_{Cohen} = 0.52, p < .001$ effektiver als für Tutees aus der Grundschule mit $d_{Cohen} = 0.34, p < .001$ oder der Universität mit $d_{Cohen} = 0.43, p < .001$. Weiterhin kann der Metaanalyse entnommen werden, dass bei über 60% aller berücksichtigten Studien die Tutoren bzw. Tutees aus der Grundschule stammen. Weniger als 20% der Studien untersuchen Tutoren bzw. Tutees aus der Sekundarstufe I. Somit liegen bezüglich Tutoren und Tutees aus der Sekundarstufe I und insbesondere für Tutoring-Angebote mit naturwissenschaftlichen Inhalten wenig belastbare Befunde vor. Als fünfte Kategorie kann die Geschlechterkomposition benannt werden, also ob die Tutoren und Tutees in gleichgeschlechtlichen Gruppen oder geschlechterübergreifend zusammenarbeiten. Gleichgeschlechtliche Gruppen scheinen mit einer Effektstärke $d_{Cohen} = 0.85, p < .001$ deutlich mehr das Wissen der Tutees zu fördern als geschlechterübergreifende Gruppen mit $d_{Cohen} = 0.36, p < .001$. Einerseits könnten sich Tutoren und Tutees in geschlechterübergreifenden Gruppen unwohl fühlen (Fogarty & Wang, 1982; Leung, 2015). Andererseits könnten vor allem weibliche Tutoren mit männlichen Tutees einen Konflikt zwischen traditionellen Geschlechterrollen und der natürlichen Autorität des Rollenverständnisses empfinden (Eagly et al., 2000; Leung, 2015). Ob dieser Effekt über die Motivation der Teilnehmenden mediiert ist, kann der Metaanalyse von Leung (2015) nicht entnommen werden.

Über die vielfältigen Gestaltungskategorien hinweg scheinen sowohl Tutoren als auch Tutees ihr Wissen bezüglich der im Tutoring thematisierten Inhalte durch eine Teilnahme im Vergleich zu nicht Teilnehmenden zu erhöhen (P. Cohen et al., 1982; Korner, 2015; Robinson et al., 2005; Topping et al., 2004). Die oben genannten Befunde geben an, in welchen Fällen Tutoring relativ größere Effektstärken hat. Entsprechend den fünf vorgestellten Kategorien können Laborhelferinnen und Laborhelfer als Cross-Age-Tutoren zur Betreuung physikalischer Experimente bezeichnet werden, welche ein Training erhalten haben und Schülerinnen oder Schüler der Sekundarstufe I sind. Sie können sowohl gleichgeschlechtlich als auch geschlechterübergreifend eingesetzt werden.

Entsprechend den oben genannten Befunden scheint der gewählte Tutoring-Typ (Cross-Age-Tutoring) irrelevant für den Wissenszuwachs der Tutees zu sein. Allerdings können Unterschiede bezüglich affektiver Variablen der Tutoren angenommen werden (Berger et al., 2017; Buchs et al., 2004; Ryan & Deci, 2017). Die Tutoren scheinen beim Cross-Age-Tutoring bezüglich ihrer Tutorentätigkeit stärker intrinsisch motiviert zu sein. Dies kann darauf zurückgeführt werden, dass sie sich beim Tutoring weniger in ihrer Kompetenz be-

droht fühlen (Buchs et al., 2004; Ryan & Deci, 2017). Im Folgenden wird dargelegt, weshalb der Alters- und Rollenunterschied der Tutoren und Tutees zu diesem Befund führt. Aufgrund ihrer Rolle als Tutoren, erhalten sie vor ihrem Einsatz im Regelfall ein Training bezüglich der zu betreuenden Inhalte, sodass sie gegenüber ihren Tutees einen Wissensvorsprung haben (z. B. Berger et al., 2017). Durch ihren Altersunterschied von mehreren Jahren sind sie vermutlich in einer höheren Klassenstufe und haben die zu betreuenden oder verwandte Themen im Regelunterricht behandelt. Somit ist davon auszugehen, dass Cross-Age-Tutoren mehr über den Lerngegenstand wissen als ihre Tutees, weshalb eine Kompetenzbedrohung ausbleibt (Buchs et al., 2004). Insbesondere könnte dies durch die unterschiedlichen Rollenselbstverständnisse begünstigt werden. Beim Cross-Age-Tutoring nehmen Tutoren die höhergestellte Rolle der Expertin oder des Experten ein, beim Same-Age-Tutoring eher die Rolle einer gleichrangigen Helferin oder Helfers. Die Motivation von Cross-Age-Tutoren bezüglich naturwissenschaftlicher Tätigkeiten steigt sogar durch ihre Tätigkeit als Tutoren signifikant (Berger et al., 2017). Neben dem gewählten Tutoring-Typ scheint auch das Vorhandensein eines Tutorentrainings irrelevant für den Wissenszuwachs der Tutees zu sein. Erneut können Unterschiede bezüglich affektiver Variablen der Tutoren und Tutees angenommen werden (Berger et al., 2017; Hänze et al., 2018; Ryan & Deci, 2017). Berger et al. (2017) haben speziell die intrinsische Motivation der Tutoren und Tutees beim Cross-Age-Tutoring im Physikunterricht untersucht. Sie ließen vier Experimente aus dem Themenbereich der Elektrizitätslehre betreuen. Entsprechend der Lernziele sollten die Tutees die verwendeten Energiequellen benennen können und ermitteln, dass ein geschlossener Stromkreis die Voraussetzung für einen Stromfluss bei Batteriespannung ist. Berger et al. (2017) bildeten zwei verschiedene Interventions- und eine Kontrollgruppe von Tutoren aus. Alle drei Gruppen erhielten eine fachliche Einführung in die oben genannten Experimente des Tutorings. So sollte ein inhaltliches Verständnis der Tutoren gesichert und somit eine Kompetenzbedrohung vermieden werden (Berger et al., 2017; Hänze et al., 2018). Die Kontrollgruppe erhielt darüber hinaus kein weiteres Training. Die beiden Experimentalgruppen lernten drei Betreuungsregeln, nämlich (1) nur zu helfen, wenn Hilfe gebraucht wird, (2) vor jedem Eingriff drei Sekunden zu warten, sodass die Tutees ihre Fehler selbst erkennen und korrigieren können und (3) Selbsthilfe zu fördern, indem die Tutoren nur so wenig helfen, dass die Tutees möglichst viel selbst machen können. Die Autoren bezeichnen dieses Verhalten als Autonomieunterstützung (Berger et al., 2017). In der Ausbildung der Tutoren der zweiten Experimentalgruppe wurden zusätzlich Videosequenzen gezeigt. Diese Videosequenzen zeigen, wie die Betreuungsregeln von Tutoren umgesetzt wurden und wie dies zum erfolgreichen Handeln der Tutees führt (Berger et al., 2017). Der Einsatz von prototypischen Videosequenzen führte erwartungskonform zu einer signifikant höheren intrinsischen Motivation der Tutoren der zweiten Experimentalgruppe gegenüber der ersten Experimental- und der Kontrollgruppe (Berger et al., 2017). Außerdem zeigte sich, dass sich die Tutoren der zweiten Experimentalgruppe signifikant autonomieunterstützender verhielten als die Tutoren in der Kontrollgruppe. Dass keinerlei Unterschiede der ersten Experimentalgruppe zu den anderen beiden Gruppen gefunden werden konnten, begründen die Autoren mit den fehlenden prototypischen Videosequenzen. In einer weiteren Studie untersuchten Hänze et al. (2018) die Auswirkung der Autonomieunterstützung auf die intrinsische Motivation der Tutees. Die zwei Tutorengruppen durchliefen ein inhaltliches Training zu den vier Betreuungsexperimenten von Berger et al. (2017). Die Experimentalgruppe durchlief zusätzlich das oben beschriebene Training zu Autonomieunterstützung mit prototypischen Videosequenzen. Die Experimentalbedingung förderte, mediiert über die von den Tutees selbst initialisierte Aktivität in Abgrenzung zu durch den Tutor initialisierte Tuteetätigkeiten, die intrinsische Motivation der Tutees (Hänze et al.,

2018). Entsprechend den Studien von Berger et al. (2017) und Hänze et al. (2018) scheint eine Ausbildung zu Autonomieunterstützung und ein entsprechendes Verhalten der Tutoren förderlich für die Motivation der Tutoren und Tutees zu sein. Der Einsatz von Schülerinnen und Schülern als Tutoren kann für Schulen eine wichtige Stütze zur Realisierung naturwissenschaftlicher außerunterrichtlicher Angebote sein. Aufgrund der vorgestellten Befundlage ist allerdings unklar, inwiefern Cross-Age-Tutoren für die Betreuung physikalischer Experimente eine Ausbildung benötigen. Ein Großteil bisheriger Studien fokussiert auf Tutoring-Angebote, die mathematische Inhalte thematisieren oder Lesefähigkeiten fördern. Es existieren wenige Studien zu Tutoren für naturwissenschaftliche Inhalte. In Physik bzw. in den Naturwissenschaften spielt das experimentelle Arbeiten eine besondere Rolle. Hierbei sind neben dem kognitiven Wissen auch prozedurale Fähigkeiten (wie in Sport und Kunst) relevant. Gerade in diesem Bereich könnte ein Peer-Tutoring besonders effektiv sein (s.o.; Leung, 2015). Deshalb besteht insbesondere ein Bedarf an Studien, die Tutoring-Angebote mit einem Fokus auf naturwissenschaftliche Inhalte beforschen (Leung, 2015). Weiterhin liegen keine belastbaren Aussagen dazu vor, inwiefern Schülerinnen und Schüler der Sekundarstufe I ein autonomieunterstützendes Verhalten erlernen und umsetzen können und welche Effekte dadurch auf die Motivation der Tutoren bzw. Tutees der Sekundarstufe I erreicht werden können.

3 Selbstbestimmungstheorie

Ein wesentlicher Ausgangspunkt für diese Arbeit ist die Selbstbestimmungstheorie nach Ryan und Deci (2017). Handlungen, die auf keine Intention zurückgeführt werden können, werden als amotiviert bezeichnet, während intentionsgeleitete Handlungen motiviert genannt werden (Deci & Ryan, 1993; Ryan & Deci, 2017). Die Motivation intendierter Handlungen wird in fünf *Motivationsqualitäten* unterteilt. Diese sind auf einem Kontinuum angeordnet (Ryan & Deci, 2017). Das Kontinuum beschreibt vereinfacht ausgedrückt, wie sehr das Individuum bei einer intendierten Handlung den Eindruck hat, die Kontrolle über eigene Handlungen zu haben (*Autonomieerleben*), wie wirksam sich das Individuum bei der Handlung erlebt (*Kompetenzerleben*) und wie sozial eingebunden es sich dabei fühlt (*Soziale Eingebundenheit*). Diese drei Bedürfnisse nach Autonomieerleben, Kompetenzerleben und sozialer Eingebundenheit bezeichnen Ryan und Deci (2017) als die *psychologischen Grundbedürfnisse*. Je stärker diese Grundbedürfnisse befriedigt sind, desto stärker ist ein Individuum intrinsisch motiviert. Es erlebt sich als *selbstbestimmt*. Je stärker die Grundbedürfnisse frustriert werden, desto mehr sinkt die intrinsische Motivation für das verlangte Verhalten und desto extrinsisch motivierter oder amotivierter wird das Individuum. Es erlebt sich als *fremdbestimmt*. Im Folgenden werden nur die Bedürfnisse nach Autonomieerleben und Kompetenzerleben dargestellt, da die soziale Eingebundenheit wenig zur intrinsischen Lernmotivation beiträgt (Jang et al., 2009; siehe auch Kapitel 5.1).

Das Bedürfnis nach Autonomieerleben

Laut Reeve und Jang (2006) umfasst das Autonomieerleben im Sinne der Selbstbestimmungstheorie drei Aspekte. Einerseits soll sich das Individuum als Handlungszentrum empfinden, also den Ursprung einer Handlung bei sich selbst und nicht bei anderen Individuen oder bei vom Selbst trennbaren Zielen wahrnehmen (*Wahrgenommener Ursprung der Handlung*). Des Weiteren umfasst das Autonomieerleben die Wahrnehmung, eine Handlung freiwillig und ohne Druck ausüben zu können (*Volition*) sowie die wahrgenommene *Wahlfreiheit*, entscheiden zu können, was zu tun ist, wie es zu tun ist und ob es überhaupt getan wird (Reeve & Jang, 2006). Besonders hervorgehoben wird, dass Autonomieerleben im Rahmen der Selbstbestimmungstheorie von Selbstständigkeit im Sinne einer Unabhängigkeit von anderen Individuen abzugrenzen ist (Ryan & Deci, 2017). Hat beispielsweise ein Schüler beim Lösen einer Aufgabe ein Problem und benötigt Hilfe, kann ihm maximale Selbstständigkeit (Autonomie) abverlangt werden, indem ihm jegliche Hilfe versagt wird. Der Schüler empfindet jedoch kein Autonomieerleben im Sinne der Selbstbestimmungstheorie, da er auf die Aufgabe bezogen nicht selbstständig handeln kann (Reeve & Jang, 2006). Da er nicht handeln kann, hat er keine Wahlfreiheit hinsichtlich der Nutzung der Hilfe und kann sich nicht als Handlungszentrum erleben.

Nicht jede Form der Wahlfreiheit führt zu einer Befriedigung des Autonomieerlebens. Reeve et al. (2003) unterscheiden zwei Arten von Wahlfreiheit, erstens die Auswahl zwischen von außen vorgegebenen Möglichkeiten (z. B. *Möchtest du mit Aufgabe 1 oder Aufgabe 2 beginnen?*) und zweitens die Wahlfreiheit als fortlaufende Folge von selbst gewählten Handlungsentscheidungen (z. B. *Was möchtest du wie lange untersuchen? Welche Hilfe*

möchtest du nutzen?). Nur die zweite Art der Wahlfreiheit befriedigt das Bedürfnis nach Autonomieerleben. Die erste Art der Wahlfreiheit erhöht zwar die wahrgenommene Wahlfreiheit, nicht aber das Autonomieerleben oder dazu verwandte Konstrukte (Reeve et al., 2003). Für das Autonomieerleben muss die Wahlfreiheit auch für das Individuum nutzbar sein (Katz & Assor, 2007). Sind die Freiheiten zu umfangreich oder komplex, kann ein Individuum die infrage kommenden Handlungsmöglichkeiten weder überblicken noch zielorientiert nutzen. So ist das Individuum in seinen Handlungsmöglichkeiten gelähmt, sein Autonomieerleben wird frustriert und das Individuum wird demotiviert (Katz & Assor, 2007). Ist einer Schülerin oder einem Schüler das Unterrichtsziel aufgrund des Fehlens einer hinreichenden Struktur nicht klar, können die (tatsächlich) vielen Handlungsmöglichkeiten nicht genutzt werden. Folglich kann diese Schülerin oder der Schüler keine Handlung festlegen, was wiederum das Autonomieerleben frustriert (Cheon et al., 2020).

Das Bedürfnis nach Kompetenzerleben

In der Selbstbestimmungstheorie wird angenommen, dass alle Menschen den natürlichen Drang haben, ihre Umgebung zu verändern und sich kompetent zu fühlen (Ryan & Deci, 2017). Entsprechend ist das Kompetenzerleben synonym zu dem Bedürfnis nach Wirksamkeit (Deci & Ryan, 1993). Ryan und Deci (2017) heben den Selbstzweck des Grundbedürfnisses nach Kompetenzerleben hervor und grenzen dieses Bedürfnis von dem Wunsch ab, ein bestimmtes Ergebnis oder eine Belohnung zu erreichen. Somit kann ein Kompetenzerleben bei einer Handlung nur entstehen, wenn die Autonomie der Handlung gewährleistet ist (Ryan & Deci, 2017). Ist die Autonomie nicht gegeben, folgt die Handlung einem externen Zweck und dient somit nicht mehr dem natürlichen Bestreben nach Kompetenzerleben. Zweckgebundene Handlungen, die den individuellen Zielen und Werten entgegenstehen, befriedigen nicht das Kompetenzerleben im Sinne der Selbstbestimmungstheorie (Ryan & Deci, 2017). Handlungen, die das Kompetenzerleben befriedigen, werden aufgrund der tätigkeitsspezifischen oder der gegenstandsspezifischen Freude ausgeübt. Erfolge bei diesen Handlungen bestärken diese Freude. Das aus dem Erfolg resultierende Kompetenzerleben führt zu einer weiteren Ausführung der Tätigkeit. Externe Ziele, z. B. eine Belohnung, stehen diesem natürlichen Bestreben entgegen (Ryan & Deci, 2017). Das Konstrukt Kompetenzerleben ist in Anlehnung an die *effectance motivation* von White (1959) definiert (Ryan & Deci, 2017). Es kann nach Krapp und Ryan (2002) auch als *Selbstwirksamkeitserwartung* aufgefasst werden. Die Befriedigung des Kompetenzerlebens resultiert aus der individuellen Bewertung, ob eine Handlung gut ausgeführt wurde sowie dem sozialen Vergleich, ob das Individuum besser bzw. effektiver als andere gehandelt hat (Markland & Hardy, 1997). In Anlehnung an die obige Definition ist das Konstrukt durch die individuelle Wahrnehmung der Wirksamkeit und nicht durch das Erreichen externer Ziele oder das Zufriedenstellen anderer Personen operationalisiert.

Die Definition und Einordnung der Motivationsqualitäten

Basierend auf der Definition der psychologischen Grundbedürfnisse können die Motivationsqualitäten definiert und auf dem Kontinuum der Selbstbestimmung eingeordnet werden. Es gibt fünf Motivationsqualitäten, die intrinsische Motivation und die vier Verhaltensregulationen der extrinsischen Motivation: die integrierte, identifizierte, introjizierte und externale Regulation. Handlungen, deren Beweggründe nicht von der Handlung getrennt werden können, werden als *intrinsisch motivierte* Handlungen bezeichnet. Beispiele sind Handlungen, die aufgrund der Freude an der Tätigkeit oder dem Gegenstand ausgeführt werden. Intrinsisch motivierte Handlungen sind im Einklang mit den individuellen Werten und das Individuum kann sich immer als Handlungszentrum erleben. Eine intrinsisch motivierte Handlung dient nicht der Erfüllung eines extern gesetzten Ziels, sondern

dem natürlichen Bestreben, die Umwelt zu beeinflussen. Somit stellen intrinsisch motivierte Handlungen den Prototypen selbstbestimmten Handelns dar (Deci & Ryan, 1993). Neben der Amotivation, bei der keine Regulation stattfindet, und der intrinsischen Motivation, existiert die *extrinsische Motivation*. Extrinsisch motivierte Handlungen werden ausgeführt, um ein von der Handlung trennbares Ziel zu erreichen. *Integriert regulierte* Verhaltensweisen sind in das Selbstkonzept des Individuums mit den zugehörigen Werten, Zielen und Normen integriert. Aufgrund dieser Integration werden bestimmte Handlungen ausgeübt, um dem eigenen Selbstkonzept gerecht zu werden. Die Handlungen sind vom Ziel trennbar. Somit ist die Motivation extrinsisch. Dennoch ist die Handlung überwiegend selbstbestimmt, da die Handlungen an den Werten und Normen des Individuums orientiert sind und durch das Individuum selbst gesteuert werden. Somit ist die integriert regulierte extrinsische Motivation zwar weniger selbstbestimmt als die intrinsische Motivation, sie ist jedoch die am meisten selbstbestimmte extrinsische Motivation. Eine *identifiziert regulierte* Handlung wird vom Individuum als subjektiv wichtig erachtet. Deci und Ryan (1993) nennen als Beispiel einen Schüler, der für seine Abiturprüfung lernt, um ein subjektiv bedeutsam wahrgenommenes Universitätsstudium zu ermöglichen. Identifiziert regulierte Handlungen sind eher selbst- als fremdbestimmt motiviert, da das Individuum das Ziel selbst wählt. Dennoch sind identifiziert regulierte Handlungen weniger selbstbestimmt als integriert regulierte Handlungen, da ein Ziel verfolgt wird, welches nicht im Selbstkonzept des Individuums integriert ist. Somit kann das Kompetenzerleben nicht optimal befriedigt werden, da ein externes statt eines subjektiv natürlichen Ziels verfolgt wird. *Introjiert regulierte* Verhaltensweisen sind Handlungen, deren Ursprung ein vom Individuum selbst auferlegter Druck ist. Die Handlung wird nicht als subjektiv bedeutsam bewertet, aber als notwendig erachtet. Ein Beispiel ist eine Handlung aufgrund eines schlechten Gewissens. Der Unterschied zwischen der identifizierten und der introjierten Regulation liegt darin, ob das Handlungsziel als subjektiv wichtig oder nur als notwendig empfunden wird. Wird das Ziel nur als notwendig erachtet, passen die individuellen Werte und Normen nicht zum Ziel. Somit orientieren sich introjiert regulierte Handlungen an dem wahrgenommenen Druck, die notwendige Tätigkeit ausführen zu müssen, und nicht an den individuellen Werten und Normen. Folglich kann sich das Individuum weder als Handlungszentrum noch als natürlich wirksam erleben. Somit sind introjiert regulierte Handlungen eher fremd- als selbstbestimmt. Die am meisten fremdbestimmte extrinsische Motivation ist die *externale Regulation*. Handlungen zur Vermeidung von Bestrafungen und zum Erhalt von Belohnungen sind Beispiele external regulierten Verhaltens. Das Handlungsziel, also die Belohnung oder die Vermeidung einer Bestrafung, liegt außerhalb der Handlung und des Individuums. Was gewünschte Handlungen sind, entscheidet nicht das Individuum, sondern die Person oder Instanz, welche die Belohnung oder die Bestrafung vergibt. Bei external regulierten Handlungen kann das Individuum kaum Autonomie erleben, da es in Einklang mit dem externen Ziel und nicht mit den eigenen Werten handelt. Dieses Ziel passt per Definition nicht zu den eigenen Werten, da die Motivationsregulation andernfalls introjiert oder identifiziert wäre (Deci & Ryan, 1993). Das Individuum kann auch kaum Kompetenzerleben empfinden, da die bewertende Instanz nicht das Individuum ist. Zusammengefasst sind intrinsisch motivierte Handlungen im Einklang mit den Werten und Normen des Individuums, während diese Passung von der integrierten bis zur externalen Regulation hin abnimmt. Bei external regulierten Handlungen orientiert sich das Individuum gänzlich an externen Einflüssen und Vorgaben. Die intrinsische Motivation ist maximal selbstbestimmt. Die external regulierte extrinsische Motivation ist maximal fremdbestimmt.

Die Relevanz der psychologischen Grundbedürfnisse

Ohne die entsprechenden Zusammenhänge explizit anzugeben, nehmen Hänze et al. (2018) im Einklang mit der Selbstbestimmungstheorie an, dass die Befriedigung der Bedürfnisse nach Autonomie- und Kompetenzerleben der Tutees eng mit ihrer intrinsischen Motivation zusammenhängt. Diese Zusammenhänge wurden sowohl bei Schülerinnen und Schülern im Regelunterricht als auch bei Studierenden gezeigt (Jang et al., 2009; Reeve et al., 2003). Die Befriedigung der psychologischen Grundbedürfnisse und die damit verbundene Förderung einer selbstbestimmten Motivation ist im Bildungskontext kein Selbstzweck. Lernende, die von einer eher selbstbestimmten Lernmotivation berichtet haben, zeigen bessere schulische Leistungen, eine höhere Kreativität, ein höheres Durchhaltevermögen, ein höheres Wohlbefinden und ein höheres Selbstwertgefühl (Für einen Überblick siehe Reeve, 2002 und Reeve, 2009). Weiterhin bevorzugen solche Lernende Herausforderungen und beziehen aus Herausforderungen Freude (Reeve, 2002). Weitere Studien zeigen Ähnliches für Schülerinnen und Schüler bzw. Studentinnen und Studenten, deren psychologische Grundbedürfnisse befriedigt sind: Jang et al. (2016) konnten zeigen, dass Studierende mit befriedigten psychologischen Grundbedürfnissen über ein höheres konzeptuelles Verständnis berichten. Darüber hinaus führt die Befriedigung der psychologischen Grundbedürfnisse von Schülerinnen und Schülern zu besseren Schulnoten (Jang et al., 2009). Cheon et al. (2012) konnten in einer experimentellen Studie zeigen, dass die Befriedigung der psychologischen Grundbedürfnisse zu einem höheren Engagement und einem höheren Lernerfolg führt. All diese Befunde lassen sich zwar vermutlich im Einklang mit der Selbstbestimmungstheorie auf Cross-Age-Tutoring-Angebote übertragen (Ryan & Deci, 2017), jedoch fehlen entsprechende Studien, um diese Übertragung empirisch abzusichern. Deshalb kann insgesamt ein Desiderat bei der Beforschung der Selbstbestimmungstheorie im Kontext von Cross-Age-Tutoring-Angeboten festgestellt werden. Aus den Desideraten zum Cross-Age-Tutoring folgt weiterhin, dass Cross-Age-Tutoring-Angebote beforscht werden sollten, bei denen physikalische Inhalte thematisiert werden und bei denen sowohl die Tutoren als auch die Tutees aus der Sekundarstufe I stammen (Siehe Kapitel 2).

4 Operationalisierung der Motivierstile

Aufgrund der oben dargestellten positiven Rolle der psychologischen Grundbedürfnisse im Bildungssektor haben viele Studien untersucht, ob und inwiefern die Befriedigung dieser Bedürfnisse in Lehr-Lern-Situationen gefördert werden kann (Für eine Übersicht siehe Reeve, 2009). Ähnlich wie Berger et al. (2017) und Hänze et al. (2018) fokussiert ein großer Teil dieser Studien auf das Verhalten der Lehrenden (Für eine Übersicht siehe Su und Reeve, 2011). Laut diesen Studien können Lehrende zu der Befriedigung und Frustration der psychologischen Grundbedürfnisse von Lernenden beitragen (Berger et al., 2017; Hänze et al., 2018; Jang et al., 2016; Jang et al., 2009; Reeve et al., 2003).

Die Tendenz von Lehrenden, das Autonomieerleben der Lernenden zu unterstützen, wird als *autonomieunterstützender Motivierstil* (autonomy-supportive motivating style) bezeichnet (Reeve et al., 2014). Entsprechend wird die Tendenz der Lehrenden zur Frustration des Autonomieerlebens der Lernenden als *kontrollierender Motivierstil* (controlling motivating style) bezeichnet (Reeve et al., 2014). Ursprünglich wurden diese Tendenzen als gegensätzliche Pole eines eindimensionalen Motivierstils definiert (Deci et al., 1981). Dieser eindimensionale Motivierstil wird in Abgrenzung zur separaten Betrachtung des autonomieunterstützenden und kontrollierenden Motivierstils als *bilanzierter Motivierstil* (net motivating style) bezeichnet und aus der Differenz der beiden berechnet (Reeve et al., 2014). Im Rahmen dieser eindimensionalen Definition wurde angenommen, dass Autonomieunterstützung das Gegenteil von Kontrolle ist. Dies ist nicht der Fall, da der autonomieunterstützende und der kontrollierende Motivierstil je nach Studie gar nicht, schwach oder höchstens mittelmäßig negativ miteinander korrelieren (Aelterman et al., 2018; Reeve et al., 1999; Reeve et al., 2014). Weiterhin führt ein niedrig ausgeprägter Motivierstil nicht dazu, dass der entsprechend andere Motivierstil hoch ausgeprägt ist (Reeve et al., 2018). Als Resultat der niedrigen Korrelationen präsentierte Reeve (2009) je eine Definition für den autonomieunterstützenden und kontrollierenden Motivierstil. Die Motivierstile wurden anschließend in einigen Studien getrennt ausgewertet (Aelterman et al., 2018; Reeve et al., 2014). Reeve et al. (2014) sowie Aelterman et al. (2018) konnten die Trennbarkeit der Motivierstile nachweisen. Deshalb werden im Folgenden der autonomieunterstützende und kontrollierende Motivierstil getrennt betrachtet.

Reeve (2009) definiert den autonomieunterstützenden Motivierstil von Lehrkräften als „interpersonal sentiment and behavior teachers provide during instruction to identify, nurture, and develop students' inner motivational resources“ (Reeve, 2009, S. 3) und den kontrollierenden Motivierstil als „interpersonal sentiment and behavior teachers provide during instruction to pressure students to think, feel, or behave in a specific way“ (Reeve, 2009, S. 3). Nach diesen Definitionen können bestimmte Verhaltensweisen als autonomieunterstützend oder kontrollierend bezeichnet werden. Zu autonomieunterstützendem Verhalten zählt, dass Lehrende die Perspektive und Gefühle ihrer Betreuten anerkennen, ihnen Informationen und Gelegenheiten für selbstbestimmte Handlungen geben, dabei möglichst wenig Zeit- und Handlungsdruck ausüben und eine informierende Sprache nutzen (Reeve & Jang, 2006). Der kontrollierende Motivierstil umfasst, geäußerte negative Gefühle zu

kontern, statt diese zu akzeptieren, Handlungs- und Zeitdruck auszuüben und eine kontrollierende Sprache zu verwenden (Reeve, 2009). In Anlehnung an Su und Reeve (2011) und Reeve (2009) können sechs Verhaltensgruppen beschrieben werden. Für jede Verhaltensgruppe werden typische autonomieunterstützende bzw. kontrollierende Verhaltensweisen benannt und, falls möglich, gegenübergestellt:

I Umgang mit Äußerungen negativer Gefühle

Betreute Schülerinnen und Schüler müssen sich an den Lernkontext (z. B. das gegebene Lernziel, Klassenkameraden oder das Betreuungsexperiment) anpassen und Aufgaben erledigen, die nicht immer den eigenen Zielen entsprechen oder Spaß machen (z. B. Klassenarbeiten oder Routineaufgaben). Folglich kann es zu Äußerungen negativer Emotionen kommen. Diese Emotionen resultieren einerseits aus mangelnder Freude am Lerngegenstand oder der Tätigkeit (Deci & Ryan, 1993). Weiterhin können diese Emotionen durch Langeweile (z. B. durch Unterforderung) oder Frustration (z. B. durch Überforderung) entstehen. Ein autonomieunterstützender Umgang mit diesen Äußerungen ist die Verbalisierung von Verständnis (z. B. *Das ist ja auch wirklich eine knifflige Aufgabe, . . .*) und Ermutigungen (*. . . aber du schaffst das bestimmt. Da bin ich mir sicher*). Wenn Betreuende ihre Betreuten ermutigen oder ihnen gegenüber Verständnis zeigen, wird den Betreuten vermittelt, dass ihre Gefühle akzeptiert werden. Die Betreuenden können auch ermutigend auf die Betreuten eingehen (*Ja, das kannst du gern ausprobieren*). Dies fördert das Autonomieerleben, da den Schülerinnen und Schülern gezeigt wird, dass ihre Empfindungen und nicht die Interessen der Betreuenden im Zentrum stehen. Bestimmte Arten von Lob sind ebenfalls eine autonomieunterstützende Maßnahme, beispielsweise Lob für Fortschritt und Erfolg (Reeve & Jang, 2006). Lob für Fortschritt und Erfolg beim Beschreiten selbstbestimmter Wege unterstützt das Autonomie- und Kompetenzerleben (z. B. *Ich kann selbstbestimmt erfolgreich handeln*). Diese spezifische Art des Lobs konnte als sehr autonomieunterstützend identifiziert werden (Reeve & Jang, 2006). Dagegen ist ein kontrollierender Umgang mit Äußerungen negativer Gefühle, das Kontern oder die Relativierung der Äußerungen (z. B. *Ach, die Aufgabe ist eigentlich ganz einfach*). Diese Aussage wird als kontrollierend wahrgenommen, weil die subjektiv von den Betreuten empfundene Schwierigkeit durch die normative Feststellung der Betreuenden abgewertet wird. Gegebenenfalls nutzen Betreuende auch Ihre rolleninhärente Macht, um etwaige Beschwerden zu übergehen oder zu unterdrücken (z. B. *Stell dich nicht so an und fang an zu arbeiten*) (Reeve, 2009). Das Übergehen oder Kontern von negativen Äußerungen der Schülerinnen und Schüler führt zu dem Eindruck, dass sie keine eigene Meinung haben dürfen und somit auch nicht selbstbestimmt handeln können oder dürfen.

II Freiräume und Hinweise statt vorgegebene Lösungs- und Handlungsschritte

Würden Schülerinnen und Schüler ihre Lernziele, Methoden und die Wege zur Erarbeitung der Lernziele vollständig selbst wählen und dabei auf keine inhaltlichen Schwierigkeiten stoßen, müssten Lehrerinnen und Lehrer nicht aktiv werden. Der Regelfall ist jedoch, dass Lehrende Lernziele an Bildungsstandards und Lehrplänen orientieren und ihre Schülerinnen und Schüler bei der Erarbeitung unterstützen müssen. Innerhalb vorgegebener Rahmenbedingungen fördern Betreuende das Autonomieerleben ihrer Betreuten, indem sie ihnen Handlungs- und Wahlmöglichkeiten eröffnen (Su & Reeve, 2011). Wie in Kapitel 3 erläutert, muss der Grad an Freiräumen an den jeweiligen Bedarf der Betreuten angepasst sein (Berger et al., 2017; Vansteen-

kiste et al., 2012; Wittwer & Renkl, 2008). Ein Beispiel ist eine insgesamt zeitlich begrenzte, nach Aufgabenschwierigkeit differenzierte Übungsphase, in der die Schülerinnen und Schüler die Übungsaufgabe nach ihrem entsprechenden Leistungsstand selbst wählen können, bei Bedarf zwischen den Aufgaben wechseln dürfen und die Bearbeitungsdauer einzelner Aufgaben flexibel festlegen können (Reeve et al., 2003). Autonomieunterstützung umfasst auch, dass Betreuende den Betreuten die nötigen Informationen geben, sodass sie selbstständig handeln können (Reeve & Jang, 2006). Beispiele sind Inputphasen, in denen inhaltliche Konzepte eingeführt werden. Weiterhin umfasst das autonomieunterstützende Bereitstellen nötiger Informationen auch das Geben von Hinweisen, wenn die Lernenden Schwierigkeiten mit der Bearbeitung einer Aufgabe haben. Betreuende sollten aber nur so viele Informationen geben, wie die Betreuten benötigen (Reeve & Jang, 2006). Dies kann auch durch an den Leistungsstand der Betreuten angepasste Scaffolds realisiert werden (Reeve, 2009). Je nach Situation kann ein allgemeiner Hinweis (*Was ist das Ziel der Aufgabe?*), ein richtungsweisender Hinweis (*Es könnte dir helfen, wenn du deinen Aufbau überprüfst*) oder in letzter Konsequenz ein konkreter Hinweis (*Du hast noch keine Batterie eingebaut*) angebracht sein (Berger et al., 2017; Reeve & Jang, 2006). Angemessene Hinweise ermöglichen den Betreuten zu handeln und verhindern, dass sie aufgrund mangelnder Informationen handlungsunfähig sind. Ein an den Stand der Betreuten angepasster Hinweis sorgt dafür, dass die Aufgabe nicht durch die betreuende Person gelöst wird, sondern dass die Betreuten selbstständig weiterarbeiten können. Die Betreuten können sich als Handlungszentrum und als fähig wahrnehmen, was sowohl das Autonomie- als auch das Kompetenzerleben erhöht (Hänze et al., 2018; Reeve, 2009).

Das Autonomie- und Kompetenzerleben von Betreuten kann durch unerwünschte Hinweise, vorgegebene bzw. demonstrierte Lösungen oder festgelegte nächste Handlungsschritte frustriert werden (Reeve, 2009; Reeve & Jang, 2006). Die Frustration resultiert aus dem Gefühl, nicht selbstbestimmt handeln zu können, da die Handlung extern festgelegt oder sogar vom Betreuenden übernommen wird (Reeve, 2009). Schülerexperimente mit unbegründeten Auflistungen von Handlungsschritten sind ein Beispiel für kontrollierend gestaltete Experimentierphasen.

III Geduld statt Ungeduld für Arbeiten im selbstbestimmten Tempo

Ist den Schülerinnen und Schülern das Unterrichts-, Aufgaben- oder Experimentierziel klar, kann ihr Autonomieerleben unterstützt werden, indem sie nur dann Hilfe bekommen, wenn sie darum bitten (Berger et al., 2017). Auf diese Art werden keine selbstbestimmten Handlungen unterbrochen oder gestört und es wird sichergestellt, dass die Betreuten auf das Ziel bezogen handeln können und sich somit als Handlungszentren wahrnehmen. Greifen Betreuende ein, weil ihnen die Betreuten zu langsam arbeiten, wird die selbstbestimmte Lerngelegenheit für die Betreuten unterbrochen. Diese Unterbrechung der selbstbestimmten Handlung frustriert das Autonomieerleben (Reeve & Jang, 2006). Anstatt ungewollt oder unnötig inhaltlich einzugreifen, können Betreuende durch Ermutigungen und Verständnis für Schwierigkeiten unterstützen (siehe vorhergehende Abschnitte; Reeve, 2009; Reeve und Jang, 2006).

IV Verwenden einer autonomieunterstützenden oder kontrollierenden Sprache

Eine druckerzeugende oder auffordernde Aussage kann als kontrollierend wahrgenommen werden, „because they functionally interrupt or stop students’ autonomous self-regulation to replace it with external regulation [...], whereas autonomy-supportive

communications are autonomy supportive because they functionally support and further students' autonomous self-regulation“ (Reeve, 2009, S. 170). Beispiele für autonomieunterstützende Kommunikation sind als Vorschläge formulierte Hinweise und z. B. das Verbalisieren von Verständnis und Ermutigungen. Beispiele für kontrollierende Kommunikation sind Aufforderungen (*Fang an!, Mach als nächstes damit weiter!*). Weitere Indikatoren kontrollierender Kommunikation sind Modalverben, die einen externen Ursprung der Handlung implizieren (*Du musst/ sollst als nächstes damit weitermachen*) und lenkende Fragen (*Magst/Kannst du mit der nächsten Aufgabe weitermachen?*), die ebenfalls kontrollierend wahrgenommen werden (Reeve & Jang, 2006).

V **Fördern intrinsischer motivationaler Ressourcen oder Nutzen Quellen extrinsischer Motivation**

Wie in Kapitel 3 erläutert, können Betreuende extrinsisch durch Anreize oder intrinsisch durch die Befriedigung der psychologischen Grundbedürfnisse motiviert werden. Die Grundbedürfnisse der Betreuten werden durch selbstbestimmtes Handeln befriedigt, indem sie die Lernziele selbst festlegen oder bei der Auswahl mitbestimmen dürfen (Autonomieerleben), ihnen herausfordernde Aufgaben präsentiert werden, die sie bewältigen können (Kompetenzerleben), oder sie ihr Arbeitsumfeld (Sozialform, Auswahl der Partner oder Gruppenmitglieder) mitbestimmen dürfen (soziale Eingebundenheit) (Ryan & Deci, 2017). Durch extrinsische Motivation (z. B. Notendruck, Belohnung für erwünschtes Verhalten oder Verweis auf Autoritäten) können Betreuende die Grundbedürfnisse dagegen frustrieren (Reeve, 2009).

VI **Identifizieren der Relevanz des Lerngegenstands und von Routineaufgaben für die Lernenden**

Neben dem selbstbestimmten Handeln müssen sich Schülerinnen und Schüler auch an äußere Bedingungen (z. B. inhaltliche Vorgaben) anpassen und unattraktive Aufgaben (Routineaufgaben, Hausaufgaben, Klassenarbeiten) erledigen (Reeve, 2009). Lehrerinnen und Lehrer können mit den Schülerinnen und Schülern den Mehrwert bzw. Sinn der Aufgaben besprechen, um Akzeptanz zu erzielen. Dadurch kann das gegebene externe Ziel internalisiert und so ein annähernd selbstbestimmtes Handeln der Lernenden ermöglicht werden (Reeve, 2009). Versäumen Lehrerinnen und Lehrer, die Bedeutung und Relevanz der Inhalte, Aufgaben und Ziele transparent zu machen, können die Schülerinnen und Schüler die gesetzten Ziele schwieriger nachvollziehen, nicht internalisieren und somit nicht selbstbestimmt handeln (Reeve, 2009).

Aus den dargestellten Verhaltensgruppen lassen sich konkrete Verhaltensweisen ableiten. Diese Verhaltensweisen wurden in den letzten Jahren untersucht, um sie als autonomieunterstützend oder kontrollierend zu bestätigen. Eine Übersicht von neunzehn Verhaltensweisen, welche in drei Studien untersucht wurden, ist in Tabelle 4.1 dargestellt. Beispielsweise untersuchten Reeve und Jang (2006) 72 Paare von Lehramtsstudierenden, die sich nacheinander gegenseitig bei der Bearbeitung einer dreidimensionalen Puzzleaufgabe namens Happy Cubes betreut haben. Die Verhaltensweisen der Betreuenden wurden eventbasiert kodiert und mit dem Autonomieerleben der Betreuten korreliert. Die Studie konnte im Einklang mit Reeve et al. (1999) und Deci et al. (1982) zeigen, dass das ungefragte Zeigen von Lösungen eine kontrollierende Verhaltensweise ist. Weiterhin bestätigen alle drei Studien, dass informatives Feedback, insbesondere Lob für Erfolg und Fortschritt eine autonomieunterstützende Verhaltensweise ist (Reeve & Jang, 2006).

Bis auf das Nennen von Fristen konnten alle Verhaltensweise durch mindestens eine Studie erwartungskonform als autonomieunterstützend beziehungsweise kontrollierend bestätigt werden. Für zehn der übrigen achtzehn Verhaltensweisen haben alle Studien, in denen sie berücksichtigt wurden, die gleichen Ergebnisse gefunden.

Allerdings sind nicht alle Verhaltensgruppen für jedes Cross-Age-Tutoring-Angebot möglich, sondern hängen von den Rahmenbedingungen ab. Kennen die Tutoren beispielsweise ihre Tutees vor der Betreuung nicht, können die Tutoren die Inhalte kaum an die Wünsche der Tutees anpassen und somit nur bedingt an die intrinsischen motivationalen Ressourcen der Tutees anknüpfen. Schließt sich an das Tutoring-Angebot kein Test an und dient das Tutoring-Angebot nicht der Routinebildung, müssen Tutoren nicht den Mehrwert bzw. Sinn der Aufgaben besprechen, um Akzeptanz für etwaige Langeweile bei Routineaufgaben zu erzielen. Je nach Gestaltung des konkreten Tutoring-Angebotes können alle sechs oder nur die Verhaltensgruppen I-IV im Tutorenttraining berücksichtigt werden.

Wie in Kapitel 2 im Detail dargestellt, wurde in der Studie von Berger et al. (2017) das Verhalten der Tutoren und die Befriedigung des Autonomieerlebens der Tutees erhoben. Die Tutoren der autonomieunterstützenden Versuchsbedingung ließen ihre Tutees mehr auf ihre Art und Weise arbeiten und halfen eher nur auf Nachfrage. Dieses Verhalten hing erwartungskonform mit der Befriedigung des Autonomieerlebens der Tutees zusammen $r = -.20, p < .05$. Somit konnten Berger et al. (2017) für Cross-Age-Tutoring-Angebote in der Physik insbesondere die Relevanz der Verhaltensgruppe II und III belegen.

	Reeve und Jang (2006)	Reeve et al. (1999)	Deci et al. (1982)	Bestätigung durch
Vermutlich autonomieunterstützende Verhaltensweisen				
Zuhörzeit	AS > C	AS > C	n/a	2 von 2; 100%
Zeit, die dem Schüler zum unabhängigen Arbeiten gegeben wurde	AS > C	AS > C	AS > C	3 von 3; 100%
Lob als informatives Feedback	AS > C	AS > C	AS > C	3 von 3; 100%
Ermutigung	AS > C	AS = C	n/a	1 von 3; 33,3%
Empathische, perspektivübernehmende Aussagen	AS > C	AS > C	n/a	2 von 2; 100%
Eingehen auf Schüleraussagen/ -fragen	AS > C	AS > C	n/a	2 von 2; 100%
Hinweis gegeben	AS > C	AS = C	AS = C	1 von 3; 33,3%
Nachfrage, was der Schüler möchte	AS > C	AS > C	AS = C	2 von 3; 66,6%
Subjektiv bedeutsamer Rahmen	AS > C	n/a	n/a	1 von 1; 100%
Vermutlich kontrollierende Verhaltensweisen				
Redezeit	AS = C	AS = C	AS < C	1 von 3; 33,3%
Zeit, in der das Instruktionsmaterial festgehalten wurde	AS < C	AS < C	n/a	2 von 2; 100%
Lösung gezeigt/ gesagt	AS < C	AS < C	AS < C	3 von 3; 100%
Aufforderungen/ Befehle	AS < C	AS < C	AS < C	3 von 3; 100%
„Sollst“-/ „Musst“-Aussagen	AS < C	AS = C	AS < C	2 von 3; 66,6%
Lenkende Fragen	AS < C	AS = C	AS < C	2 von 3; 66,6%
Nennung von Fristen	n/a	n/a	AS = C	0 von 1; 0%
Lösungen gegeben	AS < C	AS < C	AS < C	3 von 3; 100%
Kritik	AS = C	n/a	AS < C	1 von 2; 33,3%
Personenbezogenes Lob	AS = C	AS = C	AS < C	1 von 3; 33,3%

Legende

AS > C, autonomieunterstützend wahrgenommene Betreuende zeigen dieses Verhalten signifikant häufiger als kontrollierend wahrgenommene Betreuende

AS < C, kontrollierend wahrgenommene Betreuende zeigen dieses Verhalten signifikant häufiger als autonomieunterstützend wahrgenommene Betreuende

AS = C, dieses Verhalten wird von autonomieunterstützend wahrgenommenen Betreuenden weder signifikant häufiger noch signifikant seltener als von kontrollierend wahrgenommenen Betreuenden gezeigt;

n/a, dieses Verhalten wurde in der Studie nicht erhoben;

Tabelle 4.1: 19 Verhaltensweisen, die autonomieunterstützende oder kontrollierende Betreuende zeigen; Tabelle in Anlehnung an Reeve (2002) und Reeve und Jang (2006)

5 Die Einflüsse der Motivierstile

Im Folgenden werden die Einflüsse und die Wirkweise der Motivierstile von Betreuenden auf ihre Betreuten dargestellt. Anschließend werden verschiedene Konstrukte vorgestellt, welche die Motivierstile von Betreuenden möglicherweise vorhersagen können.

5.1 Wirkungsweise der Motivierstile der Betreuenden auf ihre Betreuten

Reeve (2009) hat die Ergebnisse von über 40 Studien und damit die Vorzüge von Autonomieunterstützung und die Nachteile von Kontrolle zusammengefasst. Ein ausgeprägter autonomieunterstützender und wenig ausgeprägter kontrollierender Motivierstil der Betreuenden führt bei den Betreuten zu einer höheren intrinsischen Motivation, einem höheren Engagement, einem höheren Flowerleben und einer höheren Performanz sowie einem höheren psychologischen Wohlbefinden (Für einen Überblick siehe Reeve, 2009; siehe auch Hofferber et al., 2016). Verschiedene autonomieunterstützende Verhaltensweisen von Betreuenden hängen mit der Befriedigung des Autonomieerlebens (Korrelationen zwischen $r_{min} = .23, p < .05$ und $r_{max} = .42, p < .01$) der Betreuten zusammen, welche mit einer höheren intrinsischen Motivation bezüglich der Experimentiertätigkeit ($r = .57, p < .01$), einem höheren Engagement ($r = .56, p < .01$) und besseren Leistungen ($r = .45, p < .01$; Reeve und Jang, 2006) zusammenhängt. Ebenso zeigen sich entsprechende negative Einflüsse, wenn die Betreuenden einen ausgeprägten kontrollierenden Motivierstil besitzen: Beispielsweise reagieren Schülerinnen und Schüler auf einen kontrollierenden Motivierstil ihrer Lehrkräfte mit einer erhöhten Konzentration des Stresshormons Cortisol ($F(2, 74) = 17.85, p < .01, \eta^2 = .33$; Reeve und Tseng, 2011). Insgesamt zeigen sich bei den Motivierstilen von Betreuenden erwartungskonform positive bzw. negative Einflüsse auf leistungsbezogene und affektive Variablen der Betreuten. Aus den oben genannten Befunden lässt sich nicht ableiten, ob die Motivierstile direkt oder indirekt auf die abhängigen Variablen wirken. Mehrere Studien bestätigen, dass die Motivierstile vollständig über die Befriedigung und Frustration der psychologischen Grundbedürfnisse mediiert werden (Hein et al., 2015; Jang et al., 2016b; Jang et al., 2016; Jang et al., 2009). Laut einem Strukturgleichungsmodell von Jang et al. (2009) werden die Schulnoten, das Engagement und die intrinsische Motivation von Schülerinnen und Schülern vom autonomieunterstützenden Motivierstil ihrer Lehrkräfte mediiert über die Befriedigung der zwei psychologischen Grundbedürfnisse Autonomieerleben und Kompetenzerleben beeinflusst („ $\chi^2(113, N = 256) = 246.74, p < .01, SRMR = .055, RMSEA = .068, CFI = .95$ “; Jang et al., 2009, S. 651). Die Befriedigung der sozialen Eingebundenheit scheint nachrangig zu sein (Jang et al., 2009). Das Strukturgleichungsmodell wurde in einer anderen Stichprobe mit vergleichbaren Kennwerten bestätigt (Jang et al., 2009). Jang et al. (2016b) konnten zeigen, dass die beiden Motivierstile über unterschiedliche Konstrukte mediiert wirken: Der autonomieunterstützende Motivierstil der untersuchten Dozenten sagt die Befriedigung der

psychologischen Grundbedürfnisse der Studierenden und diese wiederum das Engagement der Studierenden über drei Messzeitpunkte eines Semesters vorher. Der kontrollierende Motivierstil der Dozenten sagt die Frustration der psychologischen Grundbedürfnisse der Studierenden und diese wiederum fehlendes Engagement vorher („ $\chi^2(2787) = 1046.18, p = .999, RMSEA(90\%CI) = .000(.000, .000), SRMR = .054, CFI = 1.00, NNFI = 1.00$ “; Jang et al., 2016b, S. 33). Dabei sind die Befriedigung der psychologischen Grundbedürfnisse und deren Frustration statistisch getrennte Konstrukte. Insbesondere kann dem Modell von Jang et al. (2016b) entnommen werden, dass der kontrollierende Motivierstil weder die Befriedigung der Grundbedürfnisse noch das Engagement negativ beeinflusst. Ebenso beeinflusst der autonomieunterstützende Motivierstil weder die Frustration der Grundbedürfnisse noch das fehlende Engagement. Auf Basis der vorgestellten Studien und der angenommenen Universalität der Selbstbestimmungstheorie (Ryan & Deci, 2017) kann die Vermutung aufgestellt werden, dass alle oben genannten leistungsbezogenen und affektiven Variablen der Betreuten nicht direkt, sondern vollständig über die psychologischen Grundbedürfnisse von den Motivierstilen der Betreuenden beeinflusst werden (siehe Abbildung 5.1).

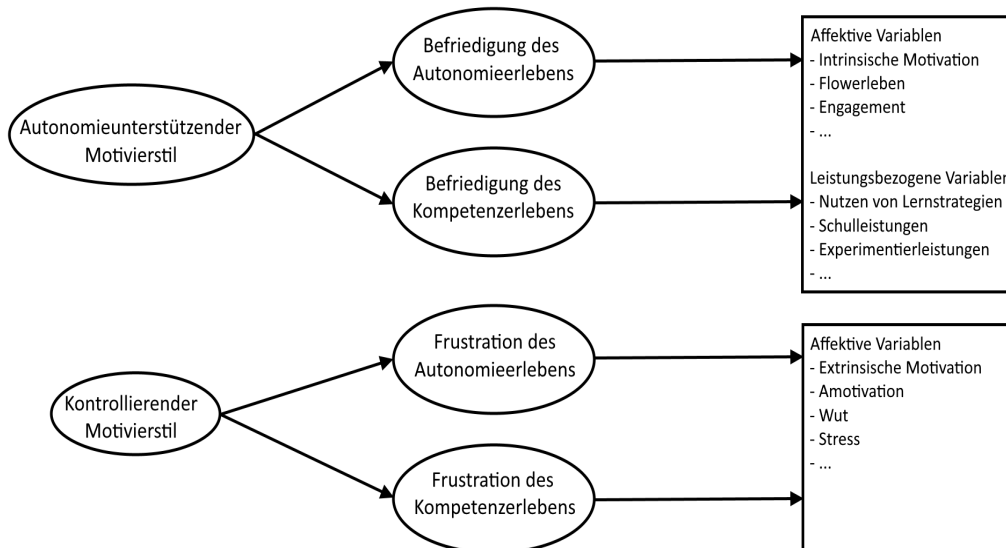


Abbildung 5.1: Vermutetes Wirkmodell der Motivierstile von Betreuenden auf leistungsbezogene und affektive Konstrukte von Betreuten

Entsprechend den vorgestellten Desideraten sind Studien zu Cross-Age-Tutoring-Angeboten bezüglich physikalischer Inhalte in der Sekundarstufe I nötig, welche die vermutete Übertragbarkeit der Befunde der Selbstbestimmungstheorie auf solche Angebote empirisch absichern (Siehe Kapitel 2 und 3). Laut Hänze et al. (2018) förderte die autonomieunterstützende Betreuung physikalischer Experimente bei Grundschülerinnen und Grundschülern, mediiert über die von den Tutees selbst initialisierte Aktivität, die intrinsische Motivation der Tutees (siehe auch Kapitel 2). Entsprechend der in der Einleitung formulierten Ziele soll in der vorliegenden Arbeit diese intrinsische Motivation von Tutees aus der Sekundarstufe I untersucht werden. Ob der autonomieunterstützende Motivierstil die Befriedigung der psychologischen Grundbedürfnisse vorhersagt und ob dieser Motivierstil ausschließlich über diese Grundbedürfnisse mediiert auf die intrinsische Motivation wirkt, ist für Cross-Age-Tutoren bisher nicht bestätigt. Laut Jang et al. (2009) ist die Befriedi-

gung des Kompetenzerlebens für die intrinsische Motivation relevanter als die Befriedigung des Autonomieerlebens. Aufbauend auf Hänze et al. (2018), Jang et al. (2009) sowie Jang et al. (2016b) wird in der vorliegenden Studie untersucht, ob der autonomieunterstützende Motivierstil der Tutoren auf die intrinsische Motivation von Tutees der Sekundarstufe I mediiert über die Befriedigung des Kompetenzerlebens wirkt (siehe Abbildung 5.2).



Abbildung 5.2: Vermutetes Wirkmodell des autonomieunterstützenden Motivierstils von Tutoren auf die intrinsische Motivation ihrer Tutees

5.2 Einflüsse auf die Motivierstile

Pelletier et al. (2002) konnten zeigen, dass die autonomieunterstützenden und kontrollierenden Motivierstile von den Betreuten beeinflusst werden. Dieser und weitere Einflüsse auf die Motivierstile werden im Folgenden vorgestellt.

Reeve (2009), Reeve et al. (2014) sowie Pelletier et al. (2002) argumentieren, dass die Motivierstile von vielen verschiedenen Einflüssen abhängen. Einerseits werden die Motivierstile von internen Faktoren beeinflusst, z. B. Überzeugungen und individuelle Dispositionen der Betreuenden (Reeve, 2009). Externe Faktoren, wie beispielsweise kulturelle Normen oder inhaltliche bzw. administrative Vorgaben, sowie das Verhalten der Betreuten können den Motivierstil der Betreuenden beeinflussen. Lehrerinnen und Lehrer neigen offenbar zu einem eher kontrollierenden Verhalten, wenn sie den Eindruck haben, dass ihre Schülerinnen und Schüler unmotiviert oder wenig engagiert sind (Reeve, 2009; Sarrazin et al., 2006). Die Motivierstile von Betreuenden hängen also von individuellen Charakteristika, Rahmenbedingungen und der Wahrnehmung der Betreuten ab. Diese Gründe lassen sich nach Grolnick und Apostoleris (2002), Pelletier et al. (2002), Reeve (2009) und Rocchi et al. (2013) in den Druck von Innen, den Druck von Unten und den Druck von Oben aufteilen.

5.2.1 Der Druck von Innen

Der Druck von Innen umfasst die Einflüsse von Persönlichkeitsmerkmalen (individuellen Dispositionen) der Betreuenden auf die Motivierstile, nämlich ihre Kausalitätsorientierungen, die Selbstbestimmtheit ihrer Arbeitsmotivation sowie ihre Überzeugungen, wobei Letztere im Rahmen des Projekts keine Berücksichtigung fanden und daher hier auch nicht vertieft betrachtet werden.

Kausalitätsorientierungen

Unterschiedliche Personen können dieselben Ereignisse und Situationen unterschiedlich autonomieunterstützend beziehungsweise kontrollierend wahrnehmen (Deci & Ryan, 1985). Dabei unterscheiden Deci und Ryan (1985) drei motivationale Ausprägungen, konkret die *autonome*, *kontrollierende* und *unpersönliche Kausalitätsorientierung*:

Personen mit einer hoch ausgeprägten *autonomen Kausalitätsorientierung* streben nach Handlungsspielräumen, um selbstbestimmt handeln zu können. Beispielsweise suchen sie

Berufe mit kreativen Anteilen oder Möglichkeiten zur Selbstregulation. Die selbstbestimmten Handlungen orientieren sich dabei an selbstgewählten Zielen und Tätigkeiten, die Freude bereiten. Derartige Personen erleben neue Ereignisse und Situationen als Möglichkeiten, selbstbestimmtes Verhalten zu initiieren und zu regulieren. Entsprechend erfahren sie in diesen Situationen ein hohes Maß an Selbstbestimmung und handeln überwiegend intrinsisch motiviert. Sie fühlen sich durch kontrollierendes Verhalten anderer (z. B. von Lehrkräften) weniger kontrolliert.

Eine hoch ausgeprägte *kontrollierende Kausalitätsorientierung* führt dazu, dass Individuen ihre Handlungen an Vorgaben und Grenzen orientieren. Deci und Ryan (1985) heben hervor, dass die Grenzen sowohl extern als auch intern sein können. Interne Grenzen können beispielsweise das Gefühl, eine Handlung machen zu müssen, oder eine selbstgesetzte Frist sein. Bei einer hoch ausgeprägten kontrollierenden Kausalitätsorientierung orientiert sich die Berufswahl eher am Gehalt oder an gesellschaftlichem Ansehen. Personen mit einer kontrollierenden Kausalitätsorientierung nehmen neuartige Situationen als Verunsicherung wahr und suchen nach Vorgaben und Grenzen zur Verhaltensregulation. Sie berichten über eine geringe Selbstbestimmung und handeln überwiegend fremdbestimmt motiviert.

Die *unpersönliche Kausalitätsorientierung* bedeutet, dass Individuen den Eindruck haben, ihr Verhalten nicht selbst kontrollieren und nicht zum gewünschten Ziel beitragen zu können. Individuen mit einer hoch ausgeprägten unpersönlichen Kausalitätsorientierung empfinden sich als inkompetent und fühlen sich von neuen Situationen überfordert. Die unpersönliche Kausalitätsorientierung wird durch negative Gefühle bezüglich aktueller Situationen, durch die Angst vor neuen Situationen und durch die unreflektierte Orientierung eigener Handlungen an Handlungen von Vorgängern operationalisiert (Deci & Ryan, 1985).

Die autonome und die kontrollierende Kausalitätsorientierung beeinflussen die Motivierstile von Lehrkräften und insbesondere ihr gezeigtes Unterrichtsverhalten (Reeve, 1998; Reeve et al., 1999; Reeve et al., 2018; van den Berghe et al., 2013). Deci und Ryan (1985) sowie Reeve et al. (2018) argumentieren, dass Individuen mit einer hoch ausgeprägten autonomen Kausalitätsorientierung zu einem autonomieunterstützenden Motivierstil neigen, weil sie die Relevanz von Selbstbestimmung aus ihrem eigenen Leben kennen und schätzen. Ist die kontrollierende Kausalitätsorientierung hoch ausgeprägt, neigen Individuen zu einem kontrollierenden Motivierstil (Reeve, 1998; Reeve et al., 1999; Reeve et al., 2018; van den Berghe et al., 2013). Eine mögliche Erklärung ist, dass sich solche Individuen im eigenen Leben an Grenzen und Vorgaben orientieren, Belohnungen und Vorgaben als Quellen motivierten Verhaltens wahrnehmen und deshalb anderen durch Grenzen und Vorgaben helfen wollen (Reeve et al., 2018). Es konnte kein Zusammenhang zwischen der unpersönlichen Kausalitätsorientierung und den Motivierstilen nachgewiesen werden (Deci & Ryan, 1985). Deshalb wird die unpersönliche Kausalitätsorientierung im Folgenden nicht weiter berücksichtigt.

Selbstbestimmtheit der Arbeitsmotivation

In Kapitel 3 wurden die verschiedenen Ausprägungen extrinsischer sowie die intrinsische Motivation vorgestellt und auf einem Kontinuum zwischen Selbstbestimmtheit und Fremdbestimmtheit angeordnet. Pelletier et al. (2002) konnten zeigen, dass Lehrkräfte mit einer relativ fremdbestimmten Arbeitsmotivation von einem relativ kontrollierenden und wenig autonomieunterstützenden Motivierstil berichten. Lehrkräfte mit einer relativ selbstbestimmten Arbeitsmotivation berichten von einem relativ autonomieunterstützenden und wenig kontrollierenden Motivierstil (Pelletier et al., 2002). Eine mögliche Erklärung für diesen Befund ist, dass Betreuende ihre eigenen motivationalen Ausprägungen auf ihre Betreuten übertragen. So könnten selbstbestimmt bzw. fremdbestimmt motivierte Betreuende

eine entsprechende Motivation fördern, da sie diese Motivation als natürlich empfinden.

Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung

Dickhäuser et al. (2002) definieren das Fähigkeitsselbstkonzept als „die Gesamtheit der kognitiven Repräsentationen eigener Fähigkeiten“ (Dickhäuser et al., 2002, S. 393). Fähigkeitsselbstkonzepte sind spezifisch, werden also bezogen auf bestimmte schulische Unterrichtsfächer, Aufgaben oder Tätigkeiten erfasst (Dickhäuser et al., 2002). Schwanzer et al. (2005) haben beispielsweise sechzehn verschiedene Selbstkonzepte untersucht, darunter tätigkeitsspezifische Fähigkeitsselbstkonzepte. Die Autoren konnten alle sechzehn Faktoren trennen und konvergente und divergente Validität der Faktoren untereinander und zu relevanten anderen Konstrukten nachweisen. Jedes Fähigkeitsselbstkonzept kann bezüglich verschiedener Bezugsnormen erfasst werden (Dickhäuser et al., 2002; Schwanzer et al., 2005). Einerseits können Fähigkeitsselbstkonzepte bezüglich einer individuellen, sozialen oder kriterialen Bezugsnorm untersucht werden (Dickhäuser et al., 2002). Andererseits können „auch absolute Fähigkeitsrepräsentationen erfasst werden, in denen keine Bezugsnorm (zumindest nicht explizit) verwendet wird“ (Dickhäuser et al., 2002, S. 395).

Kulgemeyer (2017) konnte einen Einfluss des tätigkeitsspezifischen Fähigkeitsselbstkonzeptes bezüglich des Physikerklärens von Studierenden auf deren Erklärperformanz nachweisen (siehe auch Kulgemeyer und Riese, 2018). Die Studierenden, die ein ausgeprägteres Fähigkeitsselbstkonzept bezüglich des Physikerklärens angegeben haben, zeigten eine höhere Performanz beim Physikerklären. In Anlehnung an Kulgemeyer (2017) sowie Kulgemeyer und Riese (2018) wird die Hypothese aufgestellt, dass das tätigkeitsspezifische Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung ebenfalls die Motivierstile beeinflusst. Es wird vermutet, dass ein hoch ausgeprägtes Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung positiv mit einem autonomieunterstützenden Motivierstil und negativ mit einem kontrollierenden Motivierstil zusammenhängt. Dies ist jedoch bisher nicht bestätigt und stellt eine offene Forschungsfrage dar.

5.2.2 Der Druck von Unten

Mit Druck von Unten ist der Einfluss der Betreuten auf den Motivierstil der Betreuenden gemeint. Die Wahrnehmung der Motivation, des Engagements und der Selbstständigkeit der Betreuten kann den Motivierstil als Druck von Unten beeinflussen (Pelletier et al., 2002; Reeve, 2009; Rocchi et al., 2013).

Pelletier und Vallerand (1996) zeigten den Einfluss der erwarteten Motivation der Betreuten auf die umgesetzte Autonomieunterstützung bzw. Kontrolle der Betreuenden. In der ersten Versuchsbedingung wurde den Betreuenden gesagt, dass die Betreuten die anstehende Aufgabe gerne bearbeiten und solche Rätsel als spannend und herausfordernd empfinden. In der zweiten Versuchsbedingung wurde den Betreuenden gesagt, dass die Betreuten die Aufgabe als langweilig empfinden und nur teilnehmen, weil sie dafür Geld erhalten. In der ersten Versuchsbedingung verhielten sich die Betreuenden autonomieunterstützender und weniger kontrollierend als in der zweiten Versuchsbedingung (Pelletier & Vallerand, 1996). Betreuende geben einen autonomieunterstützenderen Motivierstil an, wenn sie eher eine intrinsische Motivation ihrer Betreuten erwarten (Pelletier et al., 2002). Dabei ist nicht die tatsächliche Motivation der Betreuten ausschlaggebend, sondern nur die Erwartungshaltung der Betreuenden (Sarrazin et al., 2006). Reeve (2009) vermutet, dass Betreuende auch auf wenig engagiertes, unselbstständiges und unmotiviertes Verhalten der Betreuten unmittelbar mit einem eher kontrollierenden und wenig autonomieunter-

stützenden Verhalten reagieren. Da dieser Aspekt laut Reeve (2009) jedoch nur einmalig im Rahmen einer unveröffentlichten Dissertation untersucht wurde, besteht hier weiterer Forschungsbedarf.

Demnach müssen also zwei Ursachen des Drucks von Unten separat betrachtet werden. Einerseits werden durch vorhergehende Erfahrungen Erwartungshaltungen gebildet, die direkt oder indirekt die Motivierstile beeinflussen können. Andererseits kann auch der spontane Eindruck in der Betreuungssituation eine Veränderung im Motivierstil bewirken. Zur besseren Unterscheidung werden die beiden Arten des Drucks von Unten als Druck von Unten (Erwartungshaltung) und Druck von Unten (spontan) bezeichnet.

5.2.3 Der Druck von Oben

Der Druck von Oben beschreibt Einflüsse auf die Motivierstile, die nicht aus dem Individuum (Druck von Innen) oder aus der Wahrnehmung der betreuten Schülerinnen und Schüler (Druck von Unten) resultieren.

Deci et al. (1982) konnten zeigen, dass die Motivierstile der Betreuenden von dem Eindruck beeinflusst werden, für die Disziplin und Leistung der Betreuten verantwortlich zu sein. In der Studie lösten Schülerinnen und Schüler ein räumliches Puzzle. Bevor die Probanden den Schülerinnen und Schülern bei der Bearbeitung dieser Experimentieraufgabe geholfen haben, wurde ihnen in der Experimentalgruppe gesagt, dass es ihre Verantwortung sei, dass die Betreuten gute Leistungen erbringen. Den Probanden in der Kontrollgruppe wurde in der Einführung gesagt, dass es keine konkreten Leistungsanforderungen für die Betreuten gibt. Im Anschluss an diese Einleitung unterstützten die Probanden beider Gruppen jeweils einen Betreuten oder eine Betreute des gleichen Geschlechts bei der Bearbeitung der Experimentieraufgabe. Die Probanden der Interventionsgruppe zeigten bei der Betreuung ein signifikant kontrollierenderes Verhalten (Deci et al., 1982).

Reeve (2009) bezeichnet die Fehlvorstellung von Lehrerinnen und Lehrern, dass Kontrolle und Struktur Synonyme sind und die unzulässige Assoziation von Autonomieunterstützung mit einem Laissez-faire Stil bzw. Chaos, als eine weitere Form des Drucks von Oben. Nicht die Fehlvorstellung selbst ist der Druck von Oben, sondern die Kombination aus der Fehlvorstellung und der Erwartungshaltung der Schulleitung, der Eltern und der Kollegen, dass der Lehrende für eine Struktur und Ordnung zu sorgen hat. Wird Kontrolle mit Struktur verwechselt, kann bei Lehrenden der Eindruck entstehen, dass von ihnen ein kontrollierendes Verhalten erwartet wird. Tatsächlich korreliert ein autonomieunterstützendes Lehrerverhalten sehr hoch mit der von der Lehrkraft gewährleisteten Struktur (Aeltermann et al., 2018).

Ein dritter Aspekt des Drucks von Oben liegt vor, wenn die Lehrenden selbst kontrolliert werden, z. B. wenn sie durch eine Belohnung motiviert werden. So verhalten sich Tutoren, die eine Belohnung für Ihre Tätigkeit erhalten haben, kritischer und anspruchsvoller als Tutoren, die keine Belohnung erhalten haben (Garbarino, 1975).

In Abhängigkeit der konkreten Gestaltung einer Betreuungssituation kann der Druck von Oben unterschiedlich relevant für die Betreuenden sein. Werden die Betreuenden für die Leistungen und das Verhalten der Betreuten verantwortlich gemacht und zum Beispiel in Abhängigkeit dieser Leistungen unterschiedlich belohnt, führt dies zu einem hohen Druck von Oben und einem kontrollierenden Motivierstil (Reeve, 2009). Somit entfallen diese

Aspekte des Drucks von Oben für Cross-Age-Tutoren, wenn sie für ihre Tätigkeit nicht belohnt und nicht für die Leistung und Disziplin der Tutees verantwortlich gemacht werden.

Darüber hinaus betreuen Tutoren im Regelfall nicht ganze Klassen, sondern kleinere Gruppen. Dadurch sind die Betreuungssituationen weniger komplex. Cheon et al. (2020) nennen das klare Kommunizieren der Erwartungen, freiwillige gestufte Lernhilfen und Feedback als drei Aspekte autonomieunterstützender Struktur. Durch die weniger komplexe Betreuungssituation lassen sich diese drei Aspekte autonomieunterstützender Struktur einfacher umsetzen. Somit wirkt sich die eventuell vorhandene Fehlvorstellung weniger aus, dass mit kontrollierendem Verhalten eine bessere Struktur gewährleistet werden kann.

5.2.4 Das zu untersuchende Wirkmodell der Motivierstile und der individuellen Drücke

Die folgende Übersicht (siehe Tabelle 5.1) fasst die in den letzten Unterkapiteln vorgestellten Drücke auf Betreuende zusammen:

Druck von Innen	Druck von Unten	Druck von Oben
Autonome/ kontrollierende Kausalitätsorientierung	Wahrgenommene Motivation der Betreuten	Fehlvorstellung über den Zusammenhang von Struktur und Kontrolle
Selbstbestimmtheit der Arbeitsmotivation	Wahrgenommenes Engagement der Betreuten	Verantwortlichkeit für Leistung und Disziplin
Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung	Wahrgenommene Selbstständigkeit der Betreuten	Kontrolle (z. B. extrinsische Motivation) der Betreuenden durch Administrative

Tabelle 5.1: Übersicht über mögliche Einflüsse auf die Motivierstile von Betreuenden

Die Studien von Pelletier et al. (2002) sowie Rocchi et al. (2013) haben Wirkmodelle der Drücke auf Motivierstile untersucht. Deshalb werden im Folgenden beide Studien vorgestellt, um darauf aufbauend ein allgemeines Modell zur Wirkung der oben beschriebenen Drücke auf die Motivierstile abzuleiten. Anschließend wird dieses für Cross-Age-Tutoren konkretisiert.

Pelletier et al. (2002) haben für Lehrkräfte ein Wirkmodell untersucht. Dabei wurde die wahrgenommene Selbstbestimmtheit der Schülermotivationen als Druck von Unten (Erwartungshaltung) erfasst. Durch vier Subskalen wurden die intrinsische, die identifizierte, die introjizierte und die extrinsische Motivation erhoben. Diese vier Motivationen konnten faktorenanalytisch getrennt werden und wurden gewichtet zur Selbstbestimmtheit der Motivation zusammengefasst. Die anderen Aspekte des Drucks von Unten (Erwartungshaltung) wurden nicht erfasst. Weiterhin wurde der Druck von Unten (spontan) nicht erfasst. Pelletier et al. (2002) haben als Druck von Innen lediglich die Selbstbestimmtheit der Arbeitsmotivation erfasst, also die bilanzierte intrinsische, identifizierte, introjizierte und

extrinsische Motivation. Es wurde durch drei Subskalen der Druck von Oben erfasst: Die erste Subskala hat den wahrgenommenen Druck der Lehrkräfte durch die übrigen Lehrkräfte gemessen. Die Items erfragten, inwiefern sich die Lehrkräfte an die Lehrmethoden von Kolleginnen und Kollegen anpassen müssen. Die zweite Subskala umfasste den wahrgenommenen Druck der Lehrkräfte durch die Schulleitung. Die Items erfragten, wie sehr die Lehrkraft für die Disziplin und Leistung ihrer Schülerinnen und Schüler verantwortlich gemacht wird. Die dritte Subskala hat den wahrgenommenen Druck der Lehrkräfte bezogen auf das Schulcurriculum erfasst und erfragte, inwiefern sie es als wichtig erachten, dieses Schulcurriculum vollständig zu erfüllen. Laut Pelletier et al. (2002) zeigte das am besten passende Modell einen Zusammenhang des Drucks von Unten (Erwartungshaltung) und des Drucks von Oben. Beide beeinflussten den Druck von Innen, welcher den bilanzierten Motivierstil beeinflusste. Weder der Druck von Unten noch der Druck von Oben beeinflussten den Motivierstil direkt. Das beschriebene Modell hatte die Fitwerte „ $\chi^2(87, N = 254) = 109.69, p < .01, GFI = .95, AGFI = .92, CFI = .96, IFI = .96, PGFI = .68$ “ (Pelletier et al., 2002, S. 192).

In der Studie von Rocchi et al. (2013) wurde ein ähnliches Wirkmodell untersucht. Die Daten der Sporttrainerinnen und Sporttrainer wurden durch eine Onlineumfrage erhoben. Der Druck von Unten wurde wie bei Pelletier et al. (2002) ausschließlich durch die Selbstbestimmtheit der wahrgenommenen Motivation der Betreuten erfasst. Im Gegensatz zu Pelletier et al. (2002) wurden nicht nur die intrinsische, identifizierte, introjizierte und extrinsische Motivation, sondern auch die integrierte Motivation und Amotivation erfasst. Alle sechs Subskalen wurden entsprechend der Verortung auf dem Kontinuum der Selbstbestimmtheit (siehe Kapitel 3) gewichtet und bilanziert. Der Druck von Oben wurde durch eine auf Sporttrainerinnen und Sporttrainer angepasste Version des Fragebogens von Pelletier et al. (2002) erfasst. Laut Rocchi et al. (2013) zeigte das am besten passende Modell einen Zusammenhang des Drucks von Unten (Erwartungshaltung) und des Drucks von Oben. Beide beeinflussten den Druck von Innen, welcher den autonomieunterstützenden Motivierstil beeinflusste. Der Druck von Oben beeinflusste den autonomieunterstützenden Motivierstil nicht direkt. Der Druck von Unten (Erwartungshaltung) beeinflusste den autonomieunterstützenden Motivierstil partiell mediert über den Druck von Innen. Das beschriebene Modell hatte die Fitwerte „ $\chi^2_{49} = 82.79, p < .001, SRMR = .067, IFI = .94, RMSEA = .048$ “ (Rocchi et al., 2013, S. 856).

In beiden Studien wurde der kontrollierende Motivierstil nicht separat betrachtet. Es wurden keine weiteren Aspekte des Drucks von Innen, beispielsweise die Kausalitätsorientierungen, berücksichtigt. Diese können konzeptuell dem Druck von Innen zugeordnet werden. Ob diese Konstrukte mit der Selbstbestimmtheit der Arbeitsmotivation ein gemeinsames latentes Konstrukt bilden, ist bisher nicht untersucht. Weiterhin berücksichtigen beide Modelle nicht den Druck von Unten (spontan). Da Tutoren ihre Tutees im Regelfall nicht kennen und somit keine Erwartungshaltung Ihnen gegenüber aufbauen konnten, könnte der Druck von Unten (spontan) eine stärkere Bedeutung haben. Aufbauend auf die Modelle von Pelletier et al. (2002) und Rocchi et al. (2013) wird folgendes Modell vermutet (siehe Abbildung 5.3). Argumente für die Wirkbeziehungen werden im Folgenden passend zur jeweiligen Ziffer dargelegt.

1-4: Sowohl in der Studie von Pelletier et al. (2002) als auch in der Studie von Rocchi et al. (2013) hängen der Druck von Oben und der Druck von Unten (Erwartungshaltung) zusammen. Der Druck von Innen, repräsentiert durch die Selbstbestimmtheit der Arbeitsmotivation der Betreuenden, hat in beiden Studien einen direkten Einfluss auf die Motivierstile. Der Druck von Oben wirkt sich in beiden Studien nicht direkt, sondern vollständig

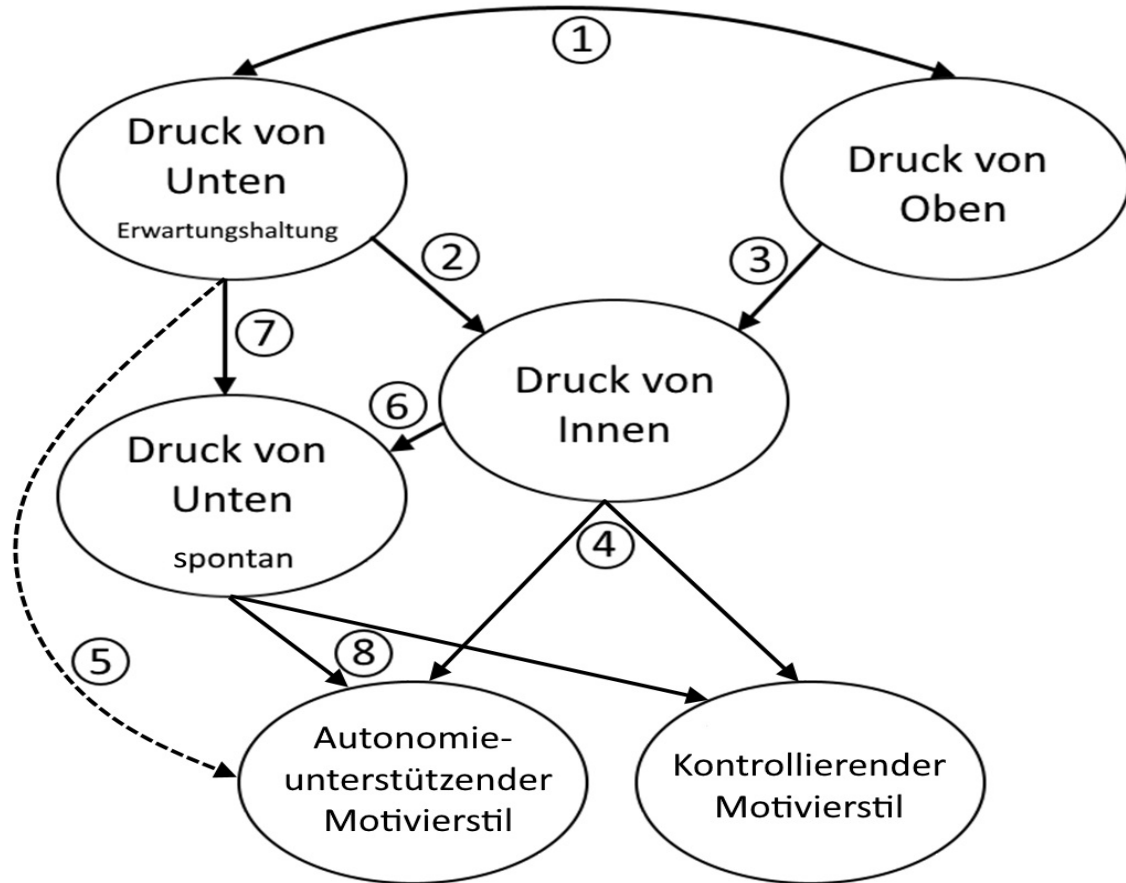


Abbildung 5.3: Allgemeines Wirkmodell des Drucks von Oben, des Drucks von Unten (Erwartungshaltung), des Drucks von Unten (spontan) und des Drucks von Innen auf die Motivierstile

mediert über den Druck von Innen auf die Motivierstile aus. Der Druck von Unten (Erwartungshaltung) sagt in beiden Studien den Druck von Innen voraus. Der Druck von Unten (Erwartungshaltung) und der Druck von Oben beeinflussen somit, wie selbstbestimmt motiviert und wie insgesamt motiviert die Betreuten vor der Betreuung ihrer Betreuten sind (Pelletier et al., 2002; Rocchi et al., 2013; Sarrazin et al., 2006). Empfinden die Betreuenden einen großen Druck von Oben oder erwarten, dass ihre Betreuten unmotiviert oder störend sein werden, sind sie insgesamt weniger motiviert und die übrige Motivation ist eher fremdbestimmt (Pelletier et al., 2002; Rocchi et al., 2013; Sarrazin et al., 2006). Dies führt wiederum zu einem eher kontrollierenden Verhalten (siehe Kapitel 5.2.1).

5: In der Studie von Rocchi et al. (2013) weist der Druck von Unten (Erwartungshaltung) einen partiell medierten Effekt auf den autonomieunterstützenden Motivierstil auf. In der Studie von Pelletier et al. (2002) wird der Druck von Unten (Erwartungshaltung) vollständig über den Druck von Innen mediert. Die unterschiedlichen Befunde könnten darauf zurückzuführen sein, dass Pelletier et al. (2002) den bilanzierten und Rocchi et al. (2013) den autonomieunterstützenden Motivierstil erhoben haben. Da der direkte Einfluss des Drucks von Unten auf die Motivierstile bei Rocchi et al. (2013) gefunden wurde und bei Pelletier et al. (2002) nicht, wird dieser gestrichelt dargestellt. Inhaltlich ausgedrückt haben Betreuende eine weniger intrinsische Motivation bezüglich ihrer Betreuungsaufgabe,

wenn sie erwarten, dass ihre Betreuten unmotiviert, unselbstständig oder wenig engagiert sind. Diese verringerte Selbstbestimmtheit der Betreuungsmotivation reduziert den autonomieunterstützenden Motivierstil und fördert den kontrollierenden Motivierstil.

6: Da Aspekte des Drucks von Innen vor einer konkreten Interaktion vorhanden sind und der Druck von Unten (spontan) erst während einer konkreten Interaktion entsteht, könnten die Aspekte des Drucks von Innen die Aspekte des Drucks von Unten (spontan) beeinflussen. Inhaltlich ausgedrückt: Die spontane Wahrnehmung der Betreuten durch die Betreuenden könnte durch deren Prädispositionen beeinflusst werden.

7: Da Aspekte des Drucks von Unten (Erwartungshaltung) vor einer konkreten Interaktion vorhanden sind und der Druck von Unten (spontan) erst während einer konkreten Interaktion entsteht, könnten die Aspekte des Drucks von Unten (Erwartungshaltung) die Aspekte des Drucks von Unten (spontan) beeinflussen. Sarrazin et al. (2006) argumentieren, dass die Erwartungshaltung von Lehrkräften über die Motivation der Schülerinnen und Schüler entsprechend dem Prinzip einer sich selbst erfüllenden Prophezeiung (Merton, 1948) bzw. entsprechend dem Pygmalion-Effekt (Rosenthal & Jacobson, 1968) die tatsächliche Motivation der Schülerinnen und Schüler und die Wahrnehmung der Lehrkräfte, also den Druck von Unten (spontan), beeinflussen kann. Inhaltlich ausgedrückt: Das spontan wahrgenommene Verhalten der Betreuten könnte durch das vom Betreuenden erwartete Verhalten der Betreuten mitbestimmt werden.

8: Reeve (2009) nimmt an, dass auch die spontanen Wahrnehmungen sowohl den autonomieunterstützenden als auch den kontrollierenden Motivierstil der Betreuenden beeinflussen können. Ähnlich wie Erwartungshaltungen zur Unmotiviertheit und Unselbstständigkeit (siehe Druck von Unten (Erwartungshaltung)) könnte auch das spontan wahrgenommene unselbstständige und unmotivierte Verhalten der Betreuten den autonomieunterstützenden Motivierstil reduzieren und den kontrollierenden fördern. In der vorliegenden Studie wird wie bei Berger et al. (2017) und Hänze et al. (2018) ein Cross-Age-Tutoring betrachtet, bei dem die Tutoren ihre Tutees vor der Betreuung nicht kennen, keine Belohnung für ihre Tätigkeit erhalten und nicht für die Leistung oder Disziplin ihrer Tutees verantwortlich gemacht werden. Dies gibt Anlass dazu, das Modell für diese Studie spezifisch zu vereinfachen. So fallen in diesem Setting mit der Belohnung und der Verantwortlichkeit wesentliche Aspekte des Drucks von Oben weg. Da der Druck von Oben zudem selbst bei Lehrkräften der am wenigsten relevante Druck ist (Pelletier et al., 2002), wird er im Modell nicht weiter berücksichtigt. Der Druck von Unten (Erwartungshaltung) kann in dem Modell vermutlich ebenfalls vernachlässigt werden, da die Betreuenden ihre Betreuten nicht kennen und keine Erwartungshaltungen bilden können. Abbildung 5.4 stellt das verbleibende, vermutete Wirkmodell für die betrachteten Cross-Age-Tutoren dar.

Bezüglich dieses Modells besteht mehrfacher Forschungsbedarf. Einerseits ist die Rolle des Drucks von Unten (spontan) weitestgehend ungeklärt. Es ist unklar, ob der Druck von Innen den Druck von Unten (spontan) beeinflusst und ob dieser Druck von Unten (spontan) die Motivierstile beeinflusst. Weiterhin ist ungeklärt, ob die verschiedenen Aspekte des Drucks von Innen beziehungsweise Unten (spontan) tatsächlich ein latentes Konstrukt ergeben oder als einzelne Prädiktoren auf die Motivierstile wirken, da Pelletier et al. (2002) und Rocchi et al. (2013) lediglich die Selbstbestimmtheit der Arbeitsmotivation als Druck von Innen erfasst haben. Darüber hinaus ist das Modell als solches für Cross-Age-Tutoren bisher nicht untersucht worden. Weder Pelletier et al. (2002) noch Rocchi et al. (2013) haben den kontrollierenden Motivierstil separat untersucht.

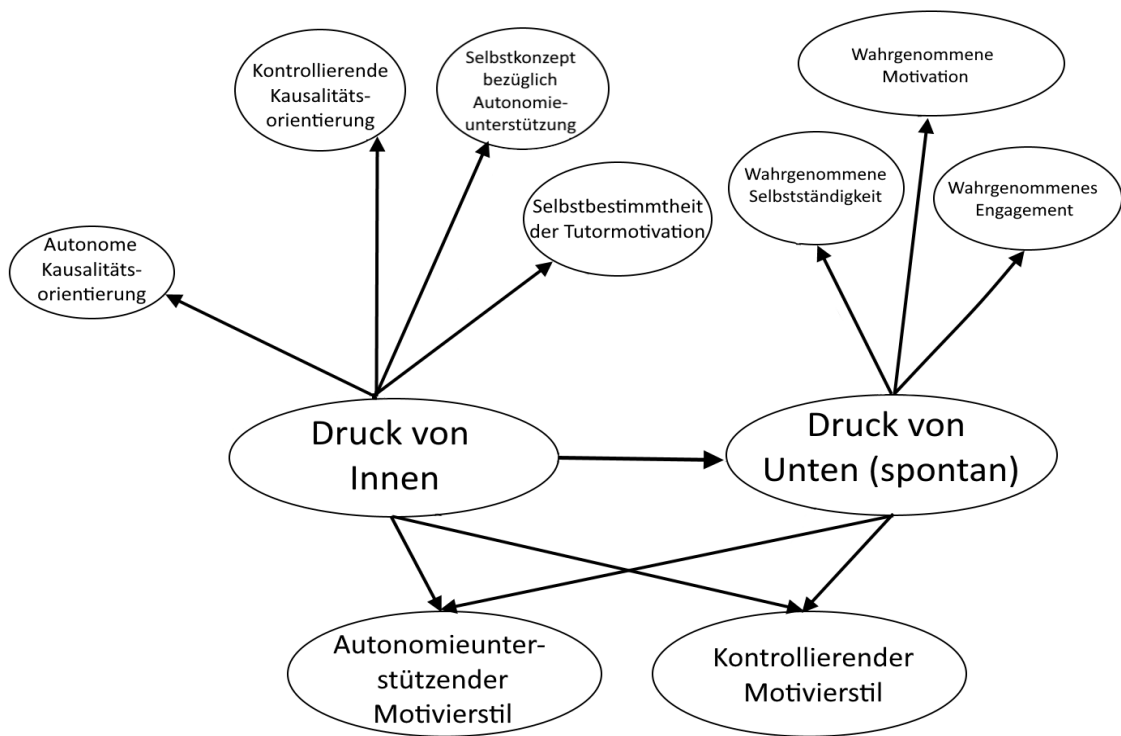


Abbildung 5.4: Auf Cross-Age-Tutoren angepasstes Wirkmodell des Drucks von Unten (spontan) und des Drucks von Innen auf die Motivierstile

6 Intervenierbarkeit der Motivierstile

Aufbauend auf den positiven Einflüssen von Autonomieunterstützung und den negativen Einflüssen von Kontrolle (siehe Kapitel 3 und 5) wurde im Rahmen einiger Studien untersucht, ob und inwiefern ein autonomieunterstützender Motivierstil gefördert und ein kontrollierender reduziert werden kann. Bevor die Veränderlichkeit diskutiert werden kann, werden die verschiedenen Perspektiven dargestellt, aus denen die Motivierstile erfasst werden können.

6.1 Perspektiven der Erfassung der Motivierstile

Entsprechend Reeve et al. (2014) und Su und Reeve (2011) können der autonomieunterstützende und der kontrollierende Motivierstil von Betreuenden aus drei Perspektiven erfasst werden. Erstens als Selbstberichte der Betreuenden, zweitens als Expertenbeobachtungen und –analysen tatsächlichen Verhaltens der Betreuenden in Betreuungssituationen und drittens durch selbstberichtete Wahrnehmung von Betreuten über ihre Betreuenden.

Grundsätzlich wird vermutet, dass die Maße der drei Perspektiven für autonomieunterstützende Motivierstile miteinander zusammenhängen und die Maße der drei Perspektiven für kontrollierende Motivierstile miteinander zusammenhängen. In der Studie von Reeve und Cheon (2016) hängen die Selbstberichte der Betreuenden und Betreuten in Höhe von $r = .40, p = .008$ zusammen. Reeve et al. (1999) haben eine Korrelation zwischen dem selbstberichteten und dem beobachteten bilanzierten Motivierstil in Höhe von $r = .42, p < .05$ gefunden. Laut Reeve und Jang (2006) korrelieren die verschiedenen autonomieunterstützenden Verhaltensweisen der Betreuenden mit dem Autonomieerleben der Betreuten zwischen $r = .23, p < .05$ und $r = .42, p < .01$. Zusammengenommen werden die drei Perspektiven der Motivierstile vermutlich in Höhe von etwa $r = .4$ zusammenhängen. Bezüglich der Zusammenhänge der drei Perspektiven zeigen sich allerdings inkonsistente Ergebnisse: Die Studie von Aelterman et al. (2014), in welcher die Zusammenhänge aller drei Perspektiven untersucht wurden, konnte keinen der drei Zusammenhänge bestätigen. Die Perspektiven hängen scheinbar gar nicht oder höchstens mittelmäßig miteinander zusammen. Deshalb können die drei Perspektiven als getrennte Teilkonstrukte eines übergeordneten autonomieunterstützenden beziehungsweise kontrollierenden Motivierstils angesehen werden (Aelterman et al., 2014; Field, 2009). Deshalb empfehlen Haerens et al. (2015), dass Studien alle drei Perspektiven erheben sollten.

6.2 Veränderlichkeit der Motivierstile

Entsprechend dem vorhergehenden Unterkapitel hängen die verschiedenen Perspektiven, aus denen die Motivierstile erfasst werden können, nicht eindeutig miteinander zusammen. Wenn die Veränderlichkeit der Motivierstile diskutiert wird, sollten idealerweise alle

drei Perspektiven berücksichtigt werden (Haerens et al., 2015). Für die erste Perspektive konnte mehrfach gezeigt werden, dass Interventionen die selbstberichteten Motivierstile von Lehrkräften und Lehramtsstudierenden verändern können (Reeve, 1998; Reeve et al., 1999; Reeve & Cheon, 2016). Reeve (1998) sowie Reeve et al. (1999) haben den bilanzierten Motivierstil betrachtet. Reeve und Cheon (2016) haben die Veränderlichkeit des selbstberichteten autonomieunterstützenden Motivierstils von Lehrkräften gezeigt. Für Cross-Age-Tutoren wurde die Veränderlichkeit der selbstberichteten Motivierstile bisher nicht untersucht. Die beobachteten Motivierstile (zweite Perspektive) von Lehrkräften und Cross-Age-Tutoren kann sich durch eine Intervention signifikant ändern (Berger et al., 2017; Reeve et al., 2004; Tessier et al., 2008). In der Studie von Jang et al. (2016) wurde der wahrgenommene autonomieunterstützende Motivierstil erfragt (dritte Perspektive). Die Betreuten der autonomieunterstützenden Versuchsbedingung gaben einen signifikant höheren wahrgenommenen autonomieunterstützenden Motivierstil an (Jang et al., 2016).

Zusammengenommen lässt sich eine Veränderbarkeit der Motivierstile bezüglich der drei Perspektiven feststellen. Diese Veränderlichkeit der Motivierstile wurde für Lehrkräfte (Reeve et al., 2004), Lehramtsanwärterinnen und -anwärter (Reeve, 1998), Sporttrainerinnen und -trainer (Mahoney et al., 2016), Managerinnen und Manager (Deci et al., 1989), Medizinstudierende (Williams & Deci, 1996) sowie Zahnärztinnen und Zahnärzte (Halvari & Halvari, 2006) bestätigt.

6.3 Gestaltungsmerkmale effektiver Interventionen zur Förderung eines autonomieunterstützenden und Reduktion eines kontrollierenden Motivierstils

Aufbauend auf die für viele Berufsgruppen gezeigte Veränderlichkeit der Motivierstile hat die Metaanalyse von Su und Reeve (2011) verschiedene Gestaltungsmerkmale und Merkmale der Teilnehmenden untersucht, um daraus Merkmale effektiver Interventionen zu identifizieren. Die Autoren untersuchten u. a. die Merkmale Berufsgruppen (Lehrkräfte, Eltern, Ärztinnen und Ärzte oder Managerinnen und Manager), Betreuungserfahrung (erfahren oder unerfahren), Interventionssetting (Laborsetting oder authentisches Setting), genutzte Medien (analoge Medien, elektronische Medien oder beides) und den Fokus des Trainings auf Wissenserwerb oder Fähigkeitstraining. Darüber hinaus untersuchten die Autoren den Einfluss der autonomen und kontrollierenden Kausalitätsorientierung der Teilnehmenden sowie das Auftreten möglicher Interventionsinhalte, also welche der Gruppen autonomieunterstützenden Verhaltens (siehe Kapitel 4) thematisiert wurden. Laut der Metaanalyse von Su und Reeve (2011) verändern solche Interventionsprogramme die Motivierstile für alle genannten Gestaltungs- und Teilnehmermerkmale signifikant, jedoch je nach konkreter Auswahl der Merkmale unterschiedlich effektiv. So ergab die Metaanalyse, dass diese Interventionsprogramme für Lehrkräfte effektiver sind als für Eltern oder Ärztinnen und Ärzte (Su & Reeve, 2011). Darüber hinaus sind die Interventionen für unerfahrene Teilnehmende effektiver als für erfahrene (Su & Reeve, 2011). Außerdem scheinen Interventionen im Umfang mehrerer Stunden an einem Tag effektiver zu sein als Interventionen, die kürzer als eine Stunde dauern oder über mehrere Wochen hinweg stattfinden. Die Trainings in authentischen Situationen, also z. B. im Klassenraum während des laufenden Schulbetriebes, sind weniger effektiv als Trainings in Laborsettings (Su & Reeve, 2011). Weiterhin scheinen Interventionen effektiver zu sein, wenn mehrere Gruppen autonomieunterstützen-

den Verhaltens (siehe Kapitel 4) statt nur einzelne thematisiert wurden. Der Einsatz von sowohl analogen als auch digitalen Medien ist effektiver als der separate Einsatz. Weiterhin sollte der Fokus auf dem Einüben konkreter autonomieunterstützender Verhaltensweisen liegen (Su & Reeve, 2011). Neben diesen Befunden der Metastudie von Su und Reeve (2011) folgt aus einzelnen Studien ein weiteres mögliches Gestaltungsmerkmal effektiver Interventionen. So scheint die Intervention für Cross-Age-Tutoren effektiver zu sein, wenn prototypische Videosequenzen zur Vermittlung der Interventionsinhalte genutzt und die Verhaltensweisen anschließend in Rollenspielen eingeübt werden (Berger et al., 2017; siehe Kapitel 2).

Ein weiteres Merkmal, dessen Einfluss in der Metastudie untersucht wurde, ist die Kausalitätsorientierung der Teilnehmenden. Der autonomieunterstützende Motivierstil wird durch eine Intervention bei Teilnehmenden mit einer hohen autonomen Kausalitätsorientierung mehr gefördert als bei Teilnehmenden mit einer relativ niedrigen autonomen Kausalitätsorientierung (Su & Reeve, 2011). Analog lässt sich der kontrollierende Motivierstil bei Teilnehmenden mit einer hohen kontrollierenden Kausalitätsorientierung nicht so effektiv reduzieren wie bei Teilnehmenden mit einer niedrigen kontrollierenden Kausalitätsorientierung (Su & Reeve, 2011). Bei Reeve (1998) konnte keine Interaktion des Interventionseffektes mit der Kausalitätsorientierung nachgewiesen werden. Bei Williams und Deci (1996) wurde lediglich der postinterventionelle autonomieunterstützende Motivierstil ohne Kontrolle des präinterventionellen Wertes mit den Kausalitätsorientierungen korreliert. Es wurden zwar Korrelationen gefunden, doch kann ein Einfluss des präinterventionellen autonomieunterstützenden Motivierstils nicht ausgeschlossen werden. Reeve et al. (2018) konnten zeigen, dass die postinterventionellen Motivierstile unter Kontrolle der präinterventionellen Werte von den jeweiligen Kausalitätsorientierungen vorhergesagt werden. Die Metaanalyse von Su und Reeve (2011), welche die Daten von Reeve (1998), Williams und Deci (1996) sowie einer weiteren Studie zusammennimmt, konnte die bilanzierte Kausalitätsorientierung als einen Moderator des Interventionseffektes nachweisen. Die Relevanz der Kausalitätsorientierung für den Interventionserfolg ist also durchaus umstritten und insbesondere für Tutoren bisher nicht untersucht. Als Erklärung für einen möglichen differentiellen Effekt der Kausalitätsorientierung wird vermutet, dass die Teilnehmenden mit einer relativ kontrollierenden Kausalitätsorientierung den Interventionsinhalten ablehnend gegenüberstehen, sie als widersprüchlich zu der eigenen motivationalen Orientierung empfinden und nur dann von der Intervention profitieren, wenn sie einen Konzeptwechsel vollziehen (Reeve, 1998, 2002; Reeve et al., 2018). Reeve (1998) konnte als Indizien dafür zeigen, dass die Zustimmung zu den Interventionsinhalten bei Teilnehmenden mit einer ausgeprägten kontrollierenden Kausalitätsorientierung ein Moderator für den Interventionserfolg ist. Empfinden Interventionsteilnehmende kontrollierendes Verhalten als effektiver oder einfacher zu implementieren, so durchlaufen sie vermutlich keinen Konzeptwechsel und profitieren weniger oder gar nicht von der Intervention (Reeve, 2002; Reeve & Cheon, 2016). Weiterhin wird vermutet, dass die Teilnehmenden mit einer ausgeprägten autonomen Kausalitätsorientierung den Interventionsinhalten schon vor der Intervention zugeneigt sind und aufgrund der Passung zur eigenen motivationalen Orientierung mehr profitieren (Reeve, 1998, 2002; Reeve et al., 2018). Da Lehrkräfte und Lehramtsstudierende eher mit kontrollierenden Lehrensätzen vertraut sind (Reeve, 1998), könnten Teilnehmende unterschiedlich viel von einer Intervention profitieren. Sind sie, zumindest implizit, mit den wesentlichen Verhaltensweisen von Autonomieunterstützung vertraut, könnten sie durch die Intervention weniger neue Ansätze lernen und somit ihren autonomieunterstützenden Motivierstil weniger ändern. Folglich könnte die vorherige Bekanntheit der Interventionsinhalte ebenfalls ein Moderator des Interventionserfolgs sein.

6.4 **Intervenierbarkeit der Perspektiven der Motivierstile von Cross-Age-Tutoren**

In Kapitel 6.1 wurde dargestellt, dass je nach Messperspektive die Interventionsergebnisse zu Motivierstilen unterschiedlich ausfallen und inkonsistent sind. Ebenso fehlt es an Befunden speziell zum Cross-Age-Tutoring. Daher können nur erste Vermutungen aufgestellt werden, wie die Selbstberichte von Cross-Age-Tutoren, deren Verhaltensanalysen und die Selbstberichte der Tutees über ihre Tutoren miteinander zusammenhängen. Haerens et al. (2015) empfehlen, dass Studien alle drei Perspektiven erheben sollten.

Da der Interventionserfolg von der Berufsgruppe der Interventionsteilnehmenden abhängt (Su & Reeve, 2011), ist unklar, ob und inwiefern die Motivierstile von Cross-Age-Tutoren interveniert werden können. Aus der Studie von Berger et al. (2017) lässt sich nur ableiten, dass entsprechend trainierte Cross-Age-Tutoren einen autonomieunterstützenderen beziehungsweise weniger kontrollierenden bilanzierten Motivierstil durch ihr Verhalten zeigen als untrainierte (Berger et al., 2017; Hänze et al., 2018). Einerseits ist nicht untersucht, ob sowohl der autonomieunterstützende als auch der kontrollierende Motivierstil durch die Intervention beeinflusst werden oder nur einer. Andererseits ist bisher nicht untersucht, ob sich durch die Intervention wirklich die Motivierstile oder nur Aspekte des Drucks von Oben oder Innen geändert haben (siehe Kapitel 5.2.4). Bei Berger et al. (2017) und Hänze et al. (2018) wurden weder die selbstberichteten Motivierstile der Cross-Age-Tutoren noch der wahrgenommene autonomieunterstützende oder kontrollierende Motivierstil erhoben. Insgesamt bleibt zu untersuchen, welche Motivierstile bei Cross-Age-Tutoren durch eine Intervention geändert werden können. Kann die Veränderlichkeit der Motivierstile nachgewiesen werden, bleibt die Übertragbarkeit der oben vorgestellten differentiellen Effekte zu zeigen. In nur wenigen Studien wurde untersucht, inwiefern die Kausalitätsorientierungen Moderatoren des Interventionseffektes sind. Insbesondere hat keine der Studien Cross-Age-Tutoren untersucht, weshalb der Befund für diese Kohorte nicht gesichert ist. Insgesamt lässt sich das oben beschriebene Desiderat bezüglich der Relevanz der Selbstbestimmungstheorie bei Cross-Age-Tutoring-Angeboten um drei differentielle Fragen erweitern:

- Welche Zusammenhänge bestehen zwischen den drei Messperspektiven der Motivierstile?
- Welche Motivierstile lassen sich durch eine Intervention fördern oder reduzieren?
- Inwiefern moderieren die Kausalitätsorientierungen, die Zustimmung zu den Interventionseinhalten und die vorherige Bekanntheit dieser Inhalte die Interventionserfolge?

7 Forschungsfragen und Hypothesen

Die vorliegende Studie untersucht die oben dargestellten Desiderate mit Fokus auf das Cross-Age-Tutoring in der Sekundarstufe I in Verbindung mit physikalischen Inhalten (siehe Kapitel 2). Bei der Formulierung der Forschungsfragen und Hypothesen sind zur einfacheren Lesbarkeit mit *Tutoren* und *Tutees* daher Schülerinnen und Schüler aus der Sekundarstufe I gemeint, die an einem Cross-Age-Tutoring-Angebot mit einem Bezug auf physikalische Experimente teilnehmen. In Analogie zu Berger et al. (2017) und das im Kapitel 5.2 hergeleitete Wirkmodell auf Motivierstile sind weiterhin nur Tutoren gemeint, die nicht durch externe Anreize zu einer Teilnahme motiviert wurden und die ihre Tutees weder aus vergangenen Lernsituationen kennen noch für deren Leistungen verantwortlich gemacht werden.

7.1 Forschungsfrage 1

In Kapitel 6.1 wurde dargestellt, dass die selbstberichteten, die aus Verhaltensanalysen stammenden und die von den Betreuten wahrgenommen Motivierstile inkonsistent miteinander zusammenhängen. Deshalb empfehlen Haerens et al. (2015), alle drei Perspektiven zu erheben und ihre Zusammenhänge zu untersuchen. Dabei wird zwischen dem autonomieunterstützenden und dem kontrollierenden Motivierstil differenziert. Gemäß Kapitel 5.1 wird sich bei den von den Tutees wahrgenommenen Motivierstilen auf den autonomieunterstützenden beschränkt. Folglich lauten die Teilforschungsfragen:

FF1a: Inwiefern hängen der von den Tutoren selbstberichtete, der von Experten beobachtete und der von den Tutees wahrgenommene autonomieunterstützende Motivierstil der Tutoren miteinander zusammen?

FF1b: Inwiefern hängen der von den Tutoren selbstberichtete und der von Experten beobachtete kontrollierende Motivierstil miteinander zusammen?

Einerseits kann eine Analyse des Zusammenhangs einen Beitrag zur psychologischen Grundlagenforschung liefern, indem untersucht wird, inwiefern und unter welchen Voraussetzungen sich die Selbstberichte in tatsächliches Handeln übertragen. Andererseits kann mit der Analyse der Zusammenhänge der Messperspektiven untersucht werden, inwiefern theoretisch angenommene autonomieunterstützende Verhaltensweisen von Tutees als solche wahrgenommen werden. Die Antworten auf die Fragen geben also Hinweise auf die Struktur der Konstrukte. In einigen Studien (z. B. Pelletier et al., 2002; Rocchi et al., 2013) wurden nur die selbstberichteten Motivierstile erhoben, so dass diese Ergebnisse bisher nicht auf das gezeigte Verhalten der Probanden übertragen werden konnten. Weitere Einflussgrößen wie die Kausalitätsorientierungen der Betreuten (Baard et al., 2004; Deci & Ryan, 1985) werden in dieser Studie nicht im Detail untersucht.

In Kapitel 5.1 wurden die inkonsistenten Befunde bezüglich der Zusammenhänge der drei Perspektiven der Erfassung der Motivierstile dargelegt. Aufgrund der nicht vorhandenen bis

mittleren Größe der Zusammenhänge wurde geschlussfolgert, dass der von den Betreuenden selbstberichtete, der im Verhalten beobachtete und der von den Betreuten wahrgenommene autonomieunterstützende Motivierstil als getrennte Teilkonstrukte eines zugrundeliegenden autonomieunterstützenden Motivierstils aufgefasst werden können. In Kapitel 5.1 wurde dargelegt, weshalb die Perspektiven der Motivierstile vermutlich in Höhe von etwa $r = .4$ zusammenhängen werden. Trotz der bisherigen inkonsistenten Befundlage werden im Einklang mit der Selbstbestimmungstheorie (Ryan & Deci, 2017) folgende Hypothesen aufgestellt:

H1a: Der selbstberichtete autonomieunterstützende Motivierstil der Tutoren, der beobachtete autonomieunterstützende Motivierstil und der von den Tutees wahrgenommene autonomieunterstützende Motivierstil hängen positiv in Höhe von $r = .4$ miteinander zusammen.

H1b: Der selbstberichtete kontrollierende Motivierstil der Tutoren und der beobachtete kontrollierende Motivierstil hängen positiv in Höhe von $r = .4$ miteinander zusammen.

7.2 Forschungsfrage 2

Im Kapitel 6 wurde dargestellt, dass Interventionen zur Förderung eines autonomieunterstützenden Motivierstils bzw. zur Reduktion eines kontrollierenden Motivierstils erfolgreicher sind, wenn sie einige Stunden umfassen, in einem Laborsetting stattfinden und das Konzept der Autonomieunterstützung sowohl durch Präsentationen als auch durch Arbeitsblätter vermitteln. Der Fokus der Intervention in dieser Studie soll daher auf dem praktischen Training autonomieunterstützender Verhaltensweisen aus verschiedenen Verhaltensgruppen liegen (siehe Kapitel 4). Diese Verhaltensweisen sollten durch prototypische Videosequenzen veranschaulicht und durch Rollenspiele eingeübt werden (siehe Kapitel 6).

Berger et al. (2017) liefern Indizien dafür, dass der bilanzierte Motivierstil von Cross-Age-Tutoren als veränderlich angenommen werden kann. Insbesondere konnte bei entsprechend trainierten Cross-Age-Tutoren ein eher autonomieunterstützenderer bilanzierter Motivierstil durch Verhaltensbeobachtung im Vergleich zu untrainierten Tutoren festgestellt werden (Berger et al., 2017; Hänze et al., 2018). Allerdings ist unklar, ob sich beide Motivierstile durch eine Intervention verändern lassen. Es wäre alternativ denkbar, dass durch die Intervention lediglich Aspekte des Drucks von Oben oder Innen und nicht die selbstberichteten Motivierstile der Teilnehmenden interveniert wurden, also die Tutoren ein erwünschtes Verhalten auswendig lernen und bei Beobachtung zeigen. Entsprechend lauten die Teilforschungsfragen:

FF2a: Inwiefern kann eine Intervention den selbstberichteten autonomieunterstützenden Motivierstil von Tutoren fördern?

FF2b: Inwiefern kann eine Intervention den selbstberichteten kontrollierenden Motivierstil von Tutoren reduzieren?

Im Folgenden werden die Forschungsfragen durch drei Gruppen an Hypothesen spezifiziert. Einerseits werden Hypothesen über die Höhe der Änderungen der Motivierstile durch die Intervention aufgestellt. Im Anschluss werden Prädiktoren dieser Änderungen der Motivierstile vermutet. Zuletzt werden unterschiedliche Relevanzen dieser Prädiktoren für Teilgruppen der Interventionsteilnehmenden vermutet.

7.2.1 Hypothesen über die Änderungen der Motivierstile durch die Intervention

Die Tutoren in der Studie von Berger et al. (2017) zeigten nach der Intervention ein signifikant ($d = .55, p < .05$) autonomieunterstützenderen beobachteten bilanzierten Motivierstil als die Tutoren der Kontrollgruppe ohne jegliches Training. Laut Su und Reeve (2011) hängt der durchschnittliche Interventionseffekt von der Messperspektive des Motivierstils ab. Weiterhin präsentieren sie für Studien mit Selbstberichtmaßen der Betreuenden Effektstärken nach J. Cohen (1988) zwischen $d_{min} = -.29$ bis $d_{max} = 1.20$. Die berichteten Effektstärken bei Su und Reeve (2011) sind Punktschätzer einzelner Studien, die verteilt um einen wahren Wert liegen. Da sich die vorliegende Studie an den Empfehlungen von Su und Reeve (2011) orientiert, werden große Effektstärken von $d = 0.8$ erwartet (J. Cohen, 1988). Zusammengenommen werden aufbauend auf die Befunde von Berger et al. (2017), Hänze et al. (2018) und der Metaanalyse von Su und Reeve (2011) folgende Hypothesen aufgestellt:

- H2a₁ : Eine Intervention kann den selbstberichteten autonomieunterstützenden Motivierstil von Tutoren mit einer Effektstärke von etwa $d = 0.8$ fördern.
- H2b₁ : Eine Intervention kann den selbstberichteten kontrollierenden Motivierstil von Tutoren mit einer Effektstärke von etwa $d = 0.8$ reduzieren.

7.2.2 Hypothesen über Prädiktoren der Änderungen der Motivierstile

In Kapitel 6 wurde dargestellt, dass die Effektivität solcher Interventionen möglicherweise von den Kausalitätsorientierungen der Teilnehmenden, ihrer Zustimmung zu den Interventionseinhalten und der vorherigen Bekanntheit dieser Inhalte abhängt. Bei Reeve (1998) konnte keine Interaktion des Interventionseffektes mit der bilanzierten Kausalitätsorientierung nachgewiesen werden. Bei Williams und Deci (1996) wurde lediglich der postinterventionelle autonomieunterstützende Motivierstil ohne Kontrolle des präinterventionellen Wertes mit den Kausalitätsorientierungen korreliert. Es wurden zwar Korrelationen gefunden, jedoch kann ein Einfluss des präinterventionellen autonomieunterstützenden Motivierstils nicht ausgeschlossen werden. Reeve et al. (2018) konnten zeigen, dass die postinterventionellen Motivierstile unter Kontrolle der präinterventionellen Werte von den jeweiligen Kausalitätsorientierungen vorhergesagt werden. Die Metaanalyse von Su und Reeve (2011), welche die Daten von Reeve (1998), Williams und Deci (1996) sowie einer weiteren Arbeit zusammennimmt, konnte die bilanzierte Kausalitätsorientierung als einen Moderator des Interventionseffektes nachweisen. Die Relevanz der Kausalitätsorientierungen für den Interventionserfolg ist also durchaus umstritten und insbesondere für Tutoren bisher nicht untersucht. Bezüglich dieser Forschungslücke ergeben sich in Anlehnung an Su und Reeve (2011) folgende spezifizierende Hypothesen zu den Forschungsfragen FF2a und FF2b:

- H2a_{2,I} : Je stärker die autonome Kausalitätsorientierung der Tutoren ausgeprägt ist, desto mehr steigt der selbstberichtete autonomieunterstützende Motivierstil durch eine entsprechende Intervention.
- H2b_{2,I} : Je stärker die kontrollierende Kausalitätsorientierung der Tutoren ausgeprägt ist, desto weniger wird der selbstberichtete kontrollierende Motivierstil durch eine entsprechende Intervention reduziert.

Reeve (1998) korrelierte die Zustimmung zu den Interventionsinhalten mit dem postinterventionellen selbstberichteten bilanzierten Motivierstil. Für Interventionsteilnehmende mit einer relativ autonomen bilanzierten Kausalitätsorientierung konnte keine Korrelation gefunden werden. Für Teilnehmende mit einer relativ kontrollierenden bilanzierten Kausalitätsorientierung wurde eine starke Korrelation in Höhe von $r = .65, p < .001$ gefunden. Reeve (1998) interpretiert dies als Indiz, dass Teilnehmende mit einer relativ kontrollierenden bilanzierten Kausalitätsorientierung einen Konzeptwechsel durchlaufen müssen, um von der Intervention zu profitieren. Da diese Korrelation den präinterventionellen bilanzierten Motivierstil und die übrigen Prädiktoren ignoriert, könnte dieser Einfluss durch die übrigen Variablen wegfallen. Entsprechend Reeve (1998) wird für die Gesamtgruppe aller Interventionsteilnehmenden ein mittlerer Effekt auf den Zuwachs bzw. die Reduktion vermutet:

H2a_{2.II} : Je stärker die Tutoren den Interventionsinhalten zustimmen, desto mehr steigt der selbstberichtete autonomieunterstützende Motivierstil ($\beta = .3$) durch eine entsprechende Intervention.

H2b_{2.II} : Je stärker die Tutoren den Interventionsinhalten zustimmen, desto mehr wird der selbstberichtete kontrollierende Motivierstil ($\beta = -.3$) durch eine entsprechende Intervention reduziert.

Tutoren, denen die Interventionsinhalte über Autonomieunterstützung bereits bekannt sind, sehen Autonomieunterstützung möglicherweise als einen sinnvollen Weg, andere zu motivieren. Folglich haben sie vermutlich einen höheren präinterventionellen autonomieunterstützenden Motivierstil, welcher vergleichsweise weniger zunehmen kann. Entsprechend sehen sie Kontrolle vermutlich als einen weniger sinnvollen Weg, andere zu motivieren. Folglich haben sie vermutlich einen niedrigeren präinterventionellen kontrollierenden Motivierstil, welcher vergleichsweise weniger reduziert werden kann:

H2a_{2.III} : Je mehr die Tutoren die Interventionsinhalte bereits vor der Intervention kennen, desto weniger steigt ihr selbstberichteter autonomieunterstützende Motivierstil durch eine entsprechende Intervention.

H2b_{2.III} : Je mehr die Tutoren die Interventionsinhalte bereits vor der Intervention kennen, desto weniger wird ihr selbstberichteter kontrollierende Motivierstil durch eine entsprechende Intervention reduziert.

7.2.3 Hypothesen über die Zustimmung zu den Interventionsinhalten als Prädiktor der Änderungen der Motivierstile für Teilnehmende mit unterschiedlichen Kausalitätsorientierungen

Wie im vorletzten Absatz erläutert, unterscheidet sich die Relevanz der Zustimmung zu den Interventionsinhalten für die Prädiktion der Interventionserfolge vermutlich für Teilgruppen unterschiedlicher Kausalitätsorientierungen (Reeve, 1998). Für Interventionsteilnehmende mit einer relativ autonomen bilanzierten Kausalitätsorientierung konnte keine Korrelation zwischen der Änderung des bilanzierten Motivierstils und der Zustimmung zu den Interventionsinhalten gefunden werden. Für Teilnehmende mit einer relativ kontrollierenden bilanzierten Kausalitätsorientierung wurde eine starke Korrelation in Höhe von $r = .65, p < .001$ gefunden. Somit ergeben sich folgende Hypothesen:

- H2a_{3,I} : Eine Intervention erhöht den selbstberichteten autonomieunterstützenden Motivierstil von Teilnehmenden mit einer relativ stark ausgeprägten autonomen Kausalitätsorientierung unabhängig von ihrer Zustimmung zu den Interventionsinhalten.
- H2b_{3,I} : Eine Intervention reduziert den selbstberichteten kontrollierenden Motivierstil von Teilnehmenden mit einer relativ schwach ausgeprägten kontrollierenden Kausalitätsorientierung unabhängig von ihrer Zustimmung zu den Interventionsinhalten.
- H2a_{3,II} : Eine Intervention erhöht den selbstberichteten autonomieunterstützenden Motivierstil von Teilnehmenden mit einer relativ schwach ausgeprägten autonomen Kausalitätsorientierung in starker Abhängigkeit ($\beta > .6$) von ihrer Zustimmung zu den Interventionsinhalten.
- H2b_{3,II} : Eine Intervention reduziert den selbstberichteten kontrollierenden Motivierstil von Teilnehmenden mit einer relativ stark ausgeprägten kontrollierenden Kausalitätsorientierung in starker Abhängigkeit ($\beta > .6$) von ihrer Zustimmung zu den Interventionsinhalten.

7.3 Forschungsfrage 3

Die Motivierstile werden von verschiedenen Konstrukten beeinflusst, die sich in den Druck von Oben, den Druck von Innen und den Druck von Unten (Erwartungshaltung und spontan) kategorisieren lassen. Der Druck von Oben hat keinen oder lediglich einen medierten Einfluss auf die Motivierstile (siehe Kapitel 5.2.4). Dagegen hat der Druck von Innen einen direkten Einfluss auf die Motivierstile (siehe Kapitel 5.2.4). Der Druck von Unten (Erwartungshaltung) wirkt partiell mediert über den Druck von Innen auf die Motivierstile (siehe Kapitel 5.2.4). Der Druck von Unten (spontan) wird vermutlich vom Druck von Innen beeinflusst und beeinflusst vermutlich die Motivierstile (siehe Kapitel 5.2.4).

Das dieser Studie zugrunde liegende Modell (siehe Kapitel 5.2.4 sowie Abbildung 5.3) lehnt sich stark an Pelletier et al. (2002) und Rocchi et al. (2013) an, welche nur einen Aspekt des Drucks von Innen erfasst haben. Weiterhin haben Pelletier et al. (2002) nur den bilanzierten und Rocchi et al. (2013) nur den autonomieunterstützenden Motivierstil erhoben. Ob und inwiefern dieses Modell für den kontrollierenden Motivierstil gilt, ist bisher nicht untersucht. Bezüglich des Modells besteht also deutlicher Forschungsbedarf. Für diese Studie vereinfacht sich das Modell, da die Tutoren nicht durch externe Anreize zu einer Teilnahme motiviert wurden und sie ihre Tutees weder aus vergangenen Lernsituationen kannten noch für deren Leistungen verantwortlich gemacht wurden. Damit fallen der Druck von Oben und der Druck von Unten (Erwartungshaltung) weg und es bleibt ein Wirkmodell übrig, bei dem der Druck von Innen und der Druck von Unten (spontan) die Motivierstile vorhersagen (siehe Kapitel 5.2.4 sowie Abbildung 5.4). Neben den in Abbildung 5.4 genannten allgemeinspsychologischen Konstrukten könnten fachbezogene Konstrukte ebenfalls für die Vorhersage der beobachteten Motivierstile relevant sein (Reeve, 2009; Su & Reeve, 2011). Beispielsweise können Tutoren mit mangelhaftem fachlichem Verständnis ihren Tutees nur bedingt inhaltlich helfen, was zu einer Frustration des Fähigkeitsselbstkonzeptes auf beiden Seiten und damit zu einem enger geleitetem Experimentierprozess führen könnte (Su & Reeve, 2011). Deshalb wird ebenfalls für das Fähigkeitsselbstkonzept der Tutoren

bezüglich Physik, das konkrete Fachwissen sowie das Verständnis der Variablen-Kontroll-Strategie kontrolliert. Insgesamt lauten die Forschungsfragen zum Modell aus Abbildung 5.4:

FF3: Inwiefern hängen die Aspekte des Drucks von Unten (spontan), die Aspekte des Drucks von Innen, der beobachtete autonomieunterstützende Motivierstil und der beobachtete kontrollierende Motivierstil bei Tutoren miteinander zusammen?

Somit kann in Anlehnung an die Herleitung des Wirkmodells (siehe Kapitel 5.2.4 sowie Abbildung 5.4) die Gültigkeit des Modells vermutet werden:

H3a: Die Aspekte des Drucks von Innen und die Aspekte des Drucks von Unten (spontan) sagen den beobachteten autonomieunterstützenden Motivierstil vorher.

H3b: Die Aspekte des Drucks von Innen und die Aspekte des Drucks von Unten (spontan) sagen den beobachteten kontrollierenden Motivierstil vorher.

H3c: Der Druck von Innen sagt die Motivierstile partiell mediiert über den Druck von Unten (spontan) vorher.

7.4 Forschungsfrage 4

Hänze et al. (2018) nehmen in Einklang mit der Selbstbestimmungstheorie an, dass die Befriedigung des Kompetenzerlebens von Tutees eng mit ihrer intrinsischen Motivation zusammenhängt, ohne die entsprechenden Zusammenhänge explizit anzugeben. Dieser Zusammenhang wurde sowohl für Schülerinnen und Schüler im Regelunterricht als auch für Studierende gezeigt (Jang et al., 2009; Reeve et al., 2003). Für Tutees von Cross-Age-Tutoren ist dieser Zusammenhang bisher nicht belegt.

In Kapitel 5.1 wurde dargelegt, dass Betreuende im Allgemeinen von einer autonomieunterstützenden Betreuung profitieren und unter einer kontrollierenden Betreuung leiden. Eine autonomieunterstützende Betreuung befriedigt sowohl das Autonomieerleben als auch das Kompetenzerleben. Dies führt zu weiteren positiven Effekten, wie beispielsweise einer gesteigerten intrinsischen Motivation. Die Autonomieunterstützung von Tutoren wirkt sich also vermutlich nicht direkt auf die intrinsische Motivation der Tutees aus, sondern wird über die Befriedigung des Autonomie- und Kompetenzerlebens mediiert. Laut Jang et al. (2009) ist die Befriedigung des Kompetenzerlebens für eine intrinsische Motivation relevanter als die Befriedigung des Autonomieerlebens. Ob der von Tutees wahrgenommene autonomieunterstützende Motivierstil tatsächlich die Befriedigung des Kompetenzerlebens vorhersagt und ob dieser Motivierstil ausschließlich über die Befriedigung des Kompetenzerlebens mediiert auf die intrinsische Motivation wirkt, ist für Cross-Age-Tutoren bisher nicht bestätigt. Weiterhin wird für die intrinsische Motivation vor der Durchführung des Experiments kontrolliert, da diese vermutlich ebenfalls die intrinsische Motivation nach der Durchführung des Experiments vorhersagt (Jang et al., 2016b).

FF4a: Inwiefern hängt der von den Tutees wahrgenommene autonomieunterstützende Motivierstil der Tutoren mit der Befriedigung des Kompetenzerlebens der Tutees und ihrer intrinsischen Motivation zusammen?

Aufbauend auf der Herleitung des Wirkmodells in Kapitel 5.1 und den Veröffentlichungen von Hänze et al. (2018), Jang et al. (2009), Jang et al. (2016b) sowie Jang et al. (2016) wird die Gültigkeit des Modells angenommen:

H4a: Unter Kontrolle der intrinsischen Motivation vor der Durchführung des Experiments, sagt der von den Tutees wahrgenommene autonomieunterstützende Motivierstil der Tutoren ihre intrinsische Motivation vorher. Dieser Zusammenhang ist nicht direkt, sondern wird vollständig über die Befriedigung des Kompetenzerlebens der Tutees mediiert.

Wie in Kapitel 2 beschrieben, ergab die Metaanalyse von Leung (2015), dass gleichgeschlechtliche Zusammensetzungen von Tutor-Tutee-Gruppen zu deutlich höheren Lernzuwächsen bei den Tutees führen. Da die Motivation häufig mit erfolgreichem Lernen assoziiert wird (z. B. Klauer und Leutner, 2012), könnte sich der von Leung (2015) dargestellte Effekt auf die Motivation der Tutees niederschlagen. Die Geschlechterkomposition könnte also eine wichtige Variable bei Tutoring-Angeboten sein. Somit ergibt sich eine weitere Forschungsfrage:

FF4b: Welche Relevanz hat die Geschlechterkomposition der Tutoren und Tutees auf den wahrgenommenen autonomieunterstützenden Motivierstil, das Kompetenzerleben und die intrinsische Motivation der Tutees?

In Anlehnung an die Befunde von Leung (2015) wird folgende Hypothese aufgestellt:

H4b: Bei gleichgeschlechtlichem Tutor geben Tutees einen signifikant ($p < .05$) höheren wahrgenommenen autonomieunterstützenden Motivierstil, eine höhere intrinsische Motivation und eine höhere Befriedigung des Kompetenzerlebens an im Vergleich zu Tutees mit einem Tutor eines anderen Geschlechts.

8 Design und Methoden

In Kapitel 7 wurden Hypothesen formuliert, welche mit der vorliegenden Studie untersucht werden. Aus diesen Hypothesen wird im Folgenden der Studienablauf hergeleitet.

8.1 Studienablauf

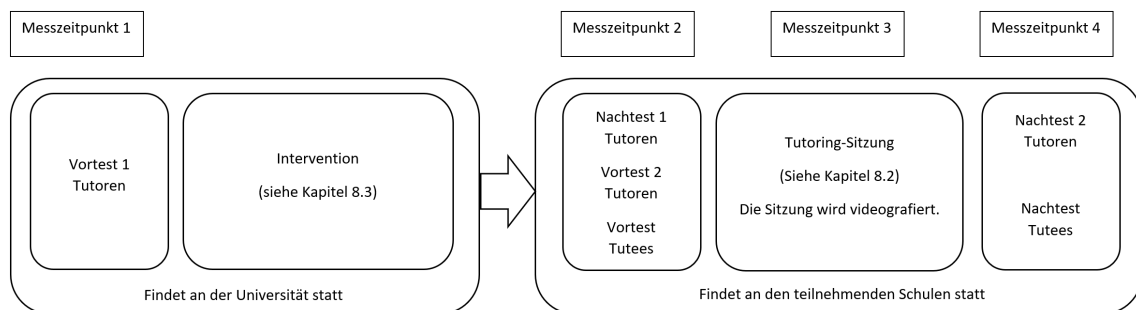


Abbildung 8.1: Studienablauf

Abbildung 8.1 stellt das Studiendesign dar, welches im Folgenden erläutert wird. Tabelle 8.1 führt auf, welche Konstrukte zu welchem Messzeitpunkt bei den Tutoren bzw. bei den Tutees erhoben wurden. Entsprechend Hypothese 2 wurde in der vorliegenden Studie der Effekt einer Intervention untersucht. Nach Bortz und Döring (2006) wurde ein Prä-Post-Design gewählt. Da Hypothese 2 keine Vermutungen über Follow-up-Ergebnisse beinhaltet, wurden zwei Messzeitpunkte benötigt. Sowohl der Vortest (Messzeitpunkt 1) als auch die Intervention fanden in den Laborräumen der Universität Duisburg-Essen statt. Die Intervention wird in Kapitel 8.3 ausführlich dargestellt. Der zweite Messzeitpunkt fand etwa zwei Wochen später (durchschnittlich 1.46 Wochen) an den teilnehmenden Projektschulen statt. Zur Untersuchung von Hypothese 3 wurden verschiedene Aspekte des Drucks von Innen und Unten (spontan) erfasst und mit dem Tutorverhalten in einer Tutoring-Sitzung in Verbindung gebracht. Um den organisatorischen Aufwand an den Schulen möglichst gering zu halten, wurde die Nachttestung der Intervention mit der Erhebung der Aspekte des Drucks von Innen zusammengelegt (Messzeitpunkt 2). Im Anschluss an den zweiten Messzeitpunkt führten die Tutees unter Betreuung eines Tutors einzeln ein standardisiertes Experiment durch, welches in Kapitel 8.2 genauer beschrieben wird. Die Tutoring-Sitzung wurde videografiert, um das Tutorverhalten nachträglich auswerten zu können (Messzeitpunkt 3). Die Aspekte des Drucks von Unten (spontan) sind die Tutorwahrnehmungen ihrer Tutees während der Tutoring-Sitzung (siehe Tabelle 5.2.4). Da die Tutoren diese Eindrücke erst in der Tutoring-Sitzung sammeln konnten, füllten sie den entsprechenden Fragebogen nach der Sitzung aus (Messzeitpunkt 4). Entsprechend Hypothese 4a wurde der Zusammenhang zwischen dem wahrgenommenen autonomieunterstützenden Motivierstil und dem

Kompetenzerleben der Tutees sowie ihrer intrinsischen Motivation unter Kontrolle ihrer intrinsischen Motivation vor der Durchführung des Betreuungsexperiments untersucht. Weiterhin wurde der Zusammenhang dieser Konstrukte zur Geschlechterzusammensetzung der Tutor-Tutee-Dyaden analysiert. Vor der Tutoring-Sitzung, also zu Messzeitpunkt 2, wurden beispielsweise das Geschlecht der Tutees und ihre intrinsische Motivation vor der Durchführung des Experiments erhoben. Zu Messzeitpunkt 4 wurden der wahrgenommene autonomieunterstützende Motivierstil, das Kompetenzerleben der Tutees sowie ihre intrinsische Motivation erfasst.

Den Tutees wurde vor der Tutoring-Sitzung erklärt, dass sie ein Experiment durchführen werden und ihnen je ein Tutor für alle Anliegen und Fragen zur Seite steht. Es wurde sichergestellt, dass sich die Tutoren und Tutees nicht kannten. Falls dies doch der Fall war, wurde bei der Zuordnung zusätzlich berücksichtigt, dass sie nicht zusammenarbeiten. Ansonsten erfolgte die Zuordnung der Tutees zu den Tutoren zufällig. Nach dieser allgemeinen Einführung erklärten die Experimentatoren den Tutoren den Ablauf und welche Materialien zur Verfügung stehen. Anschließend führten die Experimentatoren die Tutoren an ihre Experimentierplätze. Der genaue Wortlaut der Einführung kann den Interventionsmaterialien entnommen werden (siehe Kapitel 13. Interventionsmaterialien). An diesem Platz lag für die Tutoren ein Betreuerarbeitsblatt, also das Arbeitsblatt der Tutees mit Lösungen und physikalischen, inhaltlichen Hinweisen zum Experiment. Weiterhin lagen dort ein Arbeitsblatt für die Tutees und die Experimentiermaterialien. Die Tutoren hatten die Gelegenheit, sich das Schülerarbeitsblatt und ihr Betreuerarbeitsblatt vorher anzusehen. Inhaltliche Fragen wurden von den Experimentatoren nicht beantwortet, um die Studie nicht zu beeinflussen. Anschließend wurden die Tutees an einen Experimentierplatz zu ihren Tutoren geführt. Den Tutoren und Tutees wurde vorgelesen, dass sie maximal 25 Minuten Zeit haben oder das Experiment endet, sobald sie fertig sind und nicht mehr weiter experimentieren wollen. Es wurde darauf hingewiesen, dass kein Druck besteht, dass bestimmte Aufgaben erledigt werden müssen. Die Teilnahme der Tutoren und Tutees an dem Projekt sowie die Teilnahme an der Studie war freiwillig. Einverständniserklärungen der Erziehungsberechtigten liegen vor. Entsprechend der zur Verfügung stehenden Ausrüstung konnten maximal sechs Tutoren gleichzeitig videografiert werden. Bei Schulen mit mehr als sechs teilnehmenden Tutoren wurden die Tutoren nacheinander videografiert. Die Reihenfolge wurde entsprechend schulorganisatorischen Vorgaben und ansonsten zufällig bestimmt. Die Tutees wurden so ausgewählt, dass sie sich nicht mit den Tutees nachfolgender Tutoring-Sitzungen über das Experiment austauschen konnten.

	Messzeitpunkt 1	Messzeitpunkt 2	Messzeitpunkt 3	Messzeitpunkt 4
Tutoren	Selbstberichte - Motivierstile - Kausalitätsorientierungen - Alter, Geschlecht - Kontrollvariablen	Selbstberichte - Motivierstile - Kausalitätsorientierungen - Zustimmung zu den Interventioneninhalten - Vorherige Bekanntheit der Interventionen - Weitere Drücke von Innen	Analysiertes Verhalten - Motivierstile	Selbstberichte - Drücke von Unten
Tutees		Selbstberichte - Intrinsische Motivation vor der Durchführung des Experiments - Alter, Geschlecht - Kontrollvariablen		Selbstberichte - Wahrgenommener autonomieunterstützender Motivierstil - Befriedigung des Kompetenzerlebens - Intrinsische Motivation

Tabelle 8.1: Erhobene Konstrukte geordnet nach Messzeitpunkten und Tutor bzw. Tutee

8.2 Das Betreuungsexperiment der Tutoring-Sitzung

Das von den Tutees während der Tutoring-Sitzung durchgeführte Experiment hatte zum Ziel, die Leitfähigkeit von Kabeln verschiedener Längen, Dicken und Materialien anhand der Leuchtstärke von Glühlampen zu untersuchen. Konkret sollte untersucht werden, ob kurze oder lange Kabel, dicke oder dünne Kabel bzw. Kabel aus Eisen oder Konstantan Strom besser leiten. Dazu stand jedes der in Tabelle 8.2 aufgelisteten Kabel zweifach zur Verfügung, sodass mit Hilfe von zwei ebenfalls bereitgestellten 1.5V Batterien (inklusive Halterung) sowie zweier Glühlampen (1.5V; 0.2A) zwei Stromkreise gebaut werden konnten. Die Helligkeiten der Glühlampen in beiden Stromkreisen konnten verglichen werden, um die Leitfähigkeit der Kabel zu vergleichen.

Material	Dicke in mm	Länge in cm
Konstantan	0.4	45
Konstantan	0.4	15
Konstantan	0.2	45
Konstantan	0.2	15
Eisen	0.2	45
Eisen	0.2	15

Tabelle 8.2: Die vorhandenen Kabel beim Betreuungsexperiment der Tutoring-Sitzung

Für dieses Experiment ist die Variablen-Kontroll-Strategie, kurz VKS, eine relevante experimentelle Kompetenz (Schwichow et al., 2015). Bei Schwichow et al. (2015) standen Kabel mit drei unterschiedlichen Längen zur Verfügung. Für die vorliegende Studie wurden nur zwei mögliche Längen angeboten, um die Tutees nicht durch die Kabelanzahl kognitiv zu überlasten. Bei allen intendierten Vergleichen leuchteten die Lampen entweder hell, haben geglimmt oder leuchteten gar nicht, sodass die Beobachtungen eindeutig kategorisiert werden konnten.

Den Tutees stand ein Arbeitsblatt zur Verfügung, welches das Ziel des Experiments nannte und die zur Verfügung stehenden Materialien auflistete. Es schloss sich die erste Frage an, ob sich die Glühlampen zum Leuchten bringen ließen. Unter der Aufgabe stand der Hinweis, dass sie weitere Aufgaben von ihrem Tutor bekommen könnten. Auf der unteren Hälfte des Arbeitsblatts befand sich ein Bereich für Notizen. So war gewährleistet, dass alle Tutoren die Reihenfolge vorgeben oder frei von den Tutees entscheiden lassen konnten. Die Tutoren hatten die Gelegenheit, Abschweifungen zuzulassen oder zu unterbinden. Dadurch wurde bezüglich der Konstrukte die Möglichkeit für Autonomieunterstützung und Kontrolle gegeben. Die erste Aufgabe diente als Startpunkt, an dem sich die Tutoren und Tutees orientieren konnten, aber nicht mussten. Den Tutoren stand das gleiche Arbeitsblatt mit Lösungen und Kommentaren zur Verfügung (siehe Kapitel 8.1).

8.3 Konzeption und Entwicklung der Intervention

Die Motivierstile von Cross-Age-Tutoren können entsprechend Kapitel 6 als veränderlich angenommen werden. Soll eine Intervention den autonomieunterstützenden Motivierstil fördern und den kontrollierenden reduzieren, sollte die Intervention einige Stunden umfassen, in einem Laborsetting stattfinden und sowohl Präsentationen als auch Arbeitsblätter

nutzen (siehe Kapitel 6). Weiterhin sollte der Fokus der Intervention auf dem Einüben autonomieunterstützender Verhaltensweisen aus verschiedenen Verhaltensgruppen liegen. Diese Verhaltensweisen sollten durch prototypische Videosequenzen veranschaulicht und in Rollenspielen eingeübt werden (siehe Kapitel 6). Aufbauend auf diese Gestaltungshinweise und erfolgreiche Interventionen zur Veränderung von Motivierstilen von Reeve (1998), Reeve et al. (2004) sowie Berger et al., 2017 bestand die Intervention der vorliegenden Studie aus einer Inputphase, einer Übungsphase und einer Reflexionsphase. Der Ablaufplan der Pilotierung sowie der Haupterhebungsinterventionen können der Ergänzung des Anhangs entnommen werden (siehe Kapitel 13.Interventionsmaterialien).

Zu Beginn der Inputphase wurde das Betreuungsexperiment vorgestellt (siehe Kapitel 8.2), um zu gewährleisten, dass sich die Tutoren im weiteren Verlauf nicht auf fachphysikalische Inhalte konzentrieren müssen. Dadurch sollte eine hohe Lernzeit bezüglich des Konzepts der Autonomieunterstützung gewährleistet werden. Im zweiten Vortrag der Inputphase wurden ähnlich der Intervention von Reeve (1998) die Vorteile von Autonomieunterstützung und die Nachteile von kontrollierendem Verhalten dargelegt und erklärt. Es wurden nicht alle sechs Verhaltensgruppen autonomieunterstützenden Verhaltens thematisiert (siehe Kapitel 4), sondern lediglich die Verhaltensgruppen I - IV, da nicht alle sechs Verhaltensgruppen für das konkret vorliegende Cross-Age-Tutoring-Angebot sinnvoll umgesetzt werden können. Kennen die Tutoren beispielsweise ihre Tutees vor der Betreuung nicht, können die Tutoren die Inhalte kaum an die Wünsche der Tutees anpassen und somit nur bedingt an die intrinsischen motivationalen Ressourcen der Tutees anknüpfen (V). Da mit dem Tutoring-Angebot kein Test oder das routinierte Erarbeiten bestimmter Konzepte verbunden ist, fallen auch Routineaufgaben und die Identifizierung der Relevanz dieser Aufgaben weg (VI). Analog zu Berger et al. (2017) wurden die Konzepte nicht allgemein, sondern bezogen auf die Betreuung von Experimenten dargestellt, sodass die Tutoren keinen Transfer des Wissens leisten mussten. Weiterhin wurden die vier thematisierten Verhaltensgruppen autonomieunterstützenden Handelns in Form von Betreuungsregeln festgehalten (Berger et al., 2017). Die Betreuungsregeln lauten:

- Gib Sicherheit und ermutige bei Schwierigkeiten.
- Hilf nur so viel, dass dein Schüler selber weiterarbeiten kann.
- Hilf nur, wenn dein Schüler wirklich Hilfe braucht.
- Befiehl nicht, sondern berate.

Wie bei Berger et al. (2017) wurden diese Betreuungsregeln durch prototypische Videosequenzen veranschaulicht. Zum Abschluss des zweiten Vortrages wurde erneut die negative Videosequenz gezeigt. Alle Teilnehmenden konnten das Video durch einen Zwischenruf unterbrechen. Nach einer Unterbrechung wurde im Plenum diskutiert, inwiefern eine der Betreuungsregeln verletzt wurde und wie eine autonomieunterstützende Handlungsalternative aussehen könnte (Reeve, 1998).

Die Teilnehmenden konnten im ersten Teil der Übungsphase in Kleingruppen jede der vier Betreuungsregeln an jeweils einer Station erarbeiten und einüben. Je nach Gesamtzahl der Interventionsteilnehmenden wurden alle Stationen mehrfach angeboten. Jede Station umfasste vier Aufgaben (Siehe Kapitel 13.Interventionsmaterialien; Für die aktuellste Version der Materialien siehe Schlake et al., 2020). Anfänglich sollten die Teilnehmenden eine Checkliste lesen, auf welcher die Merkmale autonomieunterstützender bzw. kontrollierender Verhaltensweisen dargestellt wurden. Anschließend konnten die Teilnehmenden

gemeinsam über die Checkliste diskutieren und sich je eine prototypisch positive bzw. negative Videosequenz anschauen, in welcher die jeweilige Betreuungsregel erfolgreich bzw. mangelhaft umgesetzt wurde. Als zweite Aufgabe identifizierten die Teilnehmenden, welche Videosequenz das positive und welche das negative Beispiel war (siehe auch Berger et al., 2017). Es wurde darauf geachtet, dass die Identifikation nicht anhand oberflächlicher Merkmale (z. B. an den Schauspielenden oder dem generellen Umgangston) möglich war. Die Teilnehmenden konnten gemeinsam diskutieren, auf Basis der Checkliste eine Entscheidung fällen und diese mit einer Lösungskarte abgleichen. Als Drittes konnten die Teilnehmenden in Anlehnung an Reeve (1998) kontrollierende Verhaltensweisen aus der prototypischen Videosequenz auswählen und auf einer Fehlerliste Verbesserungsvorschläge schriftlich festhalten. Diese Verbesserungsvorschläge wurden in den Kleingruppen entwickelt und verschriftlicht. Abschließend konnten die Teilnehmenden das gelernte Wissen in einem Rollenspiel anwenden. Diese praktische Übungsphase wird von der Metaanalyse zur Wirksamkeit solcher Intervention empfohlen (Su & Reeve, 2011). Da die Tutoren möglicherweise Schwierigkeiten bei der Durchführung von Rollenspielen haben (Berger et al., 2017), wurden den Teilnehmenden Rollenkarten (Betreuer-, Schauspiel- und Beobachterkarten) zur Verfügung gestellt. Nach ca. drei Minuten sollten die Rollen getauscht werden, um zu gewährleisten, dass alle Teilnehmenden eine aktive Rolle übernehmen. Nach insgesamt 40 Minuten wechselte jede Gruppe zu einer weiteren Station. In den eingesetzten Videos konnten die Interventionsteilnehmenden sowohl männliche als auch weibliche Tutoren sehen, die ihre Tutees autonomieunterstützend bzw. kontrollierend betreuen. Außerdem konnten sie sowohl männliche als auch weibliche Tutees sehen, die autonomieunterstützend und kontrollierend betreut wurden. Diese Konstruktion hatte zum Ziel, dass die Intervention für alle Teilnehmenden unabhängig von ihrem Geschlecht gleich wirkt.

Der zweite Teil der Übungsphase fand im Plenum statt und beinhaltete ein Rollenspiel zwischen der Seminarleitung und wechselnden Teilnehmenden (Rollenspiele mit Beobachtern). Die Seminarleitung stellte einen Tutee dar und konfrontierte die Teilnehmenden mit standardisierten Betreuungsproblemen (Die Liste der standardisierten Betreuungsprobleme können den Pilotierungs- bzw. Interventionsmaterialien entnommen werden, siehe Kapitel 13. Interventionsmaterialien sowie Schlake et al., 2020). Für diese Probleme wurde jeweils direkt im Anschluss mit allen Teilnehmenden diskutiert, wie der betreuende Teilnehmende gehandelt hat und ggf. wie autonomieunterstützende Handlungsalternativen aussehen könnten.

In der Reflexionsphase hat die Seminarleitung mit einem Mitarbeiter der Universität Duisburg-Essen ein Rollenspiel vorgespielt und jede Betreuungsregel einmalig missachtet. Die Teilnehmenden wurden dazu aufgefordert, bei jedem Regelbruch „Stop“ zu rufen. Danach wurde im Plenum ähnlich der zweiten Übungsphase diskutiert, welche autonomieunterstützende Handlungen möglich und sinnvoll gewesen wären. Anschließend wurden analog zu Reeve (1998) und entsprechend der Empfehlung von Reeve et al. (2004), Reeve et al. (2014) sowie Reeve und Cheon (2016) die Effektivität und die Einfachheit der Umsetzung des autonomieunterstützenden und des kontrollierenden Ansatzes diskutiert.

Anpassungen durch die Pilotierung

Die Intervention umfasste in der Pilotierung zwei Tage. Nach der Inputphase und der anschließenden Mittagspause begann der erste Teil der Übungsphase. Die Teilnehmenden konnten zwei Betreuungsregeln an zwei Stationen erarbeiten. Daran schlossen sich die Rollenspiele im Plenum an. Nach dieser Plenumsphase konnten die Teilnehmenden eine dritte Betreuungsregel erarbeiten. Am Abend des ersten Tages wurde ein Spiel gespielt. Die Teilnehmenden haben in zwei Gruppen Fragen beantwortet und dafür Punkte erhalten. Das

Team mit den meisten Punkten gewann das Spiel. Bezüglich jeder Betreuungsregel standen fünf unterschiedlich schwierige Fragen zur Auswahl und die Teams konnten in ihrem Zug eine Frage auswählen. Am Morgen des zweiten Tages wurde die vierte Betreuungsregel erarbeitet. Der zweite Tag des Pilotierungsdurchlaufes der Intervention endete am Vormittag mit der Reflexionsphase und dem Abschlussgespräch.

Mit insgesamt 32 Schülerinnen und Schülern wurde die Intervention erprobt. Dabei sollten vor allem Unterschiede zwischen der Zeitplanung und der Umsetzung, zu komplexe oder missverständliche Materialien und falsch gelernte Inhalte identifiziert werden. Sechs der 32 Teilnehmenden wurden etwa einen Monat später bei der Betreuung je eines Fünftklässlers oder einer Fünftklässlerin gefilmt. Aus den Videos und Beobachtungen während der Durchführung des Seminars ergaben sich folgende Änderungen:

Einerseits waren die Zeiten für die erste Inputphase und die Stationsarbeitsphasen zu lang angesetzt. Statt 50 Minuten wurden für die erste Inputphase nur 30 Minuten eingeplant. Die Zeit pro Station wurde von 40 Minuten auf 30 reduziert. Weiterhin war in der Pilotierung keine Zeit zwischen den Phasen eingeplant. Da teilweise Raumwechsel nötig waren, wurden nach der Pilotierung Übergangszeiten ergänzt. Die Teilnehmenden haben während des Abschlussgesprächs geäußert, dass sie zuerst alle Stationen bearbeiten und anschließend die Plenumsübungsphase durchführen möchten. Deshalb wurde dem Wunsch der Teilnehmenden nachgekommen, sodass die Interventionen der Hauptstudie nur einen Tag umfassten und alle vier Stationen im Block erarbeitet wurden. Weiterhin wurde das oben beschriebene Spiel ersatzlos gestrichen.

Neben den Anpassungen des Ablaufs musste auch das Material angepasst werden. Bei der Pilotierung der Intervention lautete die erste Regel *Ermutige bei Schwierigkeiten und lobe bei Erfolg*. Die Teilnehmenden haben das Loben der Tutees entsprechend dieser Regel übertrieben, wodurch das Lob nicht aufrichtig wirkte. Weiterhin haben sich die Teilnehmenden teilweise so sehr zurückgehalten, dass sie ihren Tutees bei genereller Unsicherheit emotional nicht geholfen haben. Dies zeugt von einem Missverständnis der Regeln eins und drei. Deshalb wurde Regel eins von *Ermutige bei Schwierigkeiten und lobe bei Erfolg* zu *Gib Sicherheit und ermutige bei Schwierigkeiten* umformuliert. Dies soll hervorheben, dass eine emotionale Unterstützung bei Unsicherheiten keine Unterbrechung selbstregulierten Handelns darstellt. Weiterhin liegt der Fokus durch die Änderung nicht mehr auf Lob. Auf der Checkliste der ersten Betreuungsregel wurde der Satz *Als Betreuer gibst du nicht nur inhaltliche Hilfe, sondern sorgst auch dafür, dass dein Schüler sich verstanden und wohl fühlt* ergänzt. Weiterhin wurde der Stichpunkt ergänzt, dass die Tutoren ihren Tutees ein Gefühl von Sicherheit geben sollen, um anfängliche Angst vor dem Inhalt oder der Situation abzubauen. Damit diese Ergänzung nicht zu Monologen der Tutoren führt, wurde ebenfalls ein erklärender Hinweis auf der Checkliste ergänzt. In der überarbeiteten Checkliste der dritten Regel wird hervorgehoben, dass inhaltliche Hilfen zurückhaltend gegeben werden sollen, emotionale Hilfen aber jederzeit in Ordnung sind. Konkret wurde der Satz *Hilf deinem Schüler nur dann, wenn er nicht arbeitet oder dich um Hilfe bittet. Ansonsten halte dich zurück und beobachte nur* zu *Hilf deinem Schüler inhaltlich nur dann, wenn er nicht arbeitet oder dich um Hilfe bittet. Wenn dein Schüler generell unsicher ist oder Angst hat, kannst du ihn aber immer emotional unterstützen* geändert.

8.4 Konzeption und Entwicklung der Messinstrumente

Im Folgenden soll die Entwicklung der genutzten Messinstrumente dargelegt werden. Da die Motivierstile einen zentralen Teil der vorliegenden Forschungsarbeit darstellen, wird die Entwicklung der Instrumente zur Erfassung der Motivierstile ausführlich dargestellt (siehe Kapitel 8.4.1 - 8.4.3). Das Instrument zur Erfassung der Zielvariablen auf Tutorebene (intrinsische Motivation nach der Durchführung des Experiments) wird ebenfalls ausführlich beschrieben. Die Entwicklung und Pilotierung der übrigen Instrumente wird in Kapitel 8.4.5 zusammenfassend dargestellt, wobei eine detaillierte Darstellung im Anhang (siehe Kapitel 12.1) zu finden ist. Abschließend werden in Kapitel 8.4.6 Indizien konvergenter Validität vorgestellt. Die Reliabilität der Fragebögen wird im Rahmen dieser Arbeit unter anderem anhand der internen Konsistenz bewertet (Field, 2009). Interne Konsistenzen unter .70 können als mangelhaft bewertet werden (Blanz, 2015; Bühner, 2006; Field, 2009). Interne Konsistenzen über .70 können als ausreichend, Werte über .80 als gut und Werte über .90 als exzellent klassifiziert werden (Blanz, 2015; Bühner, 2006; Field, 2009). Itemselektivitäten über .30 werden als ausreichend und Werte über .60 als gut interpretiert (Clark & Watson, 1995; Field, 2009; Stevens, 2001).

8.4.1 Das Instrument zur Erfassung der selbstberichteten Motivierstile für Cross-Age-Tutoren

Durch eine Vielzahl an Studien haben sich zwei Instrumente zur Erfassung der selbstberichteten Motivierstile etabliert (Cheon & Reeve, 2013; Cheon et al., 2012; Cheon et al., 2014; Deci & Ryan, 1985; Deci et al., 1981; Reeve, 1998; Reeve et al., 1999; Reeve & Cheon, 2016; Reeve et al., 2014). Im Folgenden wird aus den Stärken und Schwächen dieser Instrumente abgeleitet, wie ein für Tutoren passendes Instrument zu konstruieren ist.

Problems in Schools Questionnaire (Deci et al., 1981): Der Fragebogen beinhaltet acht Vignetten, die motivationsbezogene Probleme von Schülerinnen und Schülern im Schulkontext darstellen. An jede Vignette schließen vier Möglichkeiten an, wie die Befragten als Lehrkraft mit diesem Problem umgehen können. Diese vier Möglichkeiten beinhalten immer eine sehr autonomieunterstützende, eine moderat autonomieunterstützende, eine moderat kontrollierende und eine sehr kontrollierende Handlungsoption. Die von den Probanden empfundene Angemessenheit der vier Handlungsoptionen kann bei jeder Vignette auf einer 7-stufigen Likert-Skala bewertet werden. Jede Subskala ergibt sich als das ungewichtete Mittel der acht Items. Der bilanzierte Motivierstil ergibt sich als gewichtetes Mittel der vier Subskalen (Deci et al., 1981; Reeve et al., 1999).

Studien bestätigen die zeitliche Stabilität und die konvergente, divergente sowie prädiktive Validität des bilanzierten Motivierstils von Lehrkräften (Deci & Ryan, 1985; Deci et al., 1981; Reeve, 1998; Reeve et al., 1999). Allerdings weist die Subskala moderat autonomieunterstützender Handlungsoptionen Validitätsprobleme auf (Reeve et al., 1999). Unter anderem deshalb sehen aktuelle Studien von einem Einsatz des Instrumentes ab (Reeve et al., 2018; Reeve et al., 2014), weshalb eine Adaption der drei validen Subskalen dieses Instruments lediglich zur Untersuchung der konvergenten Validität der selbst entwickelten Instrumente eingesetzt wird (siehe Kapitel 8.4.6). Weiterhin beinhaltet der Fragebogen Items, welche in Deutschland vermutlich nicht verstanden werden. Beispielsweise beinhaltet die Situationsbeschreibung einer Vignette eine Pfadfinder-Buchstabier-Gruppe, die in Deutschland nicht üblich ist, weshalb diese Items in der Adaption angepasst wurden.

Der *Problems In Schools* Fragebogen behandelt typische Probleme des Schulalltags von Lehrerinnen und Lehrern und ist deshalb nicht ohne weitere Anpassungen für Cross-Age-Tutoren nutzbar. Für diese Zielgruppe müsste ein auf die Tutorentätigkeit angepasstes Instrument entwickelt werden (Deci et al., 1989). Aufgrund der Validitätsprobleme der moderat autonomieunterstützenden Subskala, sollte eine Adaption lediglich die validen Subskalen enthalten (Reeve et al., 1999).

Teachers Scenario Measure (Reeve & Cheon, 2016; Reeve et al., 2014): Der Fragebogen besteht aus zwei Vignetten, welche eine sehr autonomieunterstützende, beziehungsweise kontrollierende Unterrichtsvorbereitung und -durchführung beschreiben. Daran schließen sich Items zur Erfassung der persönlichen Befürwortung sowie der Intentionen entsprechend zu handeln an. Diese Intention stellt einen Indikator der Motivierstile dar (Reeve & Cheon, 2016; Reeve et al., 2014). Sowohl das autonomieunterstützende als auch das kontrollierende Szenario orientieren sich an den Verhaltensweisen und Verhaltensgruppen, die in Kapitel 4 beschrieben wurden.

Ähnlich wie der *Problems In Schools* Fragebogen ist der *Teachers Scenario Measure* spezifisch für Lehrerinnen und Lehrer geeignet. Somit ist auch dieser Fragebogen nur eingeschränkt für Cross-Age-Tutoren einsetzbar.

Die Konzeption und geplante Auswertung eines neuen Instrumentes zur Erfassung der selbstberichteten Motivierstile von Cross-Age-Tutoren

Der *Teachers Scenario Measure* besteht nur aus einer Vignette pro Motivierstil. Laut Gollwitzer (2014) sind Intentionsfragebögen dann prädiktiv für tatsächliches Verhalten, wenn sie konkrete Handlungssituationen beinhalten und konkrete Implementationsintentionen erhoben werden. Folglich ist entsprechend Gollwitzer (2014) der *Problems In Schools* Fragebogen als prädiktiv valider zu interpretieren, weil er mehr und vor allem konkretere Handlungssituationen beinhaltet. Somit weisen beide Fragebögen Stärken und Schwächen auf. Ein Fragebogen für Cross-Age-Tutoren müsste 1.) in Anlehnung an Gollwitzer (2014), ähnlich wie der *Problems In Schools* Fragebogen, mehrere Vignetten mit konkreten Handlungsoptionen beinhalten und sich 2.) ähnlich wie der *Teachers Scenario Measure* an nachgewiesenen autonomieunterstützenden und kontrollierenden Verhaltensweisen orientieren. Darüber hinaus sollte sich dieser neue Fragebogen 3.) an der konkreten Lehr-Lern-Situation der Cross-Age-Tutoren orientieren, z. B. der Betreuung von naturwissenschaftlichen Experimenten (Berger et al., 2017). Weiterhin sollte der Fragebogen 4.) eine separate Betrachtung der Motivierstile ermöglichen.

Entsprechend der ersten und dritten Anforderung besteht das neu konzipierte Instrument aus zwölf Vignetten. Diese Vignetten bestehen immer aus einer Situationsbeschreibung und Handlungsoptionen für den Cross-Age-Tutor. Die Situationsbeschreibungen stellen typische Problemsituationen eines Cross-Age-Tutors bei der Betreuung naturwissenschaftlicher Experimente dar. Sie wurden auf Basis von Erfahrungsberichten von Cross-Age-Tutoren sowie der Analyse von videografierten Tutoringsitzungen formuliert. Die Handlungsoptionen wurden entsprechend der zweiten Anforderung an den autonomieunterstützenden und kontrollierenden Verhaltensgruppen (I-IV) orientiert (siehe Kapitel 4 und 8.3). Es gibt pro Vignette zwei mal zwei Handlungsoptionen. Jedes Paar von Handlungsoptionen stellt eine prototypisch autonomieunterstützende und eine prototypisch kontrollierende Handlungsoption dar. Jede Vignette deckt somit jeweils zwei Kategorien autonomieunterstützenden und zwei Kategorien kontrollierenden Verhaltens ab. Bei vier betrachteten Verhaltensgruppen gibt es sechs Kombinationsmöglichkeiten der Kategorien. Jede Kombination ist im Fragebogen zweifach repräsentiert.

Pilotierung

Aufbauend auf eine Präpilotierung, wurde der Fragebogen zur Erhebung der selbstberichteten Motivierstile an insgesamt 138 Schülerinnen und Schülern erprobt. Beide Subskalen weisen mit einer internen Konsistenz von $\alpha_{SMot_A} = .80$ und $\alpha_{SMot_K} = .80$ und Itemselektivitäten zwischen .33 und .60 akzeptable bis gute Kennwerte auf. Einzige Ausnahme ist das erste Item der kontrollierenden Subskala, welches eine Itemselektivität von .02 aufweist. Bei der Erhebung wurde ebenfalls der *Problems In Schools* Fragebogen eingesetzt. Die daraus resultierenden Indizien der konvergenten Validität werden in Kapitel 8.4.6 diskutiert.

Diskussion der Gütekriterien

Die Objektivität der Erfassung der selbstberichteten Motivierstile der Tutoren kann aufgrund des standardisierten Testverfahrens in Form eines Fragebogens sowie der standardisierten Testprozedur angenommen werden. Wie in Kapitel 8.1 beschrieben wurde, wurden die Instruktionen für die Tutoren wortgenau vorgegeben.

Die internen Konsistenzen der Skalen des autonomieunterstützenden und des kontrollierenden Motivierstils waren während der Pilotierung zufriedenstellend. Die internen Konsistenzen während der Hauptstudie werden im Anschluss an eine Hauptkomponentenanalyse in Kapitel 9.2 diskutiert. Die Test-Retest Reliabilität wurde in Anlehnung an die gute Test-Retest-Reliabilität des *Problems In Schools* Fragebogen angenommen (Deci et al., 1981).

Im vorhergehenden Unterkapitel wurde die Konstruktion des Instruments erläutert, aus der Indizien prädiktiver Validität (mehrere Vignetten mit konkreten Handlungsoptionen; siehe Gollwitzer, 2014), Inhaltsvalidität (Orientierung an nachgewiesenen autonomieunterstützenden und kontrollierenden Verhaltensweisen) sowie Konstruktvalidität (konkrete Lehr-Lern-Situation der Cross-Age-Tutoren; sowie separate Betrachtung der Motivierstile) folgen. Weitere Hinweise auf konvergente Validität können Kapitel 8.4.6 entnommen werden.

8.4.2 Das Instrument zur Erfassung der beobachteten Motivierstile von Cross-Age-Tutoren durch Verhaltensanalyse

Die Motivierstile von Individuen können auch erhoben werden, indem Experten gezeigtes Verhalten in einer Betreuungssituation beobachten und analysieren (Sarrazin et al., 2006). Dabei wird das Verhalten entweder von den Experten während der konkreten Betreuungssituation beobachtet und fortlaufend analysiert oder aufgezeichnet und anschließend ausgewertet (Berger et al., 2017; Reeve et al., 2004). Zwei Verfahren zur Verhaltensanalyse sind üblich (Berger et al., 2017; Deci et al., 1982). Im Rahmen des ersten Verfahrens, wird das beobachtete Verhalten bezüglich Verhaltensgruppen (siehe Tabelle 4.1) auf einer Likert-Skala bewertet. Bei dem zweiten Verfahren werden kontrollierende und autonomieunterstützende Verhaltensweisen (siehe Tabelle 4.1) kodiert. Aus den Auftretenshäufigkeiten dieser Verhaltensweisen werden die Skalen der Motivierstile gebildet.

Bewertung von Verhaltensgruppen

Berger et al. (2017), Cheon et al. (2012), Cheon und Reeve (2013), Cheon et al. (2014), Kaur et al. (2015), Reeve et al. (2004), Tessier et al. (2010) bewerteten in ihren Studien die in Kapitel 4 vorgestellten Verhaltensgruppen auf Likert-Skalen. Dabei berücksichtigten die Studien nicht alle Verhaltensgruppen, sondern beschränkten sich jeweils auf einzelne.

Beispielsweise beschränkten sich Berger et al. (2017) bei der Bewertung darauf, inwiefern die Tutoren ihren Tutees einen hinreichenden Sprachanteil und genügend Zeit zum selbstständigen Arbeiten gaben. In Anlehnung an die ursprüngliche eindimensionale Definition eines bilanzierten Motivierstils wurden in den oben genannten Studien die Verhaltensgruppen so bewertet, dass ein hoher Wert ein autonomieunterstützendes und ein niedriger Wert ein kontrollierendes Verhalten darstellt. In den genannten Studien wurden die Verhaltensgruppen (z. B. das Verwenden einer autonomieunterstützenden oder kontrollierenden Sprache) entweder einzeln ausgewertet oder ungewichtet zu einem Mittelwert zusammengefasst (Cheon et al., 2014; Reeve et al., 2004). Dieser Mittelwert kann als bilanzierter Motivierstil interpretiert werden (Cheon et al., 2014; Reeve et al., 2004).

Durch diese Art der Verhaltensanalyse können zwar die Verhaltensgruppen und der bilanzierte Motivierstil, aber nicht der autonomieunterstützende und kontrollierende Motivierstil getrennt erhoben und ausgewertet werden.

Kodierung der Häufigkeiten und der Dauer beobachteter Verhaltensweisen

Deci et al. (1982), Reeve und Jang (2006), Reeve et al. (1999), Sarrazin et al. (2006) und Tessier et al. (2008) kodierten die Auftretenshäufigkeiten und die Dauer bestimmter Verhaltensweisen. Deci et al. (1982), Reeve und Jang (2006) und Reeve et al. (1999) untersuchten, welche der kodierten Verhaltensweisen als autonomieunterstützend und welche als kontrollierend wahrgenommen werden. Sarrazin et al. (2006) und Tessier et al. (2008) kodierten Äußerungen von Lehrerinnen und Lehrern, um einen kontrollierenden und einen autonomieunterstützenden Motivierstil zu bilden. Diese so ermittelten Motivierstile wurden zur Beantwortung weiterer Forschungsfragen in Analysen berücksichtigt.

Mit der Analyse konkreter Verhaltensweisen lassen sich der autonomieunterstützende und der kontrollierende Motivierstil getrennt erheben. Insbesondere sind detailliertere Einblicke in die Bedeutung einzelner Verhaltensweisen möglich.

Erfassung der Motivierstile von Cross-Age-Tutoren durch Verhaltensanalyse

Abhängig vom Ziel einer Studie könnten beide der oben genannten Auswertungsverfahren für beobachtetes Verhalten angebracht sein. Sollen nur ein bilanzierter Motivierstil oder Verhaltensgruppen ausgewertet werden, so ist eine Bewertung der Verhaltensgruppen ausreichend. Sollen die Motivierstile getrennt ausgewertet werden, müssen die auftretenden Verhaltensweisen kodiert werden (Sarrazin et al., 2006; Tessier et al., 2008). Entsprechend den Hypothesen der vorliegenden Forschungsarbeit (siehe Kapitel 7), sollten die Motivierstile getrennt erhoben werden, weshalb die autonomieunterstützenden und kontrollierenden Verhaltensweisen der Cross-Age-Tutoren in einer Betreuungssituation kodiert wurden. Zu diesem Zweck wurde basierend auf der Übersicht autonomieunterstützender und kontrollierender Verhaltensweisen (siehe Tabelle 4.1) ein Kodiermanual erstellt.

Der Aufbau des Kodiermanuals zur Erfassung der Motivierstile von Cross-Age-Tutoren durch Verhaltensanalyse

Das Kodiermanual umfasst insgesamt 23 Seiten und beginnt mit allgemeinen Informationen und Kodierregeln. Daran schließt sich eine tabellarische Auflistung der zu kodierenden Verhaltensweisen an. Tabelle 8.3 stellt eine Übersicht der zu kodierenden Verhaltensweisen dar. Die zu kodierenden Verhaltensweisen sind vier Gruppen autonomieunterstützenden und kontrollierenden Verhaltens (T = Arbeiten im selbstbestimmten Tempo; W = Wahl- und Handlungsfreiheit; S = Sprache; E = Umgang mit Emotionen; siehe Kapitel 4) oder einer Hilfskategorie (H = Hilfebedarf des Tutee) zugeordnet.

Kürzel	Verhaltensweise	Subjekt	Verhaltensgruppe
H1	Tutee fragt nach einem/ bittet um einen Lösungshinweis	Tutee	Hilfskategorie –Tuteeprobleme
H2	Tutee fragt nach einem/ bittet um einen Handlungshinweis	Tutee	
H3	Tutee bittet um einen physischen Eingriff ins Experiment	Tutee	
H4	Tutee signalisiert verbal eine inhaltliche oder sozial-emotionale Schwierigkeit	Tutee	
T1	Der Betreuer greift verbal beziehungsweise nonverbal ein	Tutor	Geduld statt Ungeduld für Arbeiten im
T2	Der Betreuer greift physisch ein	Tutor	selbstbestimmten Tempo
W	Geben eines Hinweises (Lösungs- oder Handlungshinweis)	Tutor	Freiräume und Hinweise statt vorgegebene Lösungs- und Handlungsschritte
S1	„Musst“-/ „Sollst“- Aussagen	Tutor	Verwenden einer autonomie-
S2	Lenkende Fragen	Tutor	unterstützenden oder
S3	Imperative	Tutor	kontrollierenden Sprache
S4	Nicht-kontrollierende Aussagen/Fragen	Tutor	
E1	Ermutigung bei Schwierigkeiten	Tutor	Umgang mit Äußerungen
E2	Verständnis bei Schwierigkeiten	Tutor	negativer Gefühle
E3	Auf Tutee eingehen	Tutor	
E4	Lob für Fortschritt und Erfolg	Tutor	

Tabelle 8.3: Die zu kodierenden Verhaltensweisen

Der Kodierprozess zur Erfassung der beobachteten Verhaltensweisen

Die Tutorsitzungen wurden aufgezeichnet, sodass die oben beschriebenen Verhaltensweisen nachträglich kodiert werden konnten (Siehe Kapitel 8.1). Die Videos wurden in 10-Sekunden-Intervalle unterteilt (Seidel et al., 2003). Anhand von sechs Präpilotierungsvideos wurde die Länge der Intervalle an der erwarteten Frequenz der zu kodierenden Verhaltensweisen orientiert. Die Segmentierung sowie die Auswertung der Videos wurden mit der Software Mangold INTERACT durchgeführt (Mangold, 2020).

Für jedes Intervall wurde entschieden, ob die Tutees ein Problem geäußert haben (H1-H4) und ob und inwiefern der Tutor eingegriffen hat (T1-T2). Hat ein Tutor verbal beziehungsweise nonverbal (T1) oder physisch (T2) ins Experiment eingegriffen, wurde der gesamte fortlaufende Eingriff als ein Event kodiert. Hat eine solche Äußerung des Tutees oder ein solcher Eingriff des Tutors beziehungsweise somit ein Event länger als ein Intervall gedauert oder hat die Intervallgrenze überschritten, wurde der Kode nur im ersten Intervall gesetzt und für alle weiteren Intervalle der Kode „7!“ gesetzt. Die Kategorien W, S und E wurden nur ein Mal pro Event kodiert, also in den Intervallen, in denen der Tutor entweder

physisch oder verbal beziehungsweise nonverbal eingegriffen hat, da die Kategorien W, S und E einen Eingriff voraussetzen. Deshalb wurden in einem zweiten und dritten Kodierdurchlauf nur die Intervalle erneut betrachtet, in denen ein Event begann oder fortgeführt wurde. Im zweiten Kodierdurchlauf wurden die Kategorien W sowie E1-E4 und im dritten Kodierdurchlauf die Kategorien S1-S4 kodiert. Nach dem dritten Kodierdurchlauf hat jeder Kodierer seine Kodierdatei auf formale Fehlkodierungen hin überprüft. Hat ein Kodierer einen Fehler entdeckt, hat er das fragliche Intervall erneut angeschaut und die Codes neu gesetzt. Ein typischer Fehler war, dass durch das unbeabsichtigte Auslösen eines Tastenkürzels Codes bei W, S oder E gesetzt wurden, obwohl kein Eingriff des Tutors kodiert wurde. Es wurde also zum Beispiel kodiert, dass der Tutor einen Imperativ benutzt hat und gleichzeitig nicht gesprochen hat. Diese Korrektur ist von einer nachträglichen Änderung der Kodierentscheidungen abzugrenzen. Lag kein formaler Kodierfehler vor, wurden die gesetzten Codes nicht abgeändert.

Auswertung kodierter Verhaltensweisen zur Berechnung der Motivierstile von Cross-Age-Tutoren

Die meisten der kodierten Verhaltensweisen können als autonomieunterstützend oder kontrollierend bezeichnet werden. Dies trifft jedoch nicht auf alle Verhaltensweisen zu. Das folgende Beispiel soll diese Fälle und die geplante Auswertung dieser Verhaltensweisen veranschaulichen:

Intervall	Beispiel 1								Beispiel 2							
	H1	H2	H3	H4	T1	T2	W	...	H1	H2	H3	H4	T1	T2	W	...
10	0	0	0	0	0	0	0	...	1	0	0	0	0	0	0	...
11	0	0	0	0	2	0	1	...	0	0	0	0	1	0	1	...

Tabelle 8.4: Kodierbeispiel

Die Tabelle 8.4 zeigt zwei fiktive Kodierausschnitte. In beiden Beispielen wurden zwei Intervalle kodiert und in beiden Beispielen gibt der Tutor im elften Intervall einen inhaltlichen Hinweis (W hat den Code „1“). Wie in Kapitel 4. erläutert, ist dies je nach Hilfsbedürftigkeit des Tutees entweder autonomieunterstützend oder kontrollierend. Im ersten Beispiel benötigte der Tutee keine Hilfe (H1 – H4 haben den Code „0“) und der Tutor sagte etwas, obwohl der Tutee kein Problem geäußert hat (T1 hat den Code „2“). Somit kann der ungewünschte inhaltliche Hinweis als kontrollierend eingestuft werden. Im zweiten Beispiel hat der Tutor den inhaltlichen Hinweis als Reaktion auf eine Problemäußerung des Tutees gegeben (T1 hat den Code „1“, weil H1 zuvor den Code „1“ hatte), weshalb der Hinweis als autonomieunterstützend eingestuft werden kann. Ob der Tutor Freiräume und Hinweise gibt, statt Lösungs- und Handlungsschritte vorzugeben, kann also weder an T noch an W allein abgelesen werden. Die Auswertungskategorie ergibt sich aus einer Kombination von T und W. In Worten bedeutet die Auswertungskategorie $TW1_K$: Wenn bei W der Code „1“ gesetzt ist, dann wird die Verhaltensweise als kontrollierend gewertet, wenn bei T1 der Code „2“ gesetzt ist. Andersfalls wird das Verhalten dann in Kategorie $TW1_A$ als autonomieunterstützend gewertet, wenn bei W der Code „1“ und bei T1 der Code „1“ gesetzt wurde (siehe 8.5). Entsprechend gilt dies für $TW2_A$ und $TW2_K$.

Weiterhin lassen sich aus den Beobachtungskategorien weitere Auswertungsmaße berechnen, wie beispielsweise die Anzahl der Intervalle, in denen der Tutor kommunizierte. Aus den oben vorgestellten beobachteten Verhaltensweisen und den Codes in den entsprechenden Beobachtungskategorien werden zur weiteren Auswertung folgende Auswertungskategorien berechnet.

Auswertungskategorien	Kürzel Auswertungskategorie	Kürzel Beobachtungskategorie	Berechnungsvorschrift
Autonomieunterstützung			
Nonverbale/ Verbale inhaltliche Eingriffe, wenn der Tutee Hilfe nötig hatte (H1-H4)	TW1 _A	T1 und W	1, wenn W = „1“ und T1 = „1“ oder „3“, andernfalls, 0
Physische inhaltliche Eingriffe, wenn der Tutee Hilfe nötig hatte (H1-H4)	TW2 _A	T2 und W	1, Wenn W = „1“ und T2 = „1“, andernfalls, 0
Ermutigung bei Schwierigkeiten	E1	E1	Entspricht E1
Verständnis bei Schwierigkeiten	E2	E2	Entspricht E2
Bestätigend auf Aussagen des Tutees reagieren	E3	E3	Entspricht E3
Nutzung von Lob bei Fortschritt und Erfolg	E4	E4	Entspricht E4
Kontrolle			
Nonverbale/ Verbale inhaltliche Eingriffe, obwohl der Tutee keine Hilfe nötig hatte (H1-H4)	TW1 _K	T1 und W	1, wenn W = „1“ und T1 = „2“ andernfalls, 0
Physische inhaltliche Eingriffe, obwohl der Tutee keine Hilfe nötig hatte (H1-H4)	TW2 _K	T2 und W	1, Wenn W = „1“ und T2 = „2“ andernfalls, 0
Nutzung von druckerzeugenden Modalverben	S1	S1	Entspricht S1
Nutzung von lenkenden Fragen	S2	S2	Entspricht S2
Nutzung von Imperativen	S3	S3	Entspricht S3
Weitere Auswertungsmaße			
Redezeit	RZ	T1	Es werden alle Intervalle gezählt, in denen der Tutor verbal oder nonverbal kommuniziert
Eingriffszeit	EZ	T2	Es werden alle Intervalle gezählt, in denen der Tutor physisch in das Experiment eingreift
Eventzeit	Eventzeit	T1 und T2	Es werden alle Intervalle gezählt, in denen der Tutor verbal beziehungsweise nonverbal kommuniziert oder physisch ins Experiment eingreift

Tabelle 8.5: Die aus den Beobachtungskategorien resultierenden Auswertungskategorien

Die Auswertungskategorien liefern Häufigkeiten autonomieunterstützender und kontrollierender Verhaltensweisen. Je häufiger sich Tutoren kontrollierend verhalten, desto stärker wird das Autonomieerleben der Tutees vermutlich frustriert (Reeve & Jang, 2006). Ebenso wird angenommen, dass das Autonomieerleben der Tutees umso mehr befriedigt wird, je häufiger die Tutoren autonomieunterstützendes Verhalten zeigen (Reeve & Jang, 2006).

Die Dauer der Tutorsitzungen und damit die Gelegenheit für das Auftreten der Verhaltensweisen waren unterschiedlich. Folglich war die Anzahl der 10-Sekunden-Intervalle pro Video unterschiedlich. Damit die Tutoren untereinander verglichen werden können,

wurden ähnlich wie bei Aelterman et al. (2014) die absoluten Häufigkeiten aller Auswertungskategorien durch die Anzahl der 10-Sekunden-Intervalle des Videos geteilt. Durch diese Rechnung ergaben sich auf die auf Videolänge normierten Werte der Auswertungskategorien. Weiterhin könnten die Auswertungskategorien natürlich unterschiedlich häufig vorkommen, wodurch häufige Kategorien bei der ungewichteten Zusammenfassung relativ mehr Gewicht bekommen. Um dafür zu korrigieren und die Auswertungskategorien zu einer Skala zusammenfassen zu können, wurde jede Auswertungskategorie an der Gruppe normiert (Sarrazin et al., 2006). Ist bei einem Tutor eine Auswertungskategorie am seltensten vorgekommen, beträgt der Wert des Tutors in der Kategorie nach diesen beiden Normierungen null. Ist bei einem Tutor eine Auswertungskategorie am häufigsten vorgekommen, beträgt der Wert des Tutors in der Kategorie nach diesen beiden Normierungen eins. Die Werte aller Tutoren in einer Kategorie ergeben sich als Mittel gemäß der Formel:

$$\begin{aligned} & \text{Gruppennormierter Wert}_{\text{Kategorie, Tutor}_i} = \\ & = \frac{\text{Längennormierter Wert}_{\text{Kategorie, Tutor}_i} - \text{Längennormierter Wert}_{\text{Kategorie, Minimum}}}{\text{Längennormierter Wert}_{\text{Kategorie, Maximum}} - \text{Längennormierter Wert}_{\text{Kategorie, Minimum}}} \end{aligned}$$

Im Folgenden werden auf Videolänge und relative Ausprägung normierte Kategorien mit dem Suffix „_soz“ gekennzeichnet. Entsprechend Sarrazin et al. (2006) sowie Tessier et al. (2008) können aus diesen angepassten Auswertungskategorien die Motivierstile berechnet werden. Die weiteren Auswertungsmaße (z. B. die Redezeit) können ebenfalls für Analysen herangezogen werden, allerdings können sie nicht zur Berechnung der Motivierstile genutzt werden. Beispielsweise sind die Redezeit und die Nutzung von Imperativen abhängig voneinander, da jede Nutzung eines Imperativs das Reden des Tutors voraussetzt.

Vorgehen bei der Konstruktion und Pilotierung

Das Kodiermanual wurde als Kodierleitfaden entsprechend der qualitativen Inhaltsanalyse nach Mayring (2010) entwickelt. Bei der Konstruktion der ersten vorläufigen Version des Kodiermanuals wurde sich im besonderen Maße an Reeve und Jang (2006) orientiert (siehe auch Tabelle 4.1). Die Kategorienbildung fand also überwiegend deduktiv statt. Durch die Erprobung wurden allerdings Anpassungen nötig, weshalb einzelne Kategorien induktiv entstanden sind. Beispielsweise wurden Hinweise zu deklarativem Wissen in *W1* und Hinweise zum experimentellen Vorgehen in *W2* getrennt kodiert. Dieses vorläufige Kodiermanual sollte an wenigen Videos pilotiert werden, die Tutoren bei der Betreuung des Hauptstudienexperimentes zeigen (siehe Kapitel 8.2). Es haben sich sechs Schülerinnen und Schüler bereit erklärt, die an der Pilotierung der Intervention teilgenommen haben (siehe Kapitel 8.3). Diese haben je eine Schülerin oder einen Schüler der fünften Klasse bei der Durchführung des Experiments betreut. Dies wurde videografiert. Die Videoaufnahmen wurden dazu genutzt, das Kodiermanual zu testen und weiterzuentwickeln. Einerseits sollte ausprobiert werden, ob alle wesentlichen autonomieunterstützenden und kontrollierenden Verhaltensweisen durch die bestehenden Kategorien des Kodiermanuals abgedeckt wurden. Weiterhin sollte überprüft werden, ob diese Verhaltensweisen klar den einzelnen Beobachtungskategorien zugeordnet werden konnten. Dazu wurden die Videos von zwei weiteren Kodierern kodiert. Weiterhin wurden die Videos mit einer Woche Abstand erneut kodiert, um ein Maß für die Stabilität der Urteile zu erhalten. Als Resultat wurde das Manual überarbeitet. Beispielsweise wurde die Unterscheidung der inhaltlichen Hinweise aufgehoben und zu *Geben eines Lösungs- oder Handlungshinweises* (Kategorie *W*) zusammengefasst. Im nächsten Schritt wurde die Urteilsübereinstimmung mehrerer Kodierer mit der so entstandenen finalen Version des Manuals untersucht.

Urteilsübereinstimmung

Von den 135 Videos der Hauptstudie mussten vier aufgrund von audiovisuellen Problemen aussortiert werden. Die 131 Videos wurden von insgesamt drei Kodierern ausgewertet. Einer der Kodierer ist der Autor der vorliegenden Arbeit. Die weiteren Kodierer wurden für diesen Zweck als studentische Hilfskräfte angestellt. Sie wurden vom Autor der Arbeit an den sechs oben beschriebenen Pilotierungsvideos trainiert. Nach diesem Training haben alle Kodierer die gleichen dreizehn zufällig ausgewählten Videos kodiert. Dies entspricht ungefähr 10% des gesamten Videomaterials. Zur Überprüfung der Urteilsübereinstimmung der Kodierer wurden paarweise Cohens Kappas (J. Cohen, 1960) sowie die prozentualen Übereinstimmungen berechnet. Weiterhin wurden für die drei Kodierer *Fleiss Kappa* (Fleiss, 1971) sowie der prozentuale Anteil der Intervalle berechnet, bei denen sich alle drei Kodierer einig waren. Für alle Beobachtungs- und Auswertungskategorien liegen die prozentualen Übereinstimmungen (>84%), alle Cohen Kappas (>.71) sowie alle Fleiß Kappas (>.75) im guten bis sehr guten Bereich (Landis & Koch, 1977). Eine sehr ausführliche Darstellung der einzelnen Werte sowie einer Erklärung der Berechnung anhand mehrerer Beispiele kann dem Anhang (siehe Kapitel 12.3.1) entnommen werden.

Diskussion der Gütekriterien

Die Objektivität der Erfassung der Motivierstile von Cross-Age-Tutoren durch Verhaltensanalyse kann aufgrund des oben beschriebenen Kodiermanuals sowie der standardisierten Testprozedur angenommen werden. Wie in Kapitel 8.1 beschrieben wurde, wurden die Instruktionen für die Tutoren und Tutees wortgenau vorgegeben. Aufgrund des ausführlichen Kodiermanuals und Trainings der Kodierer wird somit angenommen, dass das Verhalten der Tutoren näherungsweise unabhängig von den Kodierern erfasst wurde.

Die Reliabilität der Erfassung der Motivierstile von Cross-Age-Tutoren durch Verhaltensanalyse kann entsprechend Cohen Kappas sowie Fleiß Kappas bewertet werden. Wie oben beschrieben wurde, sind die angegebenen Kappas Abschätzungen nach unten, da inhaltlich korrekte Kodierungen teilweise bei der Berechnung als fehlerhafte Kodierungen gewertet wurden (siehe oben). Somit können die guten bis sehr guten Cohen Kappas sowie Fleiß Kappas als Indiz für die Reliabilität gewertet werden. Da Cohens Kappa und Fleiß Kappa für zufällige Übereinstimmungen korrigieren, sind sie strengere Maße der Urteilsübereinstimmung als die prozentualen Übereinstimmungen. Folglich sind die prozentualen Übereinstimmungen ebenfalls Indizien der Reliabilität, allerdings weniger stark. Die internen Konsistenzen der daraus resultierenden Skalen werden im Anschluss an eine Hauptkomponentenanalyse in Kapitel 9.2 diskutiert.

Die Validität der Erfassung begründet sich einerseits durch die überwiegend deduktive Kategorienbildung, welche auf etablierten und mehrfach reproduzierten Studien basieren (Deci et al., 1982; Reeve et al., 1999; Reeve & Jang, 2006; Sarrazin et al., 2006). Andererseits haben Praetorius et al. (2014) in einer Videostudie die Stabilität der Instruktionsqualität von Lehrkräften untersucht, darunter auch das motivationsförderliche Lernklima. Dieses bleibt laut Praetorius et al. (2014) über mehrere Videos näherungsweise konstant, weshalb ein Video zur Erfassung dieses Klimas als ausreichend angesehen werden kann. Somit kann vermutet werden, dass dies für die Motivierstile von Tutoren ebenfalls zutrifft. Insbesondere lieferte eine Bachelorarbeit zur Erprobung des Kodiermanuals, dass die Erhebung eines Videos pro Tutor in Anlehnung an Praetorius et al. (2014) hinsichtlich der Reliabilität und inhaltlichen Validität hinreichend ist, wenn die Aspekte des Drucks von Unten erfasst werden (Wittich, 2017). Weitere Aspekte der Validität werden im Rahmen der ersten Forschungsfrage diskutiert (siehe Kapitel 10.1 und Kapitel 11.1).

8.4.3 Das Instrument zur Erfassung des wahrgenommenen autonomieunterstützenden Motivierstils von Cross-Age-Tutoren durch Selbstberichte der Tutees

Zur Erfassung des autonomieunterstützenden Motivierstils von Lehrerinnen und Lehrern, gemessen durch den wahrgenommenen autonomieunterstützenden Motivierstil von Schülerinnen und Schülern, hat sich der sogenannte *Learning Climate Questionnaire* etabliert (Black & Deci, 2000). Der *Learning Climate Questionnaire* wurde von Black und Deci (2000) in Anlehnung an den *Health-Care Climate Questionnaire* von Williams und Deci (1996) entwickelt. Black und Deci (2000) nutzten den Fragebogen, um den autonomieunterstützenden Motivierstil der Seminarleitungen an einer Universität durch die Studierenden zu erheben. Die befragten Studierenden gaben in fünfzehn Items auf einer 7-stufigen Likert-Skala an, inwiefern ihre Seminarleitung ihr Autonomieerleben unterstützte (Black & Deci, 2000). Statt der fünfzehn Items nutzen aktuelle Studien (z. B. Reeve und Cheon, 2016) eine Kurzversion des Fragebogens mit sechs Items. Sowohl die ursprüngliche Version mit fünfzehn Items als auch die Kurzversion zeigen gute interne Konsistenzen und Indizien von Validität (Black & Deci, 2000; Reeve & Cheon, 2016).

Der *Learning Climate Questionnaire* kann eingesetzt werden, um den autonomieunterstützenden Motivierstil von Cross-Age-Tutoren durch ihre Tutees zu erheben. Da der Fragebogen auf den Kontext Schule bzw. Universität angepasst ist und regelmäßigen Unterricht voraussetzt (Jang et al., 2016b), mussten die Items sinngemäß umformuliert werden. Konkret wurde der Fragebogen an das vorliegende Tutor-Setting und den Lerngegenstand angepasst, also an die Betreuung von naturwissenschaftlichen Experimenten. Zu jedem der sechs englischen Originalitems der Kurzversion des *Learning Climate Questionnaire* wurden zwischen einem und drei Items konstruiert. Dabei wurde darauf geachtet, Aspekte von Autonomieunterstützung sinngemäß in den neuen Kontext zu übertragen (siehe Kapitel 4). In manchen Fällen war der Transfer eindeutig, weshalb nur ein Item konstruiert wurde. In anderen Fällen mussten mehrere Items konstruiert werden, um die möglichen Interpretationen zu übersetzen.

Pilotierung

Die insgesamt zehn Items wurden im Rahmen von drei Vorstudien erprobt. Aufbauend auf einer Präpilotierung zur Klärung von Unklarheiten wurden in der zweiten Vorstudie die zehn Items einer Klasse von insgesamt 28 Sechstklässlerinnen und Sechstklässlern vorgelegt. Aus unterrichtsorganisatorischen Gründen konnten sie nicht das in Kapitel 8.2 beschriebene Experiment durchführen, sondern bauten verschiedene Stromkreise mit Glühlampen, Ventilatoren, Batterien und Kabeln. Um ein erstes Indiz für Validität zu untersuchen, wurde die Klasse in zwei Hälften geteilt, von denen $N_1=11$ beziehungsweise $N_2=13$ Schülerinnen und Schüler die Fragebögen vollständig ausgefüllt haben. Eine Gruppe wurde bei der Durchführung des Experimentes stark kontrollierend betreut. Die andere Gruppe wurde stark autonomieunterstützend betreut. In der Auswertung konnten sechs Items mit einer guten internen Konsistenz von $\alpha_{WMot_A}=.882$ ausgewählt werden. Diese sechs Items wurden so ausgewählt, dass die interne Konsistenz möglichst hoch ist und jedes der Originalitems (siehe Tabelle oben) genau einmal repräsentiert wurde. Die Gruppen unterschieden sich entsprechend einem Levene-Test nicht in ihrer Varianz (Field, 2009). Sowohl der gerechnete T-Test ($d = 0.63, p < .001$) als auch der Mann-Whitney-U-Test ($r = .62, p < .001$) attestierten einen höheren wahrgenommenen autonomieunterstützenden Motivierstil der Gruppe, die autonomieunterstützender betreut wurde. Aufgrund der guten Werte wurden die Items nicht weiter überarbeitet.

Die sechs ausgewählten Items wurden erneut an 51 Sechstklässlerinnen und Sechstklässlern getestet. Aus schulorganisatorischen Gründen konnten die Schülerinnen und Schüler vor dem Ausfüllen kein Experiment durchführen, weshalb sie den Fragebogen allgemein bezogen auf ihre Lehrkraft ausgefüllt haben. Die interne Konsistenz ist mit $\alpha_{WMot_A} = .724$ deutlich abgefallen. Im Rahmen dieser letzten Vorstudie wurde noch die konvergente Validität zum Kompetenzerleben und der intrinsischen Motivation der Schülerinnen und Schüler untersucht, welche in Kapitel 8.4.6 dargelegt werden.

Nr	Items der Hauptstudie
In Anlehnung an Black und Deci (2000) und Reeve und Cheon (2016)	
1	Mein Betreuer schlägt mir beim Experiment Handlungsmöglichkeiten vor und lässt mich auswählen.
2	Ich fühle mich von meinem Betreuer verstanden.
3	Mein Betreuer glaubt, dass ich das Experiment alleine durchführen kann.
4	Mein Betreuer ermutigt mich, ihn um Rat zu fragen.
5	Mein Betreuer möchte wissen, wie ich experimentieren möchte.
	Mein Betreuer interessiert sich dafür, wie ich beim Experimentieren gerne vorgehen würde.
6	Mein Betreuer macht erst neue Vorschläge zum Experimentieren, wenn er weiß, wie ich die Dinge sehe.

Tabelle 8.6: Angepasste Items des Learning Climate Questionnaires

Die gesunkene interne Konsistenz wurde mit dem fehlenden Experiment erklärt. Da die interne Konsistenz zwar niedrig, aber immer noch akzeptabel war, wurden die Items nicht überarbeitet. Es wurde lediglich eine alternative Formulierung des fünften Items hinzugenommen. Da das fünfte Item („*Mein Betreuer möchte wissen, . . .*“) möglicherweise als geringfügig kontrollierend missverstanden werden könnte, wurde die erste Satzhälfte im hinzugenommenen Item neutraler formuliert („*Mein Betreuer interessiert sich dafür, . . .*“, siehe Tabelle 8.6). Tabelle 8.6 zeigt die angepassten sieben Items der Kurzversion des Learning Climate Questionnaires (Black & Deci, 2000; Reeve & Cheon, 2016), welche in der vorliegenden Hauptstudie eingesetzt wurden.

Diskussion der Gütekriterien

Die Objektivität der Erfassung des von den Tutees wahrgenommenen autonomieunterstützenden Motivierstils kann aufgrund des standardisierten Testverfahrens in Form eines Fragebogens sowie der standardisierten Testprozedur angenommen werden. Wie in Kapitel 8.1 beschrieben wurde, wurden die Instruktionen für die Tutoren und Tutees wortgenau vorgegeben.

Die internen Konsistenzen der ausgewählten Items während der Pilotierung waren akzeptabel bis gut. Die internen Konsistenzen während der Hauptstudie werden im Anschluss an eine Hauptkomponentenanalyse in Kapitel 9.2 diskutiert.

Die Validität der Interpretation der Skalenwerte wurde einerseits durch die nahe Orientierung an dem englischen Originalinstrument begründet, welches in vielen Studien reproduzierte und theoriekonforme Ergebnisse lieferte (Black & Deci, 2000; Jang et al., 2016;

Kaur et al., 2015). Andererseits liefert der Gruppenvergleich der zweiten Vorstudie das Indiz, dass der Fragebogen die unterschiedliche Autonomieunterstützung mit einem großen Effekt messbar machen konnte. Weiterhin werden Indizien konvergenter Validität im Kapitel 8.4.6 dargelegt.

8.4.4 Intrinsische Motivation und Kompetenzerleben der Tutees

Der *Intrinsic Motivation Inventory* wurde aufbauend auf die Selbstbestimmungstheorie konstruiert (Deci et al., 1994; Ryan, 1982; Ryan & Deci, 2003). Dem Fragebogen liegt die Annahme zugrunde, dass sich die intrinsische Motivation bezüglich einer Tätigkeit von Individuen aus verschiedenen Quellen zusammensetzt (siehe Kapitel 3; Ryan und Deci, 2003). Diese sind das empfundene Vergnügen beim Ausüben der Tätigkeit, das Autonomieerleben, das Kompetenzerleben, der empfundene Druck, die Bereitschaft, die Tätigkeit trotz großem Aufwand auszuüben und dem wahrgenommenen Nutzen bzw. Wert der Tätigkeit (Deci et al., 1994; Ryan, 1982; Ryan & Deci, 2003). Die soziale Eingebundenheit während der Tätigkeit scheint für die intrinsische Motivation eine untergeordnete Rolle zu spielen (Jang et al., 2009, siehe Kapitel 3). Der *Intrinsic Motivation Inventory* besteht aus den zuvor genannten sieben Subskalen.

Die Subskala Vergnügen misst das Vergnügen an der Tätigkeit oder dem Gegenstand und ist somit prototypisch für die intrinsische Motivation bezüglich der Tätigkeit (Deci und Ryan, 1993, siehe Kapitel 3). Somit misst nur diese Skala die intrinsische Motivation bezüglich der im Fragebogen benannten Tätigkeit direkt, wohingegen die übrigen Skalen weitere Indizien bzw. Voraussetzungen für eine hohe oder niedrige intrinsische Motivation darstellen (Deci et al., 1994; Ryan, 1982; Ryan & Deci, 2003). Da entsprechend Hypothese H4a insbesondere die Zusammenhänge zwischen dem wahrgenommenen autonomieunterstützenden Motivierstil, dem Kompetenzerleben und der intrinsischen Motivation untersucht wurden, bietet sich wie bei Deci et al. (1994) die getrennte Auswertung der Subskalen an. Folglich wird die intrinsische Motivation durch die Subskala *Vergnügen*, das Kompetenzerleben durch die gleichnamige Skala und der wahrgenommene autonomieunterstützende Motivierstil durch ein separates Instrument erhoben (siehe Kapitel 8.4.3).

Im englischen Original des Fragebogens wird die Subskala *intrinsische Motivation* durch sieben Items und die Subskala *Kompetenzerleben* durch sechs Items erfasst. Ein Item der Subskala Kompetenzerleben fragt nach einer Einschätzung der erbrachten Leistung im Vergleich zu den Mitschülerinnen und Mitschülern. Dieses Item ergibt im Kontext der vorliegenden Studie keinen Sinn, da die Tutees getrennt von den anderen Tutees in einem 1:1 Verhältnis betreut wurden und somit keine Informationen über die Leistungen der anderen Tutees haben. Aus diesem Grund wurde das Item nicht eingesetzt.

Die Items der beiden Subskalen sind ähnlich aufgebaut. Sie beginnen mit einer Aussage zur experimentellen Tätigkeit des Tutees. Es folgt eine 7-stufige Likert-Skala, auf der die Tutees angeben können, in welchem Maße die Aussage für sie zutrifft. Der niedrigste Wert entspricht *stimmt gar nicht* und der höchste entspricht *stimmt völlig*. Das ungewichtete arithmetische Mittel der sieben bzw. fünf Items ergibt den Skalenwert.

In Anlehnung an den englischen Originalfragebogen und zwei deutschsprachige Übersetzungen wurden eigene Versionen der zwölf Items konstruiert (Korner, 2015; Ryan & Deci, 2003; Wilde et al., 2009). Die existierenden deutschen Übersetzungen konnten nicht genutzt werden, da in einem Fall nur eine Teilmenge der Items übersetzt wurde und diese

nicht zum Tutoringsetting passen (Wilde et al., 2009). Im anderen Fall hat die Autorin von Problemen ihrer Items berichtet (Korner, 2015).

Konstruktion und Pilotierung

Der selbst konstruierte Fragebogen wurde drei Mal pilotiert. Auf eine Präpilotierung zur Klärung von Verständnisschwierigkeiten aufbauend, wurden in einem zweiten Schritt die zwölf Items an 28 Sechstklässlerinnen und Sechstklässlern erprobt, von denen 26 die Subskala intrinsische Motivation und 22 die Subskala Kompetenzerleben vollständig ausgefüllt haben. Die internen Konsistenzen der Subskalen waren mit $\alpha_{IM} = .863$ und $\alpha_{KE} = .889$ im guten Bereich. Die Items, die nicht reverse scored formuliert waren, haben alle eine gute Itemselektivität im Intervall [.62; .85]. Die reverse scored formulierten Items hatten zwar etwas schlechtere, aber dennoch akzeptable Itemselektivitäten (.377; .453; .559). Als Resultat dieses Befundes wurden die Negierungen im Satz unterstrichen, z. B. *Dem Experiment schenkte ich überhaupt keine Aufmerksamkeit*.

Die zwölf Items wurden erneut an 51 Sechstklässlerinnen und Sechstklässlern getestet (siehe auch Kapitel 8.4.3). Aus schulorganisatorischen Gründen konnten die Schülerinnen und Schüler vor dem Ausfüllen kein Experiment durchführen, weshalb sie den Fragebogen bezogen auf den allgemeinen Unterricht ausgefüllt haben. Wie beim Fragebogen zur Erfassung des wahrgenommenen autonomieunterstützenden Motivierstils ist die interne Konsistenz der Subskala *Kompetenzerleben* mit $\alpha_{KE} = .705$ deutlich abgefallen. Folglich wird der Abfall der internen Konsistenz ebenfalls mit dem fehlenden Experiment erklärt (siehe Kapitel 8.4.3). Drei der fünf Items der Subskala weisen eine interne Konsistenz von $\alpha_{KE} = .877$ und eine gute Itemselektivität im Bereich von [.72; .81] auf. Die interne Konsistenz der Subskala *intrinsische Motivation* ist mit $\alpha_{IM} = .850$ trotz eines geringen Abfalls im guten Bereich. Bis auf ein reverse scored Item weisen die Items der Subskala *intrinsische Motivation* eine akzeptable bis gute Itemselektivität im Bereich von [.55; .79] auf. Das reverse scored Item hatte eine Itemselektivität von .166, was als mangelhaft zu bewerten ist. Bei dem problematischen Item handelte es sich um das oben benannte und überarbeitete Item *Dem Experiment habe ich überhaupt keine Aufmerksamkeit geschenkt*. Die ohnehin nicht sehr hohe Itemselektivität von .377 in der zweiten Vorstudie ergab mit dem fehlenden Experiment vermutlich die schlechte Itemselektivität in der dritten Vorstudie. Da die interne Konsistenz beider Subskalen zwar niedriger, aber trotz des fehlenden Experiments akzeptabel bis gut war, wurden die Items nicht überarbeitet und in der Hauptstudie eingesetzt. In Kapitel 8.4.6 werden die Ergebnisse der untersuchten konvergenten Validität beider Konstrukte zum wahrgenommenen autonomieunterstützenden Motivierstil dargelegt.

Diskussion der Gütekriterien

Die Objektivität der Erfassung der intrinsischen Motivation und des Kompetenzerlebens der Tutees kann aufgrund des standardisierten Testverfahrens in Form eines Fragebogens sowie der standardisierten Testprozedur angenommen werden. Wie in Kapitel 8.1 beschrieben wurde, wurden die Instruktion für die Tutees wortgenau vorgegeben.

Die internen Konsistenzen der Skalen während der Pilotierung waren im Bereich $\alpha_{KE} = [.71; .88]$ und $\alpha_{IM} = [.85; .86]$ und folglich akzeptabel bis gut. Die internen Konsistenzen während der Hauptstudie werden im Anschluss an eine Hauptkomponentenanalyse in Kapitel 9.2 diskutiert.

Die Validität der Interpretation der Skalenwerte ist einerseits durch die nahe Orientierung an dem englischen Originalinstrument begründet, welches in vielen Studien repro-

duzierte und theoriekonforme Ergebnisse lieferte (Kaur et al., 2015; Reeve & Jang, 2006; Wilde et al., 2009). Die Items der Subskala *Vergnügen* messen das Vergnügen an der Tätigkeit oder dem Gegenstand und sind somit prototypisch für die intrinsische Motivation bezüglich der Tätigkeit (siehe Kapitel 3). Bei den Items der Subskala *Kompetenzerleben* schätzen die Schülerinnen und Schüler weder ihre Leistung noch ihren Erfolg per se ein, sondern geben ihre Zufriedenheit sowie ihre wahrgenommene Wirksamkeit bezüglich kompetenzbezogener Handlungen an.

8.4.5 Übrige Instrumente

Die übrigen Instrumente wurden auf ähnliche Art und Weise adaptiert, erprobt und angepasst. Eine ausführliche Darstellung kann dem Anhang (siehe Kapitel 12.1) entnommen werden.

Autonome und kontrollierende Kausalitätsorientierung

Zur Erfassung der autonomen und kontrollierenden Kausalitätsorientierung wurde das etablierte Instrument von Deci und Ryan (1985) adaptiert. Die Adaption umfasst zwölf Situationsbeschreibungen mit jeweils einem Item pro Kausalitätsorientierung. Die Pilotierung mit insgesamt 140 Schülerinnen und Schülern ergab akzeptable interne Konsistenzen von $\alpha_{Kau_A}=.78$ und $\alpha_{Kau_K}=.74$.

Zustimmung zu den Interventionsinhalten

Reeve (1998) erfasste die Verständlichkeit, die Anwendbarkeit, die Tragfähigkeit, die Glaubwürdigkeit bzw. Authentizität und die Akzeptanz der Interventionsinhalte mit je zwei Items. Entsprechend Pintrich et al. (1993) sowie Posner et al. (1982) könnte die *Unzufriedenheit mit dem bisherigen Konzept* ebenfalls zur Zustimmung zu den Interventionsinhalten gehören, weshalb diese und die oben genannten Facetten mit jeweils zwei Items erfasst wurden. Die Pilotierung von vierundzwanzig Items an 31 Schülerinnen und Schülern ergab eine Auswahl an zwölf Items mit einer guten internen Konsistenz von $\alpha_{Zustimmung}=.846$. Aufgrund der teils schwachen Itemselektivität eines Items des Aspekts *Anwendbarkeit*, wurde dieses überarbeitet. Darüber hinaus wurde ein weiteres Item hinzugefügt. In der Haupterhebung wurden somit 13 Items eingesetzt, da der Aspekt *Anwendbarkeit* durch drei Items erhoben wurde. Mit Hilfe einer Hauptkomponentenanalyse soll anhand der Haupterhebungsdaten in Kapitel 9.2 entschieden werden, welche zwei der drei Items den Aspekt *Anwendbarkeit* repräsentieren.

Vorherige Bekanntheit der Interventionsinhalte

Die vorherige Bekanntheit der Interventionsinhalte wurde wie bei Reeve (1998) durch ein Item erfragt. Auf einer 7-stufigen Likert-Skala konnten die Teilnehmenden angeben, ob ihnen die Inhalte *1 überhaupt nicht bekannt* bis *7 sehr bekannt* waren.

Selbstbestimmtheit der Betreuungsmotivation

Thomas und Müller (2016) berechneten die Selbstbestimmtheit bezüglich des Arbeitens und Lernens in einem Schulfach aus der intrinsischen Motivation, sowie der identifizierten, introjizierten und externalen Regulation der extrinsischen Motivation. In der vorliegenden Studie wurden 17 Items zur Erfassung dieser vier Motivationsregulationen bezüglich der Tätigkeiten von Laborhelfern konstruiert und an 51 Sechstklässlerinnen und Sechstklässlern getestet. Die interne Konsistenz der vier vermuteten Skalen waren mit

$\alpha_{Intrinsisch}=.860$, $\alpha_{Identifiziert}=.650$, $\alpha_{Introjiert}=.580$, $\alpha_{External}=.685$ mangelhaft bis gut (Field, 2009). Items mit schlechten Itemselektivitäten wurden überarbeitet.

Fähigkeitsselbstkonzepte bezüglich Autonomieunterstützung und Physik

Analog zu Dickhäuser et al. (2002) beginnt jedes Item mit einer Aussage bezüglich Autonomieunterstützung oder Physik. Auf jede Aussage folgt eine 7-stufige Likert-Skala, wobei eine hohe Zahl immer ein Indiz für ein hohes Fähigkeitsselbstkonzept ist. Das ungewichtete arithmetische Mittel der Items ergibt das Fähigkeitsselbstkonzept der Tutoren bezüglich Autonomieunterstützung beziehungsweise Physik. Die Verwendung des Wortes *Autonomieunterstützung* wird insofern als unkritisch eingeschätzt, dass der Begriff bei der Intervention explizit thematisiert wurde. Fünf Items pro Fähigkeitsselbstkonzept wurden an 45 Schülerinnen und Schülern getestet. Es ergaben sich mit $\alpha_{FSK_A} = .864$ und $\alpha_{FSK_P} = .954$ gute bis exzellente interne Konsistenzen.

Wahrgenommene Selbstständigkeit, Motivation und wahrgenommenes Engagement - Der Druck von Unten (spontan)

Das wahrgenommene Engagement wurde ähnlich wie bei Sarrazin et al. (2006) durch folgendes Item auf einer 7-stufigen Likert-Skala erfasst: *Hat sich der Schüler deiner Einschätzung nach während des Experimentes angestrengt?* Da den Tutoren ein Arbeitsblatt zur Verfügung stand (siehe Kapitel 8.1), unterschied sich möglicherweise das Engagement bezüglich der Bearbeitung der Aufgaben vom Engagement beim Experimentieren. Deshalb wurde zur Erfassung des spontan wahrgenommenen Engagements folgendes Item ergänzt: *Hat sich der Schüler deiner Einschätzung nach bei der Bearbeitung der Aufgaben angestrengt?* Entsprechend ergaben sich die zwei Items zur Erfassung der wahrgenommenen Selbstständigkeit (Sarrazin et al., 2006). Die wahrgenommene intrinsische Motivation der Tutoren wurde durch drei adaptierte Items erhoben, welche ähnlich wie die Items bei Rheinberg et al. (2001) konstruiert wurden.

Verständnis der Variablen-Kontroll-Strategie

Von den 24 Items des etablierten Fragebogens zur Erfassung des Verständnisses der Variablen-Kontroll-Strategie von Schwichow et al. (2016) wurden aus testökonomischen Gründen nur elf Items eingesetzt (Konkrete Itemauswahl: ID-LS-1, ID-LS-2, UN-LS-2, ID-FL-1, UN-FL-1, ID-EIS-1, UN-EIS-1, ID-EIS-2, UN-EIS-2, ID-WS-1, ID-WS-2). Die Itemauswahl orientierte sich an der internen Konsistenz, der EAP-Reliabilität sowie der Korrelation zwischen der mittleren Personenfähigkeit geschätzt durch alle beziehungsweise die elf ausgewählten Items ($r = .89$, $\rho = .87$, $p < .001$). Letzteres kann als ein Indiz für eine valide bzw. repräsentative Itemauswahl interpretiert werden. Als weitere Nebenbedingung wurden die Items mit aufgenommen, welche thematisch zum durchzuführenden Experiment (siehe Kapitel 8.2) passen.

Physikalisches Fachwissen der Tutoren bezüglich Elektrizitätslehre

Zur Erfassung des zugehörigen Fachwissens wurden die entsprechenden Items von Christoph (2014) übernommen, der eine Variation des Experiments (siehe Kapitel 8.2) und den Fragebogen zur Erfassung des Verständnisses der Variablen-Kontroll-Strategie (siehe oben) eingesetzt hat. Am Anfang des Tests stehen vier Aussagen bezüglich der Helligkeit einer Glühlampe, welche an einer Batterie angeschlossen ist. Die Probanden können für jede Aussage ankreuzen, ob sie inhaltlich korrekt oder falsch ist. Es schließen sich die zwei offenen Aufgaben *Wovon hängt der Widerstand eines Leiters ab?* und *In einer Schaltung wird ein großer Widerstand durch einen kleineren Widerstand ersetzt. Wie verändert sich die Stromstärke, wenn die Spannung konstant bleibt?* an.

Fragebogen zur intrinsischen Motivation bezüglich der Bearbeitung der Aufgaben/ des Experiments vor der Durchführung des Experiments

Der Fragebogen zur Erfassung der aktuellen Motivation, kurz FAM, von Rheinberg et al. (2001) beinhaltet die vier Skalen *Misserfolgsbefürchtung*, *Erfolgswahrscheinlichkeit*, *Interesse* und *Herausforderung*, die sich aus insgesamt achtzehn Items zusammensetzen. Da Hypothese H4a Vermutungen über die intrinsische Motivation und das Kompetenzerleben aufstellt, wurden ähnlich zu Rheinberg et al. (2001) Items zur Erfassung der Subskalen *Interesse* und *Herausforderung* konstruiert.

Diskussion der Gütekriterien der übrigen Instrumente

Auch die Objektivität der übrigen Instrumente kann aufgrund des standardisierten Testverfahrens in Form eines Fragebogens sowie der standardisierten Testprozedur angenommen werden. Wie in Kapitel 8.1 beschrieben wurde, wurden die Instruktionen für Tutoren und die Tutees wortgenau vorgegeben.

Die internen Konsistenzen der Skalen waren während der Pilotierung in der Regel zufriedenstellend. Hatten Skalen mangelhafte interne Konsistenzen, wurden die Items mit geringen Itemselektivitäten überarbeitet. Die internen Konsistenzen während der Hauptstudie werden im Anschluss an eine Hauptkomponentenanalyse in Kapitel 9.2 diskutiert.

Die übrigen Fragebögen sind Adaptionen von etablierten Instrumenten, die sowohl national als auch international vielfach eingesetzt wurden. Diese etablierten Instrumente weisen verschiedene Validitätsindizes auf, welche den jeweils angegebenen Quellen entnommen werden können. Indizes konvergenter Validität zwischen einzelnen Fragebögen können dem folgenden Kapitel (siehe Kapitel 8.4.6) entnommen werden.

8.4.6 Konvergente Validität der Instrumente anhand einiger Pilotierungserhebungen

Hängen zwei Konstrukte theoretisch nicht zusammen, sollten die Skalenwerte der entsprechenden Fragebögen ebenfalls keine oder nur eine geringe Korrelation ($r < .3$) aufweisen. Hängen zwei Konstrukte hingegen theoretisch zusammen, sollte eine mittlere ($r > .3$) bis hohe ($r > .5$) Korrelation zu finden sein. Diese Indizes werden auch divergente Validität (oder auch diskriminante Validität) und konvergente Validität genannt (Bühner, 2006; Messick, 1995). Im Rahmen der Erprobungen der oben beschriebenen Fragebögen, wurden zwischen zehn Konstruktpaaren Korrelationen berechnet und in Tabelle 8.7 aufgelistet.

Zur Untersuchung der konvergenten Validität der Instrumente zur Erfassung der selbstberichteten Motivierstile und Kausalitätsorientierungen wurden die Skalenwerte der Instrumente mit den Skalenwerten eines angepassten *Problems in Schools* Fragebogens korreliert. Zu diesem Zweck wurden die Items des *Problems in Schools* Fragebogens übersetzt und die kulturell unverständlichen Items geringfügig angepasst (siehe Kapitel 8.4.1). Da die Fragebögen zur Erfassung der selbstberichteten Motivierstile und die Subskalen des *Problems in Schools* Fragebogens die gleichen Konstrukte erfassen sollten, wird eine hohe Korrelation ($r > .5$) erwartet. Die Motivierstile sollten im mittleren bis hohen Bereich ($r \in [.3; .6]$) mit den Kausalitätsorientierungen zusammenhängen (Kaur et al., 2015; Reeve, 1998; Reeve et al., 2018). Schwanzer et al. (2005) hat zwischen sechzehn verschiedenen Fähigkeitsselbstkonzepten Korrelationen von $r_{min} = -.29$ bis $r_{max} = .57$ gefunden. Da die Fähigkeitsselbstkonzepte bezüglich Autonomieunterstützung und Physik nicht dasselbe Fähigkeitsselbstkonzept darstellen, wird in Anlehnung an Schwanzer et al. (2005) eine

Gruppe an Konstrukten	Konstrukt 1/ Fragebogen 1	Konstrukt 2/ Fragebogen 2	N	Korrelationskoeffizient Pearson/Spearman	Erwartung
1	Selbstberichteter autonomieunterstützender Motivierstil	Selbstberichteter autonomieunterstützen der Motivierstil (Gemessen durch die Subskala HA des Problems in Schools Fragebogens)	137	.40***/.40***	$r > .5$
	Selbstberichteter kontrollierender Motivierstil	Selbstberichteter kontrollierender Motivierstil (Gemessen durch die Subskala MC des Problems in Schools Fragebogens)	138	.45***/.49***	$r > .5$
	Selbstberichteter kontrollierender Motivierstil	Selbstberichteter kontrollierender Motivierstil (Gemessen durch die Subskala HC des Problems in Schools Fragebogens)	138	.33***/.34***	$r > .5$
	Autonome Kausalitätsorientierung	Selbstberichteter autonomieunterstützender Motivierstil (Gemessen durch die Subskala HA des Problems in Schools Fragebogens)	136	.39***/.42***	$r \in [.3; .6]$
	Kontrollierende Kausalitätsorientierung	Selbstberichteter kontrollierender Motivierstil (Gemessen durch die Subskala MC des Problems in Schools Fragebogens)	135	.50***/.49***	$r \in [.3; .6]$
	Kontrollierende Kausalitätsorientierung	Selbstberichteter kontrollierender Motivierstil (Gemessen durch die Subskala HC des Problems in Schools Fragebogens)	135	.49***/.46***	$r \in [.3; .6]$
2	Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung	Fähigkeitsselbstkonzept bezüglich Physik	45	.40**/.46***	$r \in [.3; .5]$
3	Wahrgenommener autonomieunterstützender Motivierstil	Intrinsische Motivation der Tutees	50	.37**/.30*	$r \in [.2; .4]$
	Wahrgenommener autonomieunterstützender Motivierstil	Kompetenzerleben der Tutees	51	.55***/.54***	$r > .5$
	Intrinsische Motivation der Tutees	Kompetenzerleben der Tutees	51	.51***/.50***	$r > .5$

Tabelle 8.7: Korrelationen von Skalenwerten zur Untersuchung der konvergenten Validität; Signifikanzniveaus: * $p < .05$, ** $p < .01$, *** $p < .001$

Korrelation mittlerer Höhe ($r \in [.3; .5]$) erwartet. Erwartungsgemäß sollte der wahrgenommene autonomieunterstützende Motivierstil das Kompetenzerleben direkt und die intrinsische Motivation vollständig mediiert beeinflussen (siehe Kapitel 5.1). Deshalb wird zwischen dem wahrgenommenen autonomieunterstützenden Motivierstil und dem Kompetenzerleben eine hohe Korrelation erwartet ($r > .5$). Aufgrund des indirekten Effekts sollte der Zusammenhang mit der intrinsischen Motivation lediglich im niedrigen mittleren Bereich liegen ($r \in [.2; .4]$; Black und Deci, 2000; Reeve und Jang, 2006). Das Kompetenzerleben sollte die intrinsische Motivation direkt beeinflussen (siehe Kapitel 5.1), weshalb eine hohe Korrelation vermutet wird ($r > .5$).

Die Korrelationen in Tabelle 8.7 zeigen, dass die Skalen der erstellten Fragebögen in erwartungsgemäßer Richtung korrelieren. Da sowohl die Subskalen des *Problems in Schools* Fragebogens als auch die Subskalen des Instruments zur Erfassung der selbstberichteten Motivierstile die Motivierstile erfassen sollen, sind die Korrelationen unerwartet niedrig. Dies könnte auf ein Validitätsproblem der Skalen zur Erfassung der selbstberichteten Motivierstile 8.4.1 hindeuten. Alternativ könnte dies auf die bereits diskutierten Probleme des *Problems in Schools* Fragebogens für deutsche Cross-Age-Tutoren zurückgeführt werden (siehe Kapitel 8.4.1). Vor allem könnten Schülerinnen und Schüler Schwierigkeiten damit haben, Schulsituationen aus Lehrersicht zu sehen und sehr kontrollierende Handlungen nachzuvollziehen. Außerdem fällt auf, dass die konstruierte Skala zur Erfassung des kontrollierenden Motivierstils stärker mit der moderat kontrollierenden als mit der stark kontrollierenden Subskala des *Problems in Schools* Fragebogens korreliert. Folglich erfasst die Skala scheinbar vor allem einen moderat kontrollierenden Motivierstil. Da alle anderen Korrelationen in erwarteter Höhe sind, stellen sie weitere Indizien für die valide Interpretierbarkeit der Skalenwerte der vorgestellten Instrumente dar (Bühner, 2006; Messick, 1995).

8.5 Stichprobe

Die Stichprobe besteht aus insgesamt 176 Tutoren ($M_{Alter} = 13.73 \pm 0.75$ Jahren; 52.8% weiblich) und 135 Tutees ($M_{Alter} = 10.81 \pm 0.48$ Jahren; 36.3% weiblich). Diese stammen von neun städtischen Gymnasien in Nordrhein-Westfalen. Die Teilnahme der Tutoren und Tutees an der Studie sowie Teilen der Studie erfolgte freiwillig. Tabelle 8.8 zeigt, dass die Schulen mit unterschiedlich vielen Tutoren an der Intervention teilgenommen haben. Aus organisatorischen Gründen und begrenzten Forschungsmaterialien konnten maximal 16 bis 23 Schülerinnen und Schüler pro Schule videografiert werden. Aufgrund der Begrenzungen und da einige Schülerinnen und Schüler nicht gefilmt werden wollten, konnten lediglich 135 der 176 Tutoren dabei videografiert werden, wie sie je einen Tutee bei der Durchführung eines Experimentes betreuten. Die übrigen 41 Tutoren haben nur an der Intervention und am Messzeitpunkt 1 und 2 teilgenommen. Bis auf Verweigerungen erfolgte die Auswahl der Tutoren, die einen Tutee betreuen durften, randomisiert. Insgesamt liegt also eine Gelegenheitsstichprobe vor. Die neun Schulen haben an je einem der fünf Interventionstermine teilgenommen. An diesen fünf Terminen wurde immer die gleiche Intervention durchgeführt. Die ersten drei Schulen haben am ersten, die nächsten drei Schulen am zweiten und die siebte Schule am dritten Termin teilgenommen. Die achte Schule hat mit elf Schülerinnen und Schülern am vierten Termin teilgenommen und die neunte Schule mit 60 Schülerinnen und Schülern am fünften Termin. Eine Übersicht ist in Tabelle 8.8 aufgelistet.

Da die Anzahl der Teilnehmenden beim ersten, zweiten, dritten und fünften Termin über 35 lag, wurden die Gruppen für die Vorträge und die Rollenspiele mit Beobachtern (siehe Kapitel 8.3) in zwei Gruppen geteilt. Die durchführenden Experimentatoren waren der Autor dieser Arbeit sowie ein Kollege. Dieser wurde in mehreren Sitzungen vom Autor dieser Arbeit ausgebildet und hat an der Pilotierung der Intervention teilgenommen (siehe Kapitel 8.3). Um sicherzustellen, dass beide Experimentatoren die Intervention vergleichbar durchgeführt haben, waren alle Vorträge durch eine Powerpoint Präsentation vorstrukturiert. Weiterhin hatten beide eine detaillierte Liste, welche Informationen zu welcher Folie gegeben oder was die Teilnehmenden gefragt werden sollten. Diese kann den Interventionsmaterialien entnommen werden (siehe Kapitel 13. Interventionsmaterialien). Diese Listen

wurden abgehakt und anschließend verglichen. Es lagen keine Unterschiede vor, also wurden bei allen Interventionen von beiden Experimentatoren die vollständigen Informationen gegeben. Die Dauer der Vorträge war bei beiden Durchführenden annähernd gleich. 93 der 176 Probanden waren in der Gruppe des Autors der vorliegenden Studie. Um weiterhin sicherzustellen, dass diese beiden Probandengruppen zu einer Gesamtgruppe zusammengefasst werden können, wurden T-Tests und Mann-Whitney-U-Tests gerechnet (siehe Anhang Kapitel 12.3.2).

Schulnummer	Teilnahme an welcher Intervention	Anzahl der Tutoren, die an der Intervention teilgenommen haben	Anzahl der Tutoren, die an Messzeitpunkt 1 und 2 teilgenommen haben	Anzahl an Tutees, die an der videografierten Experimentiersitzung teilgenommen haben
1	1	13	13	13
2	1	11	10	10
3	1	12	12	12
4	2	19	19	19
5	2	8	8	8
6	2	18	18	16
7	3	38	32	24
8	4	11	11	11
9	5	60	53	22
Summe	—	190	176	135

Tabelle 8.8: Übersicht über die Anzahl der Tutoren und Tutees aufgeschlüsselt nach Schulnummer

Es wurde auch sichergestellt, dass die Tutees vor der Experimentiersitzung keine Elektrizitätslehre im Physikunterricht behandelt hatten und insbesondere das Experiment der Tutorsitzung (siehe Kapitel 8.2) nicht kannten. Dass die Tutees das Experiment nicht kannten, wurde zusätzlich durch ein Item abgefragt. Kein Tutee hat angegeben, die Inhalte des Experimentes vorher gekannt zu haben. Ebenso kannten die Tutoren das Konzept der Autonomieunterstützung nicht explizit. Ein implizites vorheriges Verständnis wird bei der Beantwortung der zweiten Forschungsfrage berücksichtigt (siehe Kapitel 10.2). In Kapitel 9.4 wird untersucht, inwiefern die Stichproben der neun Schulen zu einer gemeinsamen Stichprobe zusammengefasst werden können. Darauf aufbauend werden dort auch die Analysestichproben aufgeführt, welche für die Beantwortung der Forschungsfragen genutzt werden.

9 Deskriptive Ergebnisse

Anfänglich werden die Verhaltensdaten der Videoanalyse deskriptiv beschrieben. In Kapitel 9.2 werden die in Kapitel 8.4 vermuteten Skalen durch eine Hauptkomponentenanalyse abgesichert. Anschließend werden die internen Konsistenzen der eingesetzten Testinstrumente angegeben. Die parametrischen statistischen Analyseverfahren für die Beantwortung der Forschungsfragen setzen eine Normalverteilung der erhobenen Daten voraus, welche in Kapitel 9.3 überprüft wird. Der Umgang mit etwaigen Abweichungen wird ebenfalls vorgestellt. Abschließend wird in Kapitel 9.4 diskutiert, inwiefern die einzelnen Stichproben der neun Schulen (siehe Kapitel 8.5) zu einer Stichprobe zusammengefasst werden können.

Alle Berechnungen wurden entweder mit der Software *IBM SPSS Statistics 27* (IBM Corp, 2020) oder aufgrund der Limitationen von SPSS mit der Software *R* durchgeführt (R Core Team, 2020).

9.1 Deskriptive Analyse der Verhaltensdaten

Von den 135 aufgenommenen Videos mussten vier aufgrund von audiovisuellen Problemen aussortiert werden. Die übrigen 131 Videos umfassten insgesamt 15531 Intervalle, wobei die Videos im Mittel $M=118.56$, $SD=30.81$ Intervalle beziehungsweise eine Dauer von $M=19.76$, $SD=5.14$ Minuten hatten. Die Tutees haben im Mittel $M=7.22$ Mal um Hilfe gebeten oder ein Problem signalisiert. Die relativ große Standardabweichung von $SD=5.16$ weist auf die große Spannweite von minimal 0 bis maximal 26 Malen hin, in denen Tutees um Hilfe gebeten oder ein Problem signalisiert haben. Die Tutoren sind im Mittel $M=6.67$, $SD=4.13$ Mal auf die $M=7.22$, $SD=5.16$ Probleme der Tutees eingegangen. Darüber hinaus haben die Tutoren im Mittel $M=0.72$, $SD=1.05$ Mal aufgrund einer Bitte der Tutees ins Experiment eingegriffen. Neben den von den Tutees initialisierten Äußerungen der Tutoren, haben die Tutoren im Mittel $M=38.86$, $SD=16.21$ Mal von sich aus gesprochen und im Mittel $M=7.18$, $SD=7.69$ Mal ungewollt ins Experiment gegriffen. Insbesondere haben die Tutoren minimal 0 bis maximal 42 Mal ungewollt ins Experiment der Tutees eingegriffen. Ein Eingriff wird nur dann als ungewollt kodiert, wenn keinerlei Hilfe benötigt und keinerlei Problem signalisiert wurde. Deshalb sind ungewollte Eingriffe als kontrollierendes Verhalten zu werten (siehe Kapitel 4). Die $M=38.86$, $SD=16.21$ vom Tutor initialisierten Äußerungen können auch Ermutigungen oder Zuspruch enthalten. Zur besseren Differenzierung wird deshalb in den Auswertungskategorien bei den vom Tutor initialisierten inhaltlichen Eingriffen unterschieden, ob der Tutee Hilfe benötigte oder nicht (siehe Kapitel 8.4.2). Von den $M=38.86$, $SD=16.21$ von den Tutoren initialisierten Aussagen sind $M=30.53$, $SD=13.98$ verbale inhaltliche Eingriffe, in denen der Tutee keine Hilfe nötig hatte und auch kein Problem signalisiert hat. Wie Tabelle 9.1 entnommen werden kann, haben die Tutees vor allem direkt nach Hinweisen gefragt (H1 und H2), statt um einen Eingriff ins Experiment (H3) zu bitten oder Schwierigkeiten zu zeigen (H4).

Kürzel	Beobachtungskategorie	Aufsummierte Häufigkeit	Durchschnittliche Häufigkeit pro Video
H1	Tutee fragt nach einem/ bittet um einen Lösungshinweis	665	5.08(4.08)
H2	Tutee fragt nach einem/ bittet um einen Handlungshinweis	657	5.02(4.12)
H3	Tutee bittet um einen physischen Eingriff ins Experiment	12	0.09(0.29)
H4	Tutee signalisiert verbal eine inhaltliche oder sozial-emotionale Schwierigkeit	85	0.65(1.21)
H	Aggregierte Kategorie. Tutees haben mindestens in einer der Kategorien H1, H2, H3 oder H4 um Hilfe gebeten oder Schwierigkeiten signalisiert	946	7.22(5.16)
T1	Der Betreuer greift verbal ein	5965	45.53(16.75)
	• Mit Tuteeproblem	• 874	• 6.67(4.13)
	• Ohne Tuteeproblem	• 5091	• 38.86(16.21)
T2	Der Betreuer greift physisch ein	1035	7.90(8.03)
	• Mit Tuteeproblem	• 94	• 0.72(1.05)
	• Ohne Tuteeproblem	• 941	• 7.18(7.70)
Event	Intervalle, in denen der Tutor entweder physisch beziehungsweise verbal oder nonverbal eingegriffen hat (Ergibt sich aus T1 und T2)	6275	47.61(17.61)

Tabelle 9.1: Absolute und durchschnittliche Häufigkeiten der Beobachtungskategorien – Teil 1

Die Kategorien W, S und E wurden nur in den Intervallen kodiert, in denen der Tutor entweder physisch, verbal oder nonverbal eingegriffen hat, da die Kategorien W, S und E einen Eingriff des Tutors voraussetzen (siehe Kapitel 8.4.2). Auf diese Weise wurden 6275 Events identifiziert. Da diese teilweise in Folgeintervalle reichten, dauerten sie insgesamt 8160 Intervalle an. Die aufsummierten Häufigkeiten in Tabelle 9.2 beziehen sich auf die Anzahl der Events, in denen die jeweilige Kategorie mindestens einmal aufgetreten ist.

Die deutliche Mehrheit aller Tutoraussagen war nicht kontrollierend (5137 im Vergleich zu 1083 Aussagen). Weiterhin sind die Kategorien E1 und E2 mit 15 und 81 Malen sehr selten aufgetreten. Somit wurde nur in 1.6% der 946 Fälle von Tuteeschwierigkeiten mit

Kürzel Beobachtungskategorie		Aufsum. Häufigkeit	Durchschn. Häufigkeit pro Video
W	Geben eines Hinweises (Lösungs- oder Handlungshinweis)	4871	37.18(14.68)
S1	„Musst“-/ „Sollst“- Aussagen	302	2.31(2.51)
S2	Lenkende Fragen	202	1.54(1.87)
S3	Imperative	579	4.42(5.12)
S4	Nicht-kontrollierende Aussagen/Fragen	5137	39.21(14.29)
E1	Ermutung bei Schwierigkeiten	15	0.11(0.36)
E2	Verständnis bei Schwierigkeiten	81	0.62(1.09)
E3	Auf Tutee eingehen	2321	17.72(9.64)
E4	Lob für Fortschritt und Erfolg	187	1.43(1.79)

Tabelle 9.2: Absolute und durchschnittliche Häufigkeiten der Beobachtungskategorien - Teil 2

E1 und nur mit 8.6% der Fälle mit E2 reagiert.

Da nicht alle Beobachtungskategorien per se autonomieunterstützendes oder kontrollierendes Verhalten darstellen, wurden aus den Beobachtungskategorien die Auswertungskategorien berechnet (siehe Kapitel 8.4.2). Die aufsummierten Häufigkeiten beziehen sich auf die 6275 Events, in denen das beschriebene Verhalten mindestens einmal aufgetreten ist. Die durchschnittlichen Häufigkeiten pro Video sind die Durchschnitte der jeweiligen Kategorie bezogen auf die 131 berücksichtigten Videos.

Die Tutorsitzungen dauerten durchschnittlich $M=19.76$, $SD=5.14$ Minuten, wobei sie nach spätestens 25 Minuten durch den Experimentator beendet wurden. Um die Häufigkeiten der Auswertungskategorien trotz der unterschiedlichen Sitzungslängen miteinander vergleichbar zu machen, wurden alle Auswertungskategorien durch die Anzahl der Intervalle des jeweiligen Videos geteilt (Reeve & Jang, 2006). Weiterhin könnten die Auswertungskategorien von Natur aus unterschiedlich häufig vorkommen, weshalb alle Kategorien relativ zu den übrigen Tutoren normiert wurde (Sarrazin et al., 2006; siehe Kapitel 8.4.2).

Von den insgesamt 15531 Intervallen, haben die Tutoren in 7959 Intervallen (51.2%) gesprochen und in 1561 (10.0%) Intervallen physisch ins Experiment eingegriffen. In Anlehnung an Reeve und Jang (2006) sowie Sarrazin et al. (2006) sollten die auf Videolänge und relative Ausprägung normierten Kategorien $TW1_A_soz$, $TW2_A_soz$, $E1_soz$, $E2_soz$, $E3_soz$ und $E4_soz$ den beobachteten autonomieunterstützenden Motivierstil und die Kategorien $TW1_K_soz$, $TW2_K_soz$, $S1_soz$, $S2_soz$, $S3_soz$ den beobachteten kontrollierenden Motivierstil ergeben. Die Skalenbildung baut auf der Annahme auf, dass das Autonomieerleben der Tutees stärker unterstützt bzw. frustriert wird, je häufiger die jeweiligen Verhaltensweisen auftreten. Alternativ könnte man die auf die Anzahl der Events normierte Skala betrachten. Beispielsweise könnte ein Tutor zwei Aussagen mit Imperativen gefällt haben. Entsprechend der durchschnittlichen Häufigkeit pro Video in Tabelle 9.2, wäre diese Anzahl durchschnittlich. Hat der Tutor allerdings insgesamt nur zwei Aussagen gefällt, war jede seiner Aussagen eine Aufforderung. Der auf die Anzahl der Event normierte Skalenwert würde 1 betragen. Hat der Tutor hingegen zwanzig Aussagen gefällt, wären nur 10% der Aussagen Aufforderungen. Der auf die Anzahl der Events normierte Skalenwert würde .1 betragen, wenn der Tutor nie physisch ins Experiment eingegriffen hat. Bezüglich der auf die auf Videolänge und relative Ausprägung normierten Skala hätte der Tutor in beiden Fällen mit zwei auffordernden Aussagen den gleichen Wert. Die auf Videolänge und relative

Ausprägung normierte Skala vergleicht die Tutorsitzungen hinsichtlich der Auftretenshäufigkeit vermuteter autonomieunterstützender und kontrollierender Verhaltensweisen. Die auf die Anzahl der Events normierte Skala (im Folgenden mit dem Suffix „_ev“ gekennzeichnet) korrigiert für den Einfluss der Häufigkeit und stellt den relativen Anteil einer Kategorie an der Gesamtheit aller Verhaltensweisen dar.

Kürzel	Auswertungskategorie	Aufsum. Häufigkeit	Durchschn. Häufigkeit pro Video	M(SD)_ev	M(SD)_soz
Autonomieunterstützung					
TW1 _A	Verbale inhaltliche Eingriffe, wenn der Tutee Hilfe nötig hatte	867	6.62(4.14)	.15(.10)	.32(.20)
TW2 _A	Physische inhaltliche Eingriffe, wenn der Tutee Hilfe nötig hatte	83	.63(.95)	.01(.02)	.11(.17)
E1	Ermutigung bei Schwierigkeiten	15	0.11(0.36)	<.01(<.01)	.07(.21)
E2	Verständnis bei Schwierigkeiten	81	0.62(1.09)	.01(.02)	.13(.21)
E3	Bestätigend auf Aussagen des Tutees reagieren	2321	17.72(9.64)	.36(.13)	.32(.18)
E4	Nutzung von Lob bei Fortschritt und Erfolg	187	1.43(1.79)	.03(.04)	.20(.25)
Kontrolle					
TW1 _K	Verbale inhaltliche Eingriffe, obwohl Tutee keine Hilfe nötig hatte	4000	30.53(13.98)	.63(.14)	.47(.20)
TW2 _K	Physische inhaltliche Eingriffe, obwohl Tutee keine Hilfe nötig hatte	610	4.66(5.42)	.09(.08)	.14(.16)
S1	Nutzung von druckerzeugenden Modalverben	302	2.31(2.51)	.05(.05)	.18(.21)
S2	Nutzung von lenkenden Fragen	202	1.54(1.87)	.03(.04)	.17(.21)
S3	Nutzung von Imperativen	579	4.42(5.12)	.08(.08)	.21(.24)
Weitere Auswertungsmaße					
RZ	Anzahl der Intervalle, in denen der Tutor verbal beziehungsweise nonverbal kommuniziert hat (Redezeit)	7959	60.76(21.89)	1.28(.15)	.51(.20)
EZ	Anzahl der Intervalle, in denen der Tutor physisch ins Experiment eingegriffen hat (Eingriffszeit)	1561	11.92(13.51)	.23(.22)	.15(.17)
Eventzeit	Anzahl der Intervalle, in denen der Tutor verbal beziehungsweise nonverbal kommuniziert oder physisch ins Experiment eingegriffen hat	8160	62.29(22.70)	1.31(.14)	.52(.21)

Tabelle 9.3: Absolute und durchschnittliche Häufigkeiten der Auswertungskategorien; x_ev: Auf die Anzahl der Events normierte Kategorie X; x_so: Auf Videolänge und relative Ausprägung normierte Kategorie X.

9.2 Struktur und Interne Konsistenzen der Instrumente

In Kapitel 8.4 wurden die Instrumente vorgestellt, welche in der vorliegenden Studie eingesetzt wurden. Da einige Instrumente von Grund auf neu entwickelt wurden (siehe z. B. Kapitel 8.4.1) oder mehr Items zur späteren Auswahl eingesetzt wurden (siehe Kapitel 8.4.5), wird im Folgenden die Struktur der Instrumente untersucht. Wie bei der Vorstellung der Instrumente (siehe Kapitel 8.4) wird die Untersuchung der Struktur der Instrumente zur Erfassung der Motivierstile sowie der Zielvariablen auf Tutorebene (intrinsische Motivation nach der Durchführung des Experiments) ausführlich beschrieben. Die Untersuchung der Struktur der übrigen Instrumente wird zusammenfassend dargestellt, wobei eine detaillierte Analyse Kapitel 12.2 entnommen werden kann.

Die Stichprobengröße kann als hinreichend angesehen werden, wenn pro untersuchtem Item einer Hauptkomponentenanalyse mindestens 4 Probanden gegeben sind (Little, 2013). Da der umfangreichste Fragebogen vierundzwanzig Items enthält, sollten mindestens 96 Probanden vorhanden sein. Weiterhin empfiehlt Gorsuch (2013) ein Minimum von 100 Fällen. In jeder der nachfolgenden Analysen sind mindestens 116 Probanden berücksichtigt, weshalb die oben genannten Voraussetzungen erfüllt sind.

Um die Struktur der Fragebögen zu untersuchen, wurden Hauptkomponentenanalysen durchgeführt (Field, 2009). Nur Komponenten mit einem Eigenwert größer als 1, die mindestens 10% der Gesamtvarianz ausmachen, wurden ausgewählt (Urda, 2011). Interne Konsistenzen unter .70 können als mangelhaft bewertet werden (Blanz, 2015; Bühner, 2006; Field, 2009). Interne Konsistenzen über .70 können als ausreichend, Werte über .80 als gut und Werte über .90 als exzellent klassifiziert werden (Blanz, 2015; Bühner, 2006; Field, 2009). Itemselektivitäten über .30 werden als ausreichend und Werte über .60 als gut interpretiert (Clark & Watson, 1995; Field, 2009; Stevens, 2001). Kaiser-Meyer-Olkin-Maße der Stichprobenadäquanz (KMO) unter .50 deuten auf eine mangelhafte Eignung der Daten für eine Hauptkomponentenanalyse hin (Field, 2009; Hutcheson, 1999; Kaiser, 1974). KMOs über .50 können als ausreichend und Werte über .80 als gut klassifiziert werden (Field, 2009; Hutcheson, 1999; Kaiser, 1974).

9.2.1 Motivierstile der Tutoren erfasst durch Selbstberichte

Zur Untersuchung der Struktur wurden sowohl die präinterventionellen als auch die postinterventionellen Items berücksichtigt. Die Kaiser-Meyer-Olkin-Maße der Stichprobenadäquanz (KMO) aller Analysen können als gut charakterisiert werden. Der Bartlett-Test zur Sphärizität war mit $p < .001$ signifikant. Da der Fragebogen erwartungsgemäß zwei Komponenten beinhalten sollte (Reeve et al., 2014), die plausibel miteinander korrelieren könnten (Reeve et al., 2014), wurde die Oblimin-Rotation genutzt (Field, 2009). Das Item P1_TE_K funktioniert präinterventionell nicht hinreichend. Da es postinterventionell gut funktioniert, wurde es nicht verworfen. Die Items gruppieren sich sowohl präinterventionell als auch postinterventionell erwartungsgemäß in die Hauptkomponenten des autonomieunterstützenden (..._A) und kontrollierenden (..._K) Motivierstils. Mit Ausnahme des Items P1_TE_K erweisen sich die Itemselektivitäten präinterventionell mit .30-.62 als ausreichend bis gut und postinterventionell mit .54-.80 als ausreichend bis gut. Die internen Konsistenzen sind mit $\alpha_{SMot_A_Prä}=.781$, $\alpha_{SMot_K_Prä}=.772$, $\alpha_{SMot_A_Post}=.907$ und $\alpha_{SMot_K_Post}=.920$ als ausreichend bis exzellent zu bewerten. Die Hauptkomponenten korrelieren präinterventionell nicht signifikant und postinterventionell in einer Höhe von $r=-.329$ ($p < .001$) miteinander.

Itemname	präinterventionell		postinterventionell	
	KMO = .706		KMO = .884	
	Komp. 1	Komp. 2	Komp. 1	Komp. 2
P1_TE_A	.452		.602	-.333
P2_SE_A	.596		.744	
P3_TW_A	.509		.757	
P4_TS_A	.561		.759	-.407
P5_WE_A	.397		.601	
P6_TW_A	.503		.739	
P7_TS_A	.471		.705	
P8_WS_A	.577		.753	
P9_WS_A	.630		.732	
P10_WE_A	.658		.804	
P11_TE_A	.552		.740	
P12_SE_A	.542		.617	
<hr/>				
P1_TE_K				.620
P2_SE_K		.581		.724
P3_TW_K		.404		.758
P4_TS_K		.546		.817
P5_WE_K		.423		.656
P6_TW_K		.592		.771
P7_TS_K		.496		.774
P8_WS_K		.613		.830
P9_WS_K		.754		.842
P10_WE_K		.687	-.318	.844
P11_TE_K		.501		.554
P12_SE_K		.498		.557

Tabelle 9.4: Hauptkomponentenanalyse der Items zur Erfassung der selbstberichteten Motivierstile. Ladungen $< .3$ wurden unterdrückt.

9.2.2 Motivierstile der Tutoren erfasst durch Verhaltensbeobachtung

Da das autonomieunterstützende und kontrollierende Verhalten plausibel miteinander korrelieren könnten, wurde die Oblimin-Rotation genutzt (Field, 2009).

Itemname	Verhaltensweisen gemeinsam			Nur Autonomie- unterstützende		Nur Kontrol- lierende	Nur Kontrol- lierende
	KMO = .616			KMO = .565		KMO = .655	KMO = .662
	Komp. 1	Komp. 2	Komp. 3	Komp. 1	Komp. 2	Komp. 1	Komp. 1
Autonomieunterstützung							
TW1 _A _soz			.840		.803	—	—
TW2 _A _soz			.641		.789	—	—
E1_soiz		.753		.717		—	—
E2_soiz	.305	.591		.648		—	—
E3_soiz	.653			.319		—	—
E4_soiz		.702		.725		—	—
<hr/>							
Kontrolle							
TW1 _K _soz	.892			—	—	.884	.872
TW2 _K _soz	.690		.308	—	—	.721	.747
S1_soiz	.379		.585	—	—	.585	.613
S2_soiz	.475	-.314		—	—	.370	—
S3_soiz	.608			—	—	.695	.698

Tabelle 9.5: Hauptkomponentenanalyse der auf die auf Videolänge und relative Ausprägung normierten Verhaltensweisen (..._soz) zur Erfassung der beobachteten Motivierstile. Ladungen $< .3$ wurden unterdrückt.

Betrachtet man die Verhaltensweisen gemeinsam, so ergeben sich nicht die erwarteten Skalen autonomieunterstützenden und kontrollierenden Verhaltens. Das Kaiser-Meyer-Olkin-Maß der Stichprobenadäquanz (KMO) kann als ausreichend charakterisiert werden. Der Bartlett-Test zur Sphärizität war mit $p < .001$ signifikant. Die erste Komponente stellt die kontrollierende Skala dar, wobei E3 ebenfalls gut zu dieser Komponente passt, S1 hingegen passt eher zur dritten Komponente. Die zweite Komponente besteht im Wesentlichen aus drei der vier Verhaltensweisen des Umgangs mit Tuteeäußerungen negativer Gefühle. Die dritte Komponente besteht einerseits aus den inhaltlichen Eingriffen der Tutoren, die von den Tutees erwünscht bzw. nötig waren. Allerdings passt das Nutzen von „musst“-/„sollst“-Aussagen ebenfalls dazu. Die drei Komponenten haben interne Konsistenzen von $\alpha_{Komp1} = .699$, $\alpha_{Komp2} = .515$, $\alpha_{Komp3} = .507$. Folglich kommt lediglich eine Betrachtung der ersten Komponente in Frage.

Betrachtet man nur die vermutet autonomieunterstützenden Verhaltensweisen ergeben sich zwei Komponenten. Ähnlich der gemeinsamen Analyse sind sowohl die inhaltlichen Eingriffe der Tutoren als auch der Umgang mit Tuteeäußerungen negativer Gefühle getrennte Komponenten. E3 lädt zwar auf die letztgenannte Hauptkomponente, allerdings nur schwach mit einem Wert von .319. Die interne Konsistenz der ersten Komponente entspricht somit erneut $\alpha_{Komp2} = .515$, weshalb eine weitere Betrachtung nicht in Frage kommt. Eine Korrelation der beiden Kategorien autonomieunterstützender inhaltlicher Eingriffe der Tutoren nach Spearman beträgt lediglich $\rho = .193$, $p = .027$, weshalb die Items ebenfalls nicht zu einer Skala zusammengefasst werden können. Zusammengefasst konnte aus den sechs vermutet autonomieunterstützenden Verhaltensweisen keine Hauptkomponente zur weiteren Analyse extrahiert werden.

Betrachtet man nur die vermutet kontrollierenden Verhaltensweisen, kann das Kaiser-Meyer-Olkin-Maß der Stichprobenadäquanz (KMO) als ausreichend charakterisiert werden. Der Bartlett-Test zur Sphärizität war mit $p < .001$ signifikant. Nur eine Komponente hatte einen Eigenwert über 1 mit einer aufgeklärten Varianz von 45,2%. Der Screeplot legt ebenfalls nur eine einzelne Komponente nahe (Field, 2009). Die Kategorie kontrollierender Fragen S2_soz passt mit einer Ladung von .370 nicht gut zur Komponente (Field, 2009). S2 hat eine Item-Skalen-Korrelation von .208. Deshalb zeigt der Vergleich der internen Konsistenzen $\alpha_{MitS2} = .665$ bzw. $\alpha_{ohneS2} = .705$ auch eine bessere interne Konsistenz ohne S2_soz (Field, 2009). Insgesamt ergeben die auf Videolänge und relative Ausprägung normierten Kategorien TW1K_soz, TW2K_soz, S1_soz und S3_soz mit einer akzeptablen internen Konsistenz von $\alpha = .705$ den beobachteten kontrollierenden Motivierstil.

Da die gemeinsame Faktorenanalyse eine Passung der Kategorie E3_soz zum beobachteten kontrollierenden Motivierstil nahelegt, ergibt sich die Frage, ob E3_soz als weiterer Indikator hinzugenommen wird. Aufgrund des geringfügig besseren KMO, der geringfügig besseren internen Konsistenz sowie der besseren theoretischen Passung werden im Folgenden TW1K_soz, TW2K_soz, S1_soz und S3_soz ohne E3_soz als beobachteter kontrollierender Motivierstil betrachtet. Die Rolle von E3 sowie der Zusammenhang zum beobachteten kontrollierenden Motivierstil werden in Kapitel 10.3 sowie 11.1 diskutiert. Um ebenfalls die Struktur der auf die Anzahl der Events normierten Verhaltensweisen zu untersuchen, wurde eine weitere Analyse durchgeführt.

Betrachtet man die gemeinsamen oder nur die autonomieunterstützenden Verhaltensweisen, ergibt sich kein ausreichendes Kaiser-Meyer-Olkin-Maß der Stichprobenadäquanz (siehe 9.6). Das Kaiser-Meyer-Olkin-Maß der Stichprobenadäquanz ist bei der Betrachtung der vermutet kontrollierenden Verhaltensweisen mit .595 zwar ausreichend, doch ist die interne Konsistenz der Skala aus den Items mit $\alpha_{BMot_K_ev} = .475$ ungenügend.

Itemname	Nur Kontrollierende KMO = .595	
	Komp. 1	Komp. 2
TW1 _{K_ev}	.667	.420
TW2 _{K_ev}	.674	
S1_ev	.468	
S2_ev		.938
S3_ev	.671	

Tabelle 9.6: Hauptkomponentenanalyse der auf die Anzahl der Events normierten Verhaltensweisen zur Erfassung der beobachteten Motivierstile. Ladungen $< .3$ wurden unterdrückt.

Zusammengenommen konnte so aus den auf die Anzahl der Events normierten Verhaltensweisen keine Komponente zur weiteren Analyse extrahiert werden. Lediglich die auf Videolänge und relative Ausprägung normierten kontrollierenden Verhaltensweisen ergeben eine Skala.

9.2.3 Motivierstile der Tutoren erfasst durch Berichte der Tutees/ Wahrgenommener autonomieunterstützender Motivierstil

Zu jedem der sechs Originalitems der Kurzversion des Learning Climate Questionnaires wurde ein Item in der Hauptstudie eingesetzt. Das fünfte Item wurde durch zwei Items repräsentiert (siehe Kapitel 8.4.3). Das Kaiser-Meyer-Olkin-Maß der Stichprobenadäquanz (KMO) kann als gut charakterisiert werden. Der Bartlett-Test zur Sphärizität war mit $p < .001$ signifikant.

Itemname	Analyse 1 (KMO = .750)		Analyse 2 (KMO = .770)
	Komponente 1	Komponente 2	Komponente 1
LCQ_1	.672		.641
LCQ_2	.733		.688
LCQ_3		.920	—
LCQ_4	.515		.523
LCQ_5a	.589	.315	.669
LCQ_5b	.744		.750
LCQ_6	.451	.614	.649

Tabelle 9.7: Hauptkomponentenanalyse der Items zur Erfassung des von den Tutees wahrgenommenen autonomieunterstützenden Motivierstils. Ladungen $< .3$ wurden unterdrückt.

Die erste Analyse zeigt, dass das dritte Item nicht zur restlichen Skala passt. Die auf den Kontext angepasste Übersetzung lautet *Mein Betreuer glaubt, dass ich das Experiment alleine durchführen kann*. Möglicherweise wurde das Item von einigen Probanden missverstanden. Im englischen Original steht das Vertrauen des Tutors in die Fähigkeiten des Tutees im Vordergrund. Dieser Fokus ist durch die Übersetzung möglicherweise untergegangen. Folglich wurde das Item entfernt und die Analyse wiederholt. Da beide Versionen des fünften Items gut funktionieren und zuvor ein Item gestrichen werden musste, wurden sowohl LCQ_5a als auch LCQ_5b behalten. Die Itemselektivitäten erwiesen sich mit

.35-.57 als ausreichend. Die interne Konsistenz ist mit $\alpha_{WMot_A}=.723$ als ausreichend zu bewerten.

9.2.4 Intrinsische Motivation und Kompetenzerleben der Tutees

Das Kaiser-Meyer-Olkin-Maß der Stichprobenadäquanz (KMO) kann als gut charakterisiert werden. Der Bartlett-Test zur Sphärizität war mit $p < .001$ signifikant. Da der Fragebogen erwartungsgemäß zwei korrelierende Komponenten beinhalten sollte, wurde die Oblimin-Rotation genutzt (Field, 2009).

Itemname	Analyse 1 (KMO = .891)		Analyse 2 (KMO = .893)	
	Komp. 1	Komp. 2	Komp. 1	Komp. 2
IMI_IE1	.793	.678	.803	.686
IMI_IE2_R	.849	.308	.842	.353
IMI_IE3	.906	.484	.905	.516
IMI_IE4	.705	.688	.708	.706
IMI_IE5	.653	.613	.666	.612
IMI_IE6_R		.406	—	—
IMI_IE7	.912	.485	.913	.515
IMI_PC1	.474	.862	.471	.879
IMI_PC2_R	.383	.647	.383	.657
IMI_PC3	.359	.852	.358	.859
IMI_PC4	.532	.783	.529	.804
IMI_PC5	.451	.751	.456	.755

Tabelle 9.8: Hauptkomponentenanalyse der Items zur Erfassung der intrinsischen Motivation und des Kompetenzerlebens. Ladungen $< .3$ wurden unterdrückt.

Das sechste Item der Skala zur Erfassung der intrinsischen Motivation lädt auf keine der beiden Hauptkomponenten stark. Insbesondere lädt es kaum auf die erste Hauptkomponente, welche die intrinsische Motivation darstellt. Bei dem problematischen Item handelt es sich um das Item, welches während der Pilotierung mehrfach überarbeitet wurde und teils mangelhafte Itemselektivitäten hatte. Folglich wird das Item entfernt und die Analyse wiederholt. In der zweiten Analyse können weiterhin zwei Hauptkomponenten identifiziert werden. Für beide Komponenten gibt es allerdings starke Crossladungen. Die Itemselektivitäten der Items beider Komponenten erwiesen sich mit .52-.84 als ausreichend bis gut. Die internen Konsistenzen $\alpha_{IM} = .898$ und $\alpha_{KE} = .848$ können als gut bewertet werden. Die erwarteten Hauptkomponenten korrelieren in einer Höhe von $\rho = .598, p < .001$ miteinander. Da die Komponenten theoriekonform sehr hoch miteinander zusammenhängen sollten (z. B. Kapitel 8.4.6), werden die hohen Crossladungen akzeptiert und die starke Korrelation als Indiz für Validität interpretiert.

9.2.5 Übrige Instrumente

Die Struktur und interne Konsistenz der übrigen Instrumente wurde auf ähnliche Art und Weise untersucht. Eine ausführliche Darstellung kann dem Anhang (siehe Kapitel 12.2) entnommen werden.

Autonome und kontrollierende Kausalitätsorientierung

Vier der 24 Items weisen geringfügige Crossladungen $<.4$ auf. Ansonsten ergeben je zwölf Items sowohl präinterventionell als auch postinterventionell die erwarteten Kausalitätsorientierungen. Die Itemselektivitäten erwiesen sich mit $.34-.77$ als ausreichend bis gut. Die internen Konsistenzen sind mit $\alpha_{Kau_A_Prä}=.770$, $\alpha_{Kau_K_Prä}=.789$, $\alpha_{Kau_A_Post}=.788$ und $\alpha_{Kau_K_Post}=.856$ als ausreichend bis gut zu bewerten.

Zustimmung zu den Interventionsinhalten

Das Instrument zur Erfassung der Zustimmung zu den Interventionsinhalten umfasst sechs Facetten, welche durch insgesamt 13 Items erfasst wurden. Die Hauptkomponentenanalyse ergibt, dass die beiden Items der Facette *Unzufriedenheit* nicht zu der restlichen Skala passen, weshalb diese Facette wie bei Reeve (1998) im Folgenden nicht berücksichtigt wird. Eines der drei Items der Facette *Anwendbarkeit* sowie ein Item der Facette *Verständlichkeit* wurden aufgrund niedriger Faktorladungen entfernt. Die Ladungen und Selektivitäten der übrig gebliebenen neun Items erwiesen sich als ausreichend bis gut. Die interne Konsistenz kann mit $\alpha_{Zustimmung}=.907$ als exzellent bewertet werden.

Selbstbestimmtheit der Betreuungsmotivation

Die Analyse ergibt drei Hauptkomponenten, die als die intrinsische Motivation, die identifizierte Regulation und als eine Mischung der introjizierten mit der externalen Regulation der extrinsischen Motivation identifiziert werden können. Fünf der 17 Items haben entweder niedrige Ladungen, zu hohe Crossladungen oder weisen zu niedrige Itemselektivitäten auf. Die Itemselektivitäten der ersten beiden Komponenten (intrinsische Motivation und identifizierte Regulation) erwiesen sich mit $.50-.77$ als ausreichend bis gut. Die Itemselektivitäten der dritten Komponente sind mit $.32-.55$ ausreichend. Die internen Konsistenzen $\alpha_{SRQ_intr}=.831$, $\alpha_{SRQ_ident}=.753$ und $\alpha_{SRQ_introj+extrin}=.658$ können als mangelhaft bis gut bewertet werden. Da die dritte Komponente die theoretisch vermutete Trennung der introjizierten und externalen Regulation nicht abbildet und die interne Konsistenz trotz der Zusammenfassung nicht ausreichend ist, wird die dritte Komponente im Folgenden nicht berücksichtigt. Da erwartungswidrig keine vier Komponenten identifiziert werden konnten, konnte die Selbstbestimmtheit nicht berechnet werden. Folglich werden die intrinsische Motivation und identifizierte Regulation der extrinsischen Motivation im Folgenden getrennt voneinander in Analysen berücksichtigt.

Fähigkeitsselbstkonzepte bezüglich Autonomieunterstützung und Physik

Bis auf geringe Crossladungen können die beiden Hauptkomponenten erwartungsgemäß als das Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung beziehungsweise Physik interpretiert werden. Die Itemselektivitäten erwiesen sich mit $.61-.86$ als gut bis sehr gut. Die internen Konsistenzen $\alpha_{FSK_A}=.851$ und $\alpha_{FSK_P}=.926$ können als gut bis exzellent bewertet werden. Die erwarteten Hauptkomponenten korrelieren in einer erwarteten Höhe von $\rho = .310, p < .001$ miteinander (vgl. mit der entsprechenden Korrelation in Kapitel 8.4.6).

Wahrgenommene Selbstständigkeit, Engagement und Motivation (Druck von Unten)

Unter Berücksichtigung des Screeplots der Hauptkomponentenanalyse ergeben sich die drei erwarteten Komponenten des wahrgenommenen Engagements, der wahrgenommenen Selbstständigkeit sowie der wahrgenommenen Motivation mit Ladungen über $.9$. Die drei Items der wahrgenommenen Motivation haben Crossladungen zu den ersten beiden Komponenten. Eine konfirmatorische Faktorenanalyse bestätigt mit einem Signifikanzniveau

von $p < .001$, dass das dreifaktorielle Modell die Daten besser als jedes zweidimensionale Modell beschreibt. Die interne Konsistenz der wahrgenommenen Tuteemotivation ist mit $\alpha_{TW_IM} = .876$ gut. Die Items des vermuteten wahrgenommenen Engagements und der wahrgenommenen Selbstständigkeit korrelieren jeweils hinreichend hoch miteinander ($\rho_{TW_En} = .650, p < .001$ sowie $\rho_{TW_Se} = .747, p < .001$).

Verständnis der Variablen-Kontroll-Strategie

Die Hauptkomponentenanalyse hat eine vierdimensionale Struktur offenbart. Die erste Komponente besteht aus den vier UN-Items und hat einen Eigenwert von 2.5 mit einer aufklärten Varianz von 25,4%. Die übrigen drei Komponenten haben Eigenwerte zwischen 1.0 und 1.3 beziehungsweise aufklärte Varianzen zwischen 10.4% und 12.9%. Deshalb wurde für die weitere Analyse ausschließlich die erste Komponente berücksichtigt, wobei eines der vier Items aufgrund niedriger Itemselektivität ausgeschlossen wurde (Field, 2009). Die finale Skala ergibt sich also aus den Items VKS_UNLS2, VKS_UNFL1 und VKS_UNEIS1. Die interne Konsistenz ist $\alpha_{VKS_UN_OhneUNEIS2} = .711$ als ausreichend zu bewerten.

Physikalisches Fachwissen der Tutoren bezüglich Elektrizitätslehre

Der Fachwissenstest besteht aus vier Items, deren Aussagen als richtig oder falsch bewertet werden müssen und zwei Items mit offenen Fragen (siehe Kapitel 8.4.5). Die Hauptkomponentenanalyse ergab zwei Hauptkomponenten, welche die Antwortformate widerspiegeln. Die Hauptkomponenten ergaben sowohl einzeln als auch gemeinsam mangelhafte interne Konsistenzen $< .6$. Zusammengenommen konnte aus den sechs Fachwissensitems keine Komponente extrahiert werden, weshalb dieses Konstrukt in den künftigen Analysen nicht berücksichtigt wird. Da das physikalische Fachwissen der Tutoren keine der zentralen Variablen ist, können alle Hypothesen untersucht werden.

Fragebogen zur intrinsischen Motivation bezüglich der Bearbeitung der Aufgaben/ des Experiments vor der Durchführung des Experiments

Die Hauptkomponentenanalyse ergab die erwarteten zwei Hauptkomponenten Interesse und Herausforderung. Die fünf Items der Facette Interesse bilden mit $\alpha_{FAM_Interesse} = .750$ eine ausreichend intern konsistente Skala. Die Items der Facette Herausforderung bilden mit $\alpha_{FAM_Herausforderung} = .552$ keine intern konsistente Skala. Somit ergibt sich die intrinsische Motivation vor der Durchführung des Experiments aus den fünf Items der Facette Interesse (siehe auch Kapitel 8.4.4 und Ryan und Deci, 2017).

9.2.6 Zusammenfassung der internen Konsistenzen und inhaltliche Einordnung

Als Hinweise auf die Reliabilitäten der eingesetzten Instrumente werden im Folgenden ihre internen Konsistenzen zusammenfassend dargestellt. Da die Stichprobengröße teilweise stark variieren, werden diese ebenfalls aufgeführt.

Der beobachtete autonomieunterstützende Motivierstil, die Fachwissensskala und die Skala der introjizierten beziehungsweise externalen Regulation der extrinsischen Tutormotivation hatten keine hinreichenden internen Konsistenzen von mindestens $.70$. Diese vier Konstrukte werden im Folgenden nicht berücksichtigt. Somit können die Hypothesen H3b und H3c nur mit einer eingeschränkten Auswahl an Konstrukten untersucht werden. Details werden in Kapitel 10.3 dargestellt. Alle anderen Konstrukte haben hinreichende bis

sehr gute interne Konsistenzen. Da alle anderen Konstrukte mit hinreichender Güte erfasst werden konnten (siehe auch Kapitel 8.4), können alle anderen Hypothesen planmäßig ausgewertet werden.

Konstrukt Tutoren	Itemanzahl	<i>N</i>	$\alpha_{Cronbach}$
Selbstberichteter autonomieunterstützender Motivierstil			
Prä	12	174	.781
Post	12	169	.907
Selbstberichteter kontrollierender Motivierstil			
Prä	12	175	.772
Post	12	169	.920
Beobachteter autonomieunterstützender Motivierstil	—	—	—
Beobachteter kontrollierender Motivierstil	4	131	.705
Autonome Kausalitätsorientierung			
Prä	12	175	.770
Post	12	171	.788
Kontrollierende Kausalitätsorientierung			
Prä	12	176	.789
Post	12	171	.856
Zustimmung zu den Interventionsinhalten	9	170	.907
Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung	5	172	.851
Fähigkeitsselbstkonzept bezüglich Physik	5	170	.926
Selbstbestimmtheit der Betreuungsmotivation (Intrinsisch)	4	170	.831
Selbstbestimmtheit der Betreuungsmotivation (Identifiziert)	3	170	.753
Selbstbestimmtheit der Betreuungsmotivation (introjiziert)	—	—	—
Selbstbestimmtheit der Betreuungsmotivation (extrinsisch)	—	—	—
Wahrgenommenes Engagement (Druck von Unten)	2	134	.650 ¹
Wahrgenommene Selbstständigkeit (Druck von Unten)	2	133	.747 ¹
Wahrgenommene intrinsische Motivation (Druck von Unten)	3	135	.846
Fachwissen	—	—	—
Verständnis der Variablen-Kontroll-Strategie	3	160	.711
Konstrukt Tutees			
Intrinsische Motivation vor der Durchführung des Experiments	6	135	.784
Wahrgenommener autonomieunterstützender Motivierstil	6	135	.723
Intrinsische Motivation	6	133	.898
Kompetenzerleben	5	135	.848

Tabelle 9.9: Übersicht über die internen Konsistenzen der aus den Hauptkomponentenanalysen resultierenden Skalen; ¹Das wahrgenommene Engagement sowie die wahrgenommene Selbstständigkeit wurden mit je zwei Items erfasst, weshalb keine interne Konsistenz bestimmt werden konnte. Deshalb wurden ersatzweise jeweils die Korrelationen der beiden Items als Gütemaß zur Skalenbildung herangezogen. Die Items korrelieren jeweils grenzwertig beziehungsweise hinreichend hoch miteinander ($\rho_{TW_En} = .650, p < .001$ sowie $\rho_{TW_Se} = .747, p < .001$).

9.3 Untersuchung der Verteilung der Skalen und geplante Auswertungsverfahren

Um nach Field (2009) das geeignete statistische Verfahren (parametrisch oder nicht-parametrisch) für die Beantwortung der Forschungsfragen wählen zu können, müssen die Verteilungen der Daten überprüft werden. Im folgenden Unterkapitel wird überprüft, welche Skalen als weitestgehend normalverteilt angesehen werden können. Weichen einzelne Skalen stark von der Normalverteilung ab, können nicht-parametrische Testverfahren genutzt werden oder die Ergebnisse parametrischer Testverfahren durch ein nicht-parametrisches Bootstrapping-Verfahren abgesichert werden. Der konkrete Umgang mit etwaigen Verletzungen der Normalverteilungen wird am Ende des Kapitels vorgestellt. Für gängige statistische Analyseverfahren, darunter auch lineare Modelle und prä-post-Vergleiche, kann eine Robustheit gegenüber Abweichungen der Normalverteilung für Stichprobengrößen über dreißig Probanden angenommen werden (Horton, 1978). Deshalb und aufgrund der Vielzahl der Tests wird für die Verfahren zur Überprüfung der Normalverteilung ein strenges Signifikanzniveau von .001 angesetzt. Die ersten beiden Verfahren zur Überprüfung der Datensätze auf Normalverteilung sind der Shapiro-Wilk-Test und der Kolmogorov-Test (Field, 2009; Shapiro & Wilk, 1965). Weiterhin werden für alle Skalen die Schiefe und Kurtosis herangezogen (George & Mallery, 2010). Je höher der Quotient aus dem Betrag der Schiefe bzw. Kurtosis und dem jeweiligen Standardfehler ist, desto stärker weicht die Verteilung von einer Normalverteilung ab (George & Mallery, 2010). Bei einem Signifikanzniveau von .05 müsste der Quotient über einem Wert von 1.96 liegen, um eine Abweichung zu implizieren (George & Mallery, 2010). Bei dem angesetzten Signifikanzniveau von .001 weisen die Quotienten auf eine Abweichung von der Normalverteilung hin, wenn sie über 3.29 liegen (Field, 2009). In der nachfolgenden Tabelle 9.10 sind vier Indizien für die Überprüfung der Normalverteilung angegeben.

Der (postinterventionelle) selbstberichtete kontrollierende Motivierstil sowie der wahrgenommene autonomieunterstützende Motivierstil sind näherungsweise normalverteilt. Der postinterventionelle selbstberichtete autonomieunterstützende Motivierstils sowie der beobachtete kontrollierende Motivierstil weichen leicht von der Normalverteilung ab, weshalb bei der Beantwortung der ersten Forschungsfrage alle parametrischen Pearson-Korrelationen zwischen den Perspektiven der Motivierstile durch nicht-parametrische Spearman-Korrelationen abgesichert werden.

Die prä- beziehungsweise postinterventionellen selbstberichteten Motivierstile sind näherungsweise normalverteilt, weshalb für Forschungsfrage 2 T-Tests gerechnet werden können. Aufgrund der oben berichteten leichten Abweichungen des postinterventionellen selbstberichteten autonomieunterstützenden Motivierstils, werden ebenfalls Mann-Whitney-U-Tests gerechnet. Die Zeit zwischen Intervention und Messzeitpunkt 2 sowie die Zustimmung zu den Interventionsinhalten weisen starke Abweichungen von der Normalverteilung auf. Die vorherige Bekanntheit der Interventionsinhalte weicht leicht von der Normalverteilung ab. Deshalb werden die parametrischen Regressionen zur Analyse der Prädiktoren der Interventionseffekte durch ein Bootstrapping-Verfahren abgesichert. Weiterhin werden für jede Regression die Probanden mit residualen Abweichungen über drei Standardabweichungen für die Rechnung entfernt (Field, 2009).

Da die intrinsische und identifizierte Tutormotivation sowie die wahrgenommene Tuteeselbstständigkeit und Tuteemotivation leicht von der Normalverteilung abweichen und das Verständnis der Variablenkontrollstrategie und des wahrgenommenen Tuteeengagements

9 Deskriptive Ergebnisse

Konstrukte	N	M(SD)	Q. Schiefe	Q. Kurtosis	p_{K.S.}	p_{S.W.}
Tutoren - Demografische Daten						
Geschlecht (0 = weiblich/ 1 = männlich)	176	.528 (.501)	0.628	5.522	<.001	<.001
Mathematiknote (von 1 - 6)	175	2.38 (1.05)	2.348	0.973	<.001	<.001
Deutschnote (von 1 - 6)	175	2.49 (0.82)	0.196	1.334	<.001	<.001
Physiknote (von 1 - 6)	174	2.01 (0.87)	2.255	1.899	<.001	<.001
Tutoren						
Selbstberichteter autonomieunterstützender Motivierstil präinterventionell	174	.626 (.102)	2.418	3.423	>.200	.016
Selbstberichteter autonomieunterstützender Motivierstil (postinterventionell)	169	.768 (.119)	2.444	0.089	.003	.011
Selbstberichteter autonomieunterstützender Motivierstil (prä-post-Differenz)	167	.144 (.117)	0.718	1.85	>.200	0.357
Selbstberichteter kontrollierender Motivierstil (präinterventionell)	175	.602 (.097)	0.701	0.378	>.200	>.200
Selbstberichteter kontrollierender Motivierstil (postinterventionell)	169	.372 (.158)	0.487	0.447	>.200	>.200
Selbstberichteter kontrollierender Motivierstil (prä-post-Differenz)	169	-.230 (.146)	0.316	0.795	>.200	>.200
Beobachteter kontrollierender Motivierstil	131	.249 (.148)	3.627	0.361	.009	<.001
Autonome Kausalitätsorientierung (präinterventionell)	175	.782 (.099)	2.429	2.397	>.200	.020
Autonome Kausalitätsorientierung (postinterventionell)	171	.790 (.105)	1.43	0.404	>.200	.066
Kontrollierende Kausalitätsorientierung (präinterventionell)	176	.484 (.138)	1.322	0.255	>.200	>.200
Kontrollierende Kausalitätsorientierung (postinterventionell)	171	.442 (.163)	0.398	0.753	>.200	>.200
Zustimmung zu den Interventionsinhalten	170	.762 (.150)	4.511	1.27	<.001	<.001
Vorherige Bekanntheit der Interventionsinhalte	170	.411 (.262)	0.382	2.189	<.001	<.001
Zeit zwischen Intervention und Messzeitpunkt 2	177	13.04 (6.00)	7.246	4.314	<.001	<.001
Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung	172	.682 (.122)	0.978	0.91	<.001	.032
Fähigkeitsselbstkonzept bezüglich Physik	170	.728 (.162)	1.516	0.268	.017	.003
Selbstbestimmtheit der Betreuungsmotivation (intrinsisch)	170	.776 (.161)	2.935	0.005	<.001	<.001
Selbstbestimmtheit der Betreuungsmotivation (identifiziert)	170	.584 (.231)	2.032	1.073	<.001	.001
Verständnis der Variablen-Kontrollstrategie	160	.625 (.381)	2.443	3.278	<.001	<.001
Wahrgenommenes Engagement (Druck von Unten)	134	.813 (.157)	8.469	14.954	<.001	<.001
Wahrgenommene Selbstständigkeit (Druck von Unten)	134	.707 (.238)	3.617	0.195	<.001	<.001
Wahrgenommene intrinsische Motivation (Druck von Unten)	135	.652 (.221)	2.852	0.295	<.001	<.001
Tutees						
Intrinsische Motivation vor der Durchführung des Experiments	135	.722 (.166)	4.852	4.633	<.001	<.001
Wahrgenommener autonomieunterstützender Motivierstil	135	.715 (.169)	2.344	0.556	.015	.005
Intrinsische Motivation	133	.874 (.161)	10.062	13.564	<.001	<.001
Kompetenzerleben	135	.788 (.168)	5.770	4.072	<.001	<.001

Tabelle 9.10: Übersicht über die Mittelwerte, die Quotienten aus den Schiefen und den zugehörigen Standardfehlern, die Quotienten aus den Kurtosisen und den zugehörigen Standardfehlern sowie der Signifikanz des Kolmogorov-Smirnov- und des Shapiro-Wilk-Tests. Alle Skalen haben einen Wertebereich von 0 bis 1. Insbesondere wurden Items mit einer 7-stufigen Likert-Skala entsprechend $Skalenwert = (SummederItems - AnzahlItems)/(6 \cdot AnzahlItems)$ normiert. Ausnahme sind die Noten (1-6) und die Zeit zwischen Intervention und Posterhebung, welche in Minuten angegeben ist. Weiterhin haben die Mittelwertsunterschiede der Motivierstile einen Wertebereich von -1 bis 1. Indizien für eine Abweichung der Normalverteilung ($p < .001$) sind fett gedruckt.

stark von der Normalverteilung abweichen, werden die kovarianzbasierten Strukturgleichungsmodelle zur Beantwortung der Forschungsfrage 3 mit dem MLM-Schätzer gerechnet (Satorra & Bentler, 1994). Das Kompetenzerleben sowie die intrinsische Motivation vor und nach der Durchführung des Experiments weisen starke Abweichungen von der Normalverteilung auf. Deshalb werden die kovarianzbasierten Strukturgleichungsmodelle zur Beantwortung der vierten Forschungsfrage mit dem MLM-Schätzer gerechnet (Satorra & Bentler, 1994). Die Mediationsanalysen werden durch nicht-parametrische Bootstrapping-Verfahren abgesichert. Für die Mittelwertvergleiche zur Beantwortung von Forschungsfrage 4b werden neben T-Tests auch Mann-Whitney-U-Tests gerechnet.

9.4 Vergleichbarkeit der Substichproben zur Bildung einer Gesamtstichprobe

Um zu untersuchen, ob die Teilstichproben der neun Schulen für die Analysen zusammengefasst werden können, werden Unterschiede der demographischen und präinterventionellen Tutordaten gesucht. Das Geschlecht kann nach Reeve (1998) möglicherweise mit den Motivierstilen der Tutoren zusammenhängen. Als Indiz möglicher Unterschiede im Fachwissen Physik wurde die letzte Physiknote erhoben. Weiterhin wurde die letzte Zeugnisnote im Fach Mathematik erhoben. Da die Fragebögen und die Intervention einige Texte umfassen, könnte das Leseverständnis der Tutoren beeinflussen, inwiefern die Tutoren von der Intervention profitieren können (z. B. Hartmann, 2013). Als Indiz für das Leseverständnis wird die Deutschnote erfasst (Gröhlich & Guill, 2009).

Um mit Hilfe einer ANOVA überprüfen zu können, inwiefern die Tutoren der neun Schulen zu einer Gesamtstichprobe zusammengefasst werden können, muss zuvor die Varianzhomogenität der Variablen in den neun Gruppen gewährleistet werden. Gemäß dem Levene-Test kann die Varianzhomogenität der Mathematiknote, der Physiknote, der Kausalitätsorientierungen und der Motivierstile angenommen werden ($p > .05$). Die Deutschnote und das Geschlecht erfüllen dieses Kriterium nicht. Die signifikant unterschiedlich verteilte Deutschnote könnte ein Indiz für ein unterschiedlich ausgeprägtes Leseverständnis sein (Gröhlich & Guill, 2009). Folglich wird die Deutschnote vor allem bei der statistischen Analyse der zweiten Forschungsfrage als Kontrollvariable berücksichtigt. Die Geschlechter sind unterschiedlich verteilt, weshalb das Geschlecht ebenfalls als Kontrollvariable berücksichtigt wird. Tabelle 9.11 stellt die Ergebnisse einer ANOVA der Variablen in Abhängig-

keit der Schulnummer dar. Es gab keinen statistisch signifikanten Unterschied zwischen den Schülerinnen und Schülern der verschiedenen Schulen bezüglich ihrer Mathematiknote ($F(8, 166) = 1.33, p = .234$) und ihres Physikvorwissens (gemessen durch die Physiknote, $F(8, 165) = 0.95, p = .479$). Da sich die Mathematik- und Physiknoten in Abhängigkeit der Schulen nicht signifikant unterscheiden, werden über die Schulen hinweg ähnliche Mathematikvorkenntnisse und Physikvorkenntnisse vorausgesetzt. Weiterhin gab es keine statistisch signifikanten Unterschiede zwischen den Schülerinnen und Schüler der verschiedenen Schulen bezüglich ihrer Motivierstile ($F(8, 165) = 0.56, p = .812$ beziehungsweise $F(8, 166) = 0.98, p = .457$) und ihrer Kausalitätsorientierungen ($F(8, 166) = 0.50, p = .856$ beziehungsweise $F(8, 167) = 1.63, p = .121$). Insgesamt werden die neun Stichproben unter Berücksichtigung des Geschlechts und der Deutschnote zu einer Stichprobe zusammengefasst.

Variable	<i>p</i> Levene	F-Statistik	<i>p</i>
Geschlecht	<.001	F(8, 167) = 3.15	.002
Mathematiknote	.051	F(8, 166) = 1.33	.234
Deutschnote	.018	F(8, 166) = 2.32	.022
Physiknote	.746	F(8, 165) = 0.95	.479
Präinterventioneller autonomieunterstützender Motivierstil	.061	F(8, 165) = 0.56	.812
Präinterventioneller kontrollierender Motivierstil	.070	F(8, 166) = 0.98	.457
Präinterventionelle autonome Kausalitätsorientierung	.148	F(8, 166) = 0.50	.856
Präinterventioneller kontrollierende Kausalitätsorientierung	.830	F(8, 167) = 1.63	.121

Tabelle 9.11: ANOVAs bezüglich der Unterschiede verschiedener Konstrukte zwischen den neun Schulen

Wie in Kapitel 8.5 beschrieben, nahmen die insgesamt 176 Probanden verteilt auf fünf Termine an der Intervention teil. Da die Anzahl der Teilnehmenden beim ersten, zweiten, dritten und fünften Termin über 35 war, wurden die Gruppen für die Vorträge und die Rollenspiele mit Beobachtern (siehe Kapitel 8.3) in zwei Gruppen geteilt. 93 der 176 Probanden waren in der Gruppe des Autors der vorliegenden Studie. Sowohl T-Tests als auch Mann-Whitney-U-Tests bestätigen (siehe Anhang Kapitel 12.3.2), dass diese beiden Probandengruppen keine signifikanten Unterschiede aufweisen und deshalb zu einer Gesamtgruppe zusammengefasst werden können.

Die Auswahl der Analysestichproben resultierte aus fehlerhaften Datenerhebungen sowie statistischen Empfehlungen. Beispielsweise mussten von den 135 aufgenommenen Videos vier aufgrund von audiovisuellen Problemen aussortiert werden (siehe Kapitel 9.1). Für die Hypothesen ergeben sich somit folgende Analysestichproben (siehe Tabelle 9.12):

Name	Anzahl Tutoren	Anzahl Tutees	Hypothesen
Stichprobe 1	176	—	H1a, H2
Stichprobe 2	131	116/131	H3
Stichprobe 3	—	133	H4
Stichprobe 4	131	131	H1b

Tabelle 9.12: Stichproben zur Untersuchung der Hypothesen

10 Ergebnisse zur Beantwortung der Forschungsfragen

Im Folgenden werden die Rechnungen zur Beantwortung der Forschungsfragen präsentiert.

10.1 Ergebnisse der Forschungsfrage 1

Entsprechend der ersten Forschungsfrage wurde einerseits untersucht, inwiefern die durch Selbstberichte erfassten Motivierstile der Tutoren mit ihrem tatsächlichen Verhalten zusammenhängen. Andererseits wurde untersucht, inwiefern ihr Verhalten, das theoretisch als autonomieunterstützend klassifiziert wird, als solches von den Tutees wahrgenommen wurde. Da der wahrgenommene kontrollierende Motivierstil nicht erhoben wurde, liegen mit den drei Messperspektiven des autonomieunterstützenden und zwei Arten des kontrollierenden Motivierstils fünf Variablen vor. Korreliert man alle fünf miteinander, ergeben sich zehn Korrelationen, weshalb das übliche Signifikanzniveau von .001 entsprechend der Bonferonni-Korrektur auf ein alphasfehlerkorrigiertes Signifikanzniveau von $.001/10 = .0001$ herabgesetzt wird (Armstrong, 2014). Da einige Skalen näherungsweise normalverteilt sind und andere nicht (siehe Kapitel 9.3), wurden sowohl die parametrischen Korrelationen nach Pearson als auch die nicht-parametrischen Korrelationen nach Spearman berechnet und in Tabelle 10.1 aufgelistet.

Variable	1.	2.	3.	4.
<i>Autonomieunterstützender Motivierstil</i>				
1. Selbstbericht Tutoren	—			
2. Analysiertes Verhalten	—	—		
3. Selbstbericht Tutees	-.07/-.04	—	—	
<i>Kontrollierender Motivierstil</i>				
4. Selbstbericht Tutoren	-.33*** / -.37***	—	.15/.15	—
5. Analysiertes Verhalten	-.04/-.09	—	.16/.16	.48*** / .49***

Tabelle 10.1: Pearson Korrelationen r / Spearman Korrelationen ρ der Perspektiven der Motivierstile; * $p < .005$ (Alphafehler korrigiertes .05) *** $p < .0001$ (Alphafehler korrigiertes .001). Alle übrigen Korrelationen haben ein Signifikanzniveau $> .05$

Alle Korrelationen mit dem beobachteten autonomieunterstützenden Motivierstil fehlen, da die Hauptkomponentenanalyse in Kapitel 9.2 gezeigt hat, dass aus den verschiedenen Verhaltensweisen der Tutoren keine Skala autonomieunterstützenden Verhaltens gebildet werden konnte. Die beiden gefundenen Korrelationen haben Signifikanzniveaus unter der alphasfehlerkorrigierten Grenze von .0001. Alle übrigen Korrelationen haben ein Signifikanzniveau über .05, also wären sie auch ohne Alphafehlerkorrektur nicht signifikant.

Entsprechend der signifikanten Korrelationen halten Tutoren ein kontrollierendes Verhalten für umso angemessener, je weniger sie ein autonomieunterstützendes Verhalten für angemessen erachten. Weiterhin verhalten sich Tutoren in einer tatsächlichen Betreuungssituation umso kontrollierender, je mehr sie kontrollierendes Verhalten als angemessen empfinden.

Der selbstberichtete kontrollierende Motivierstil schlägt sich also auf das Verhalten in einer tatsächlichen Betreuungssituation nieder. Ob es von den Tutees auch als kontrollierend wahrgenommen wird, konnte nicht untersucht werden, da der wahrgenommene kontrollierende Motivierstil nicht erhoben wurde. Inwiefern sich der selbstberichtete autonomieunterstützende Motivierstil auf das Verhalten auswirkt und ob das autonomieunterstützende Verhalten als solches wahrgenommen wird, konnte ebenfalls nicht untersucht werden. Es wurde allerdings keine Korrelation zwischen dem selbstberichteten autonomieunterstützenden Motivierstil der Tutoren und dem wahrgenommenen autonomieunterstützenden Motivierstil gefunden. Somit scheinen sich Tutees von Tutoren, die ein autonomieunterstützendes Verhalten als angemessen empfinden, nicht autonomieunterstützender betreut zu fühlen.

10.2 Ergebnisse der Forschungsfrage 2

Entsprechend den Hypothesen H2a₁ und H2b₁ wird in Kapitel 10.2.1 untersucht, in welchem Umfang eine Intervention den autonomieunterstützenden Motivierstil fördern und den kontrollierenden Motivierstil reduzieren kann. In Kapitel 10.2.2 werden die Prädiktoren der Mittelwertsunterschiede und in Kapitel 10.2.3 etwaige differentielle Effekte untersucht.

10.2.1 Die Mittelwertunterschiede der Motivierstile und Kausalitätsorientierungen vor und nach der Intervention

Variable	Mittelwertdifferenz (SD)	T-Statistik	d_{Cohen}	p	$Z_{Wilcoxon}$	r	p
SMot_A	.144 (.117)	15.900	1.230	<.001	-10.221	-.793	<.001
SMot_K	-.223 (.146)	-20.516	-1.578	<.001	-10.983	-.847	<.001
Kau_A	.006 (.075)	0.978	0.075	.329	-0.966	-.074	.334
Kau_K	-.039 (.094)	-5.432	-0.415	<.001	-4.916	-.377	<.001

Tabelle 10.2: Mittelwertunterschiede der Motivierstile und Kausalitätsorientierungen vor und nach der Intervention sowie die statistische Signifikanz dieser Unterschiede; SMot_A/K = Selbstberichteter autonomieunterstützender/ kontrollierender Motivierstil; Kau_A/K = autonome/ kontrollierende Kausalitätsorientierung.

Da die Gruppen der beiden Experimentatoren zusammengefasst werden können (siehe Kapitel 9.4 und Kapitel 12.3.2), wurden die Prä-Post-Unterschiede der Motivierstile und der Kausalitätsorientierungen für alle Probanden gemeinsam ermittelt (Analog z. B. Aeltermann et al., 2014 und Reeve und Cheon, 2016). Aufgrund der leichten Abweichungen des postinterventionellen autonomieunterstützenden Motivierstils von der Normalverteilung (siehe Kapitel 9.3), wurden neben T-Tests auch Mann-Whitney-U-Tests gerechnet.

Die Ergebnisse der Prä-Post-Mittelwertvergleiche sind in Tabelle 10.2 zusammengefasst und in Abbildung 10.1 dargestellt.

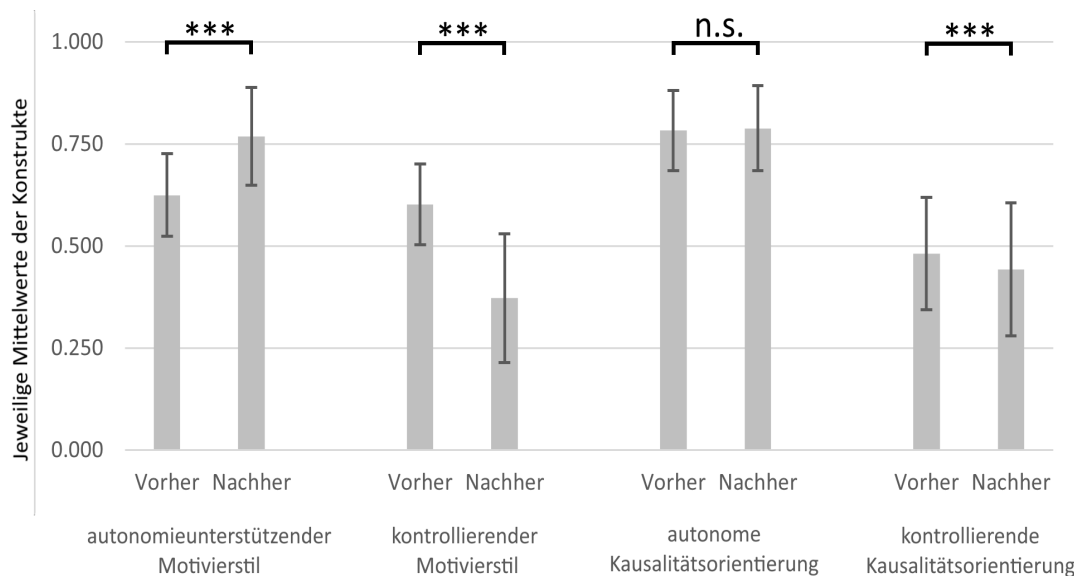


Abbildung 10.1: Jeweilige Mittelwerte des autonomieunterstützenden und des kontrollierenden Motivierstils sowie der autonomen und der kontrollierenden Kausalitätsorientierung. Fehlerbalken stellen zwei Standardabweichungen dar, also eine in jede Richtung. *** $p < .001$.

Alle Skalen haben einen Wertebereich von 0 bis 1. Der Mittelwert des autonomieunterstützenden Motivierstils steigt von .62 vor der Intervention auf .77 nach der Intervention, was nach J. Cohen (1988) einer Effektstärke von $|d_{Cohen}|=1.23$, $p < .001$ entspricht. Der Mittelwert des kontrollierenden Motivierstils ist von .60 auf .37 gefallen, was einer Effektstärke von $|d_{Cohen}|=1.58$, $p < .001$ entspricht. Die Mittelwertvergleiche zeigen keine signifikanten Änderungen der autonomen Kausalitätsorientierung. Der Mittelwert ändert sich von .78 auf .79, was einer Effektstärke von $|d_{Cohen}|=0.08$, $p = .329$ entspricht. Die Mittelwerte der kontrollierenden Kausalitätsorientierung sinken von .48 vor der Intervention auf .44 nach der Intervention. Diese Änderung entspricht einem kleinen bis mittleren Effekt von $|d_{Cohen}|=0.415$, $p < .001$ (J. Cohen, 1988). Die Mann-Whitney-U-Tests unterstützen die Befunde der T-Tests.

10.2.2 Untersuchung der Prädiktoren der Mittelwertsunterschiede

Entsprechend den Hypothesen H2a_{2.I-III} und H2b_{2.I-III} wurde weiterhin untersucht, inwiefern die Kausalitätsorientierungen der Probanden, ihre Zustimmung zu den Interventionen und die vorherige Bekanntheit dieser Inhalte den Interventionseffekt vorhersagen. Wie in Kapitel 8.1 angedeutet, konnte die Zeit zwischen der Intervention und Messzeitpunkt 2 aus schulorganisatorischen Gründen nicht schulübergreifend konstant gehalten werden. Reeve (1998) führte im Anschluss an seine Interventionsstudie einen Monat später eine Follow-Up-Erhebung durch. Er konnte für Lehramtsstudierende und Lehrende zeigen, dass die Interventionsgruppe einen signifikant autonomieunterstützenderen bilanzierten Motivierstil beibehalten hat. Weiterhin interagiert der Interventionseffekt nicht

mit der Wartezeit von einem Monat Reeve (1998). Somit konnte er nachweisen, dass der Interventionseffekt über einen Monat hinweg stabil geblieben ist. Da die Zeit zwischen Intervention und Messzeitpunkt 2 für die Teilnehmenden zwischen minimal vier Tagen und maximal achtundzwanzig Tagen variiert, wird vermutet, dass diese Zeit in der vorliegenden Studie nicht die Änderungen der Motivierstile vorhersagt. Nichtsdestotrotz wurde diese Zeit als weitere Kontrollvariable für Forschungsfrage 2 berücksichtigt. Weiterhin waren das Geschlecht der Tutoren sowie ihre Deutschnote zwischen den verschiedenen Schulen nicht vergleichbar (siehe Kapitel 8.5). Die signifikant unterschiedlich verteilte Deutschnote könnte ein Indiz für ein unterschiedlich ausgeprägtes Leseverständnis sein (Gröhllich & Guill, 2009). Entsprechend Reeve (1998) berichten weibliche Probanden einen signifikant autonomieunterstützenderen Motivierstil, wodurch sie möglicherweise ein geringeres Entwicklungspotential haben. Folglich werden das Geschlecht und die Deutschnote als weitere Kontrollvariablen berücksichtigt. Aus den zusätzlichen Kontrollvariablen ergeben sich folgende zusätzliche Zusammenhangshypothesen:

- H2a_{2.IV}: Der Zuwachs des selbstberichteten autonomieunterstützenden Motivierstils ist unabhängig von der vergangenen Zeit zwischen der Intervention und der Posterhebung, also dem zweiten Messzeitpunkt.
- H2b_{2.IV}: Die Reduktion des selbstberichteten kontrollierenden Motivierstils ist unabhängig von der vergangenen Zeit zwischen der Intervention und der Posterhebung, also dem zweiten Messzeitpunkt.
- H2a_{2.V}: Je besser die Deutschnote der Tutoren ist, desto mehr nimmt ihr selbstberichteter autonomieunterstützender Motivierstil zu.
- H2b_{2.V}: Je besser die Deutschnote der Tutoren ist, desto mehr nimmt ihr selbstberichteter kontrollierender Motivierstil ab.
- H2a_{2.VI}: Der selbstberichtete autonomieunterstützende Motivierstil von weiblichen Teilnehmenden nimmt durch eine Intervention weniger zu.
- H2b_{2.VI}: Der selbstberichtete kontrollierende Motivierstil von weiblichen Teilnehmenden nimmt durch eine Intervention weniger ab.

In den Tabellen 10.3 und 10.4 sind hierarchische lineare Regressionen sowie zugehörige Kennwerte aufgelistet, bei denen die Änderungen der Motivierstile als die abhängigen Variablen gewählt wurden. Da hierarchische lineare Regressionen durchgeführt wurden, wurden die unabhängigen Variablen entsprechend ihrer Wichtigkeit schrittweise hinzugefügt (Field, 2009). In den ersten beiden Schritten wurden die Kontrollvariablen hinzugefügt, wobei im ersten Schritt lediglich die Kontrollvariablen Geschlecht und Deutschnote entsprechend Hypothese H2a_{2.V-VI} und H2b_{2.V-VI} berücksichtigt wurden. Die Zeit zwischen der Intervention und Messzeitpunkt 2 wurde erst im zweiten Schritt hinzugenommen (H2a_{2.IV} und H2b_{2.IV}), um die Änderung der aufgeklärten Varianz sichtbar zu machen (Field, 2009). Das dritte Modell berücksichtigt gemäß H2a_{2.I} und H2b_{2.I} die Kausalitätsorientierungen der Interventionsteilnehmenden als die ersten vermuteten Prädiktoren. Entsprechend H2a_{2.II-III} und H2b_{2.II-III} wurden im vierten Schritt die Zustimmung zu den Interventionsinhalten und die selbst angegebene vorherige Kenntnis dieser Inhalte als weitere Prädiktoren hinzugefügt. Entsprechend Field (2009) wurden Probanden mit residualen Abweichungen über drei Standardabweichungen als Ausreißer ausgeschlossen und die Rechnung wiederholt. Dieser Prozess wurde so lange wiederholt, bis keine weiteren Probanden mehr ausreißende Abweichungen aufwiesen. Für die Regressionen bezüglich der Änderung des selbstberichteten autonomieunterstützenden Motivierstils mussten deshalb drei bzw.

vier Probanden ausgeschlossen werden. Für die Regressionen bezüglich der Änderung des kontrollierenden Motivierstils mussten keine Probanden ausgeschlossen werden. Weiterhin empfiehlt Field (2009), die Rechnungen zu wiederholen und im abschließenden Modell lediglich die unabhängigen Variablen zu berücksichtigen, die im vollständigen Modell signifikante Prädiktoren waren. Da das Geschlecht der Tutoren, die Zeit zwischen Intervention und Messzeitpunkt 2 sowie die Zustimmung zu den Interventionsinhalten starke Abweichungen von der Normalverteilung haben und die vorherige Bekanntheit der Interventionsinhalte eine leichte Abweichung von der Normalverteilung aufweist, werden die Signifikanzniveaus der Regressionen durch Bootstrapping-Verfahren abgesichert (Field, 2009).

Nr	Variable	$\Delta\text{SMot_A}$			$\Delta\text{SMot_K}$		
		β	p	$p_{\text{Bootstrap}}$	β	p	$p_{\text{Bootstrap}}$
1	(Konstante)		<.001	<.001		<.001	<.001
	Geschlecht	-.074	.368	.370	.106	.186	.171
	Deutschnote	-.076	.350	.347	.035	.661	.677
2	(Konstante)		<.001	<.001		<.001	<.001
	Geschlecht	-.068	.401	.409	.094	.226	.206
	Deutschnote	-.073	.371	.352	.033	.675	.688
	Zeit	-.137	.085	.095	.256	.001	.002
3	(Konstante)		.185	.094		.047	.014
	Geschlecht	-.052	.533	.548	.064	.422	.440
	Deutschnote	-.063	.441	.427	.019	.810	.823
	Zeit	-.131	.109	.110	.257	.001	.002
	Kau_A_Prä	.082	.316	.318	-.092	.235	.222
	Kau_K_Prä	-.034	.683	.665	.093	.237	.241
4	(Konstante)		.244	.104		.503	.191
	Geschlecht	-.017	.825	.845	.041	.586	.588
	Deutschnote	-.142	.067	.052	.066	.373	.381
	Zeit	-.117	.120	.166	.239	.001	.002
	Kau_A_Prä	.016	.840	.842	.002	.983	.982
	Kau_K_Prä	-.023	.769	.748	.043	.563	.553
	Zustimmung	.260	.001	.003	-.362	<.001	<.001
	Bekanntheit	-.332	<.001	<.001	.144	.049	.040
5	(Konstante)		.138	.135		.283	.262
	Zeit	—	—	—	.235	.001	<.001
	Zustimmung	.300	<.001	<.001	-.379	<.001	<.001
	Bekanntheit	-.326	<.001	<.001	.134	.057	.037

Tabelle 10.3: Standardisierte Gewichte, Signifikanzniveaus sowie durch ein Bootstrapping-Verfahren mit 10000 Stichproben abgesicherte Signifikanzniveaus der Regressionen der Änderungen der Motivierstile auf vermutete Prädiktoren; $\Delta\text{SMot_A/K}$ = Änderung des autonomieunterstützenden/ kontrollierenden Motivierstils; Zeit = Zeit zwischen Intervention und der Posterhebung; Kau_A/K_Prä = präinterventionelle autonome/ kontrollierende Kausalitätsorientierung; Zustimmung = Zustimmung zu den Interventionsinhalten; Bekanntheit = Vorherige Bekanntheit der Interventionsinhalte.

Nr	$\Delta\text{SMot_A}$			$\Delta\text{SMot_K}$		
	$R^2(\Delta R^2)$	F-Statistik	$p_{\text{Modell}}(p_{\Delta R})$	$R^2(\Delta R^2)$	F-Statistik	$p_{\text{Modell}}(p_{\Delta R})$
1	.014(.014)	1.06	.349(.349)	.014(.014)	1.14	.322(.322)
2	.032(.019)	1.72	.166(.085)	.079(.065)	4.56	.004(.001)
3	.040(.007)	1.26	.285(.558)	.095(.016)	3.30	.007(.255)
4	.192(.152)	5.10	<.001(<.001)	.225(.130)	6.43	<.001(<.001)
5	.178	16.88	<.001	.220	15.11	<.001

Tabelle 10.4: Liste der aufgeklärten Varianzen, F-Statistiken und Signifikanzniveaus der Regressionsmodelle, welche in Tabelle 10.3 dargestellt sind. $\Delta\text{SMot_A/K}$ = Änderung des autonomieunterstützenden/ kontrollierenden Motivierstils.

Bezüglich der Änderung des autonomieunterstützenden Motivierstils ist die Regression mit den fünf Prädiktoren Geschlecht, Deutschnote, Zeit zwischen Intervention und Messzeitpunkt 2 sowie den Kausalitätsorientierungen mit $F(5, 152) = 1.26, p = .285 (R^2 = .04)$ nicht signifikant. Durch Hinzunahme der Zustimmung zu den Interventionsinhalten und die vorherige Kenntnis dieser, wird das Modell mit $F(7, 150) = 5.10, p < .001 (R^2 = .19)$ insgesamt signifikant. Die beiden zuletzt hinzugenommenen unabhängigen Variablen sind jedoch die einzigen signifikanten Prädiktoren des Modells. Entsprechend der oben genannten Empfehlung von Field (2009) wurde die Regression ausschließlich mit den zuvor signifikanten Prädiktoren wiederholt. Das Modell ist mit $F(2, 156) = 16.88, p < .001 (R^2 = .18)$ insgesamt signifikant und beide unabhängigen Variablen bleiben signifikante Prädiktoren. Je stärker die Probanden den Inhalten der Intervention zustimmten, desto größer war der Zuwachs des autonomieunterstützenden Motivierstils ($\beta = .300, p < .001$). Je mehr die Probanden die Interventionsinhalte schon vor der Intervention kannten, desto kleiner war der Zuwachs des autonomieunterstützenden Motivierstils ($\beta = -.326, p < .001$).

Bezüglich der Änderung des kontrollierenden Motivierstils ist die erste Regression mit den Prädiktoren Geschlecht und Deutschnote mit $F(2, 160) = 1.14, p = .322 (R^2 = .01)$ nicht signifikant. Durch Hinzunahme der Zeit zwischen Intervention und Messzeitpunkt 2 wird das Modell mit $F(3, 159) = 4.56, p = .004 (R^2 = .08)$ insgesamt signifikant. Die Kausalitätsorientierungen sind auch für die Änderung des kontrollierenden Motivierstils keine Prädiktoren. Das Gesamtmodell ist mit $F(7, 155) = 6.43, p < .001 (R^2 = .10)$ insgesamt signifikant, wobei die vorherige Bekanntheit der Interventionsinhalte mit einem Signifikanzniveau von $p = .049$ beziehungsweise $p_{\text{Bootstrap}} = .040$ knapp signifikant geworden ist. Da im Gesamtmodell die Zeit zwischen Intervention und Messzeitpunkt 2, die Zustimmung zu den Interventionsinhalten und der vorherigen Bekanntheit dieser die drei signifikanten Prädiktoren waren, wird die Regression entsprechend der Empfehlung von Field (2009) wiederholt. Das finale Modell ist mit $F(3, 161) = 15.11, p < .001 (R^2 = .22)$ insgesamt signifikant und alle drei unabhängigen Variablen bleiben signifikante Prädiktoren. In diesem Modell ist die vorherige Bekanntheit der Interventionsinhalte mit einem parametrisch geschätzten Signifikanzniveau von $p = .057$ knapp nicht mehr signifikant. Allerdings weist das durch Bootstrapping robust geschätzte Signifikanzniveau $p_{\text{Bootstrap}} = .037$ auf Signifikanz hin. Je stärker die Probanden den Inhalten der Intervention zustimmten, desto größer war die Reduktion des kontrollierenden Motivierstils ($\beta = -.379, p < .001$). Je mehr Zeit zwischen der Intervention und der Posterhebung, also dem zweiten Messzeitpunkt, vergangen ist, desto geringer war die Reduktion des kontrollierenden Motivierstils ($\beta = .235, p = .001$). Je mehr die Probanden die Interventionsinhalte schon vor der Intervention kannten, desto geringer war die Reduktion des kontrollierenden Motivierstils ($\beta = .134, p_{\text{Bootstrap}} = .037$).

10.2.3 Die Zustimmung zu den Interventionsinhalten als Prädiktor der Änderungen der Motivierstile für Teilnehmende mit unterschiedlichen Kausalitätsorientierungen

In Anlehnung an die Befunde von Reeve (1998) wurden die Hypothesen H2a_{3,I-II} und H2b_{3,I-II} aufgestellt, dass sich die Relevanz der Zustimmung zu den Interventionsinhalten als Prädiktor der Änderungen der Motivierstile ändert, wenn Teilnehmende mit relativ hoch und relativ niedrig ausgeprägten Kausalitätsorientierungen voneinander abgegrenzt werden. Insbesondere wurde vermutet, dass Teilnehmende mit einer ausgeprägten autonomen Kausalitätsorientierung ohnehin empfänglich für die Interventionsinhalte sind, also unabhängig von ihrer Zustimmung zu den Interventionsinhalten profitieren sollten. Da Teilnehmende mit einer ausgeprägten kontrollierenden Kausalitätsorientierung den Inhalten vermutlich grundsätzlich kritisch gegenüberstehen, müssen sie einen Konzeptwechsel durchlaufen, weshalb die Zustimmung zu den Interventionsinhalten vermutlich besonders wichtig ist, um ihren kontrollierenden Motivierstil zu reduzieren. Um dies zu untersuchen, wurden alle Interventionsteilnehmenden mit Hilfe von Mediansplits auf zwei Gruppen aufgeteilt. Die Mediane entsprechen ungefähr den Mittelwerten ($M_{Kau_A_Prä}=.7815$; $Median_{Kau_A_Prä}=.7778$; $M_{Kau_K_Prä}=.4838$; $Median_{Kau_K_Prä}=.4861$), da die präinterventionellen Kausalitätsorientierungen näherungsweise normalverteilt sind (siehe Kapitel 9.3). Entsprechend ergeben sich die in Tabelle 10.2 dargestellten Teilnehmendengruppen durch Mediansplits:

	Autonome Kausalitätsorientierung	Kontrollierende Kausalitätsorientierung
Kleiner oder gleich dem Median	87	85
Größer als der Median	88	91
Summe Teilnehmende	175	176

Abbildung 10.2: Probandenanzahl in den Mediansplit-Gruppen

Wie bei den Regressionen 10.2.2 wurden die möglichen Prädiktoren der Änderungen der Motivierstile stückweise hinzugenommen. Entsprechend der Empfehlung von Field (2009) wurden die Regressionen abschließend lediglich mit den signifikanten Prädiktoren wiederholt. Diese finalen Regressionsmodelle sind in Tabelle 10.5 und 10.6 dargestellt. Die ausführlichen hierarchischen Regressionsmodelle, ihre Erläuterung sowie detaillierte Angaben zu den Modellgütern können dem Anhang (siehe Kapitel 12.3.3) entnommen werden. Konkret sind in Tabelle 10.5 und 10.6 die linearen Regressionen bezüglich der Änderung des autonomieunterstützenden (oben) beziehungsweise des kontrollierenden (unten) Motivierstils für Teilnehmende mit einer entsprechend hoch (links) beziehungsweise niedrig (rechts) ausgeprägten Kausalitätsorientierung dargestellt.

Das Regressionsmodell der Teilgruppe mit einer hoch ausgeprägten autonomen Kausalitätsorientierung wird mit $F(2, 73) = 10.05, p < .001 (R^2 = .22)$ insgesamt signifikant. Für die Gruppe mit einer niedrig ausgeprägten autonomen Kausalitätsorientierung wird das Regressionsmodell mit $F(2, 79) = 7.21, p = .001 (R^2 = .15)$ insgesamt signifikant. In beiden Gruppen wird die Zustimmung zu den Interventionsinhalten signifikant. Somit gilt für beide Teilgruppen: je stärker die Probanden den Inhalten der Intervention zustimmten, desto größer war der Zuwachs des autonomieunterstützenden Motivierstils ($\beta = .263, p = .014$ bzw. $\beta = .275, p = .010$).

Variable	$\Delta\text{SMot_A}$ (hohe Kau_A_Prä)			$\Delta\text{SMot_A}$ (niedrige Kau_A_Prä)		
	β	p	$p_{\text{Bootstrap}}$	β	p	$p_{\text{Bootstrap}}$
(Konstante)		.172	.046		.216	.213
Zustimmung	.263	.014	.003	.275	.010	.018
Bekanntheit	-.424	<.001	.001	-.288	.007	.004

Tabelle 10.5: Standardisierte Gewichte, Signifikanzniveaus sowie durch ein Bootstrapping-Verfahren mit 10000 Stichproben abgesicherte Signifikanzniveaus der Regressionen der Änderungen des autonomieunterstützenden Motivierstils für Teilnehmende mit einer hoch (links) beziehungsweise niedrig (rechts) ausgeprägten autonomen Kausalitätsorientierung auf vermutete Prädiktoren; $\Delta\text{SMot_A}$ = Änderung des autonomieunterstützenden Motivierstils; Zeit = Zeit zwischen Intervention und der Posterhebung; Zustimmung = Zustimmung zu den Interventionsinhalten; Bekanntheit = Vorherige Bekanntheit der Interventionsinhalte.

Variable	$\Delta\text{SMot_K}$ (hohe Kau_K_Prä)			$\Delta\text{SMot_K}$ (niedrige Kau_K_Prä)		
	β	p	$p_{\text{Bootstrap}}$	β	p	$p_{\text{Bootstrap}}$
(Konstante)		.422	.369		.323	.331
Zeit	.163	.118	.031	.321	.001	.003
Zustimmung	-.302	.005	.004	-.419	<.001	.001
Bekanntheit	—	—	—	.229	.016	.013

Tabelle 10.6: Standardisierte Gewichte, Signifikanzniveaus sowie durch ein Bootstrapping-Verfahren mit 10000 Stichproben abgesicherte Signifikanzniveaus der Regressionen der Änderungen des kontrollierenden Motivierstils für Teilnehmende mit einer hoch (links) beziehungsweise niedrig (rechts) ausgeprägten kontrollierenden Kausalitätsorientierung auf vermutete Prädiktoren; $\Delta\text{SMot_K}$ = Änderung des kontrollierenden Motivierstils; Zeit = Zeit zwischen Intervention und der Posterhebung; Zustimmung = Zustimmung zu den Interventionsinhalten; Bekanntheit = Vorherige Bekanntheit der Interventionsinhalte.

Das Regressionsmodell der Teilgruppe mit einer hoch ausgeprägten kontrollierenden Kausalitätsorientierung wird mit $F(2, 83) = 5.33, p = .007 (R^2 = .11)$ insgesamt signifikant. Für die Gruppe mit einer niedrig ausgeprägten kontrollierenden Kausalitätsorientierung wird das Regressionsmodell mit $F(3, 76) = 13.74, p < .001 (R^2 = .35)$ insgesamt signifikant. In beiden Gruppen wird die Zustimmung zu den Interventionsinhalten signifikant. Somit gilt für beide Teilgruppen: je stärker die Probanden den Inhalten der Intervention zustimmten, desto größer war die Reduktion des kontrollierenden Motivierstils ($\beta = -.304, p = .005$ bzw. $\beta = -.419, p < .001$).

10.3 Ergebnisse der Forschungsfrage 3

In Kapitel 5.2 wurde ein Wirkmodell (Modell 1) zur Vorhersage der beobachteten Motivierstile herausgearbeitet. Demnach sollten das Fähigkeitsselbstkonzept bezüglich Autono-

mieunterstützung, die Kausalitätsorientierungen sowie die Arten der Tutormotivation den sogenannten Druck von Innen bilden. Entsprechend Kapitel 7.3 wird ebenfalls für fachbezogene Konstrukte (Fähigkeitsselbstkonzept bezüglich Physik, Verständnis der Variablen-Kontroll-Strategie, Fachwissen) kontrolliert. Das wahrgenommene Tuteengagement, die wahrgenommene Tuteemotivation sowie die wahrgenommene Selbstständigkeit der Tutees sollten den Druck von Unten bilden. Der Druck von Unten (spontan) sollte die Motivierstile beeinflussen. Der Druck von Innen sollte sowohl den Druck von Unten (spontan) als auch die Motivierstile beeinflussen. Ob der Druck von Innen jedoch nur die Motivierstile beeinflusst, wird mit dem Vergleich zu Modell 2 überprüft. Beide Modelle werden gegen das Modell 3 verglichen, in dem alle Konstrukte direkt auf die Motivierstile wirken und ansonsten miteinander korrelieren. Das Fachwissen muss aufgrund der mangelhaften internen Konsistenz von .598 ausgeschlossen werden. Die UN-Dimension des Verständnisses der Variablenkontrollstrategie kann aufgrund der hinreichenden internen Konsistenz von .711 berücksichtigt werden. Von den vier Arten der Tutormotivation hatten lediglich die intrinsische und identifizierte Motivation eine für die weitere Berücksichtigung hinreichende interne Konsistenz. Das wahrgenommene Engagement sowie die wahrgenommene Selbstständigkeit wurden jeweils mit nur zwei Items erfasst, weshalb keine interne Konsistenz bestimmt werden konnte. Deshalb wurde die Korrelation der beiden jeweiligen Items ersatzweise als Gütemaß zur Skalenbildung herangezogen. Die Items korrelieren jeweils grenzwertig bis hinreichend hoch miteinander ($\rho_{Eff} = .650, p < .001$ sowie $\rho_{AU} = .747, p < .001$). Für den beobachteten autonomieunterstützenden Motivierstil konnte keine Komponente extrahiert werden, weshalb in den Modellen nur der kontrollierende Motivierstil berücksichtigt wird. Folglich wird auch lediglich die kontrollierende und nicht mehr die autonomieunterstützende Kausalitätsorientierung im Modell berücksichtigt, da letztgenannte nur den autonomieunterstützende Motivierstil vorhersagen sollte Reeve et al. (2018).

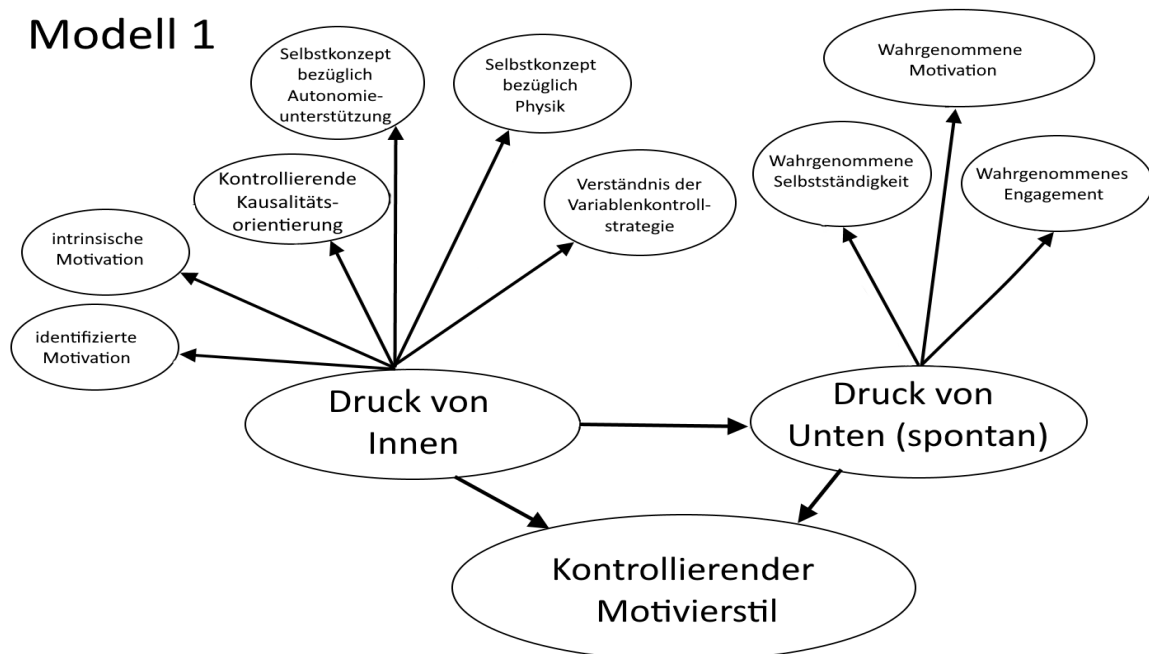


Abbildung 10.3: Modell 1; Erstes vermutetes Modell der Wirkbeziehung der Aspekte des Drucks von Innen und der Aspekte des Drucks von Unten (spontan) auf den beobachteten kontrollierenden Motivierstil.

Modell 2

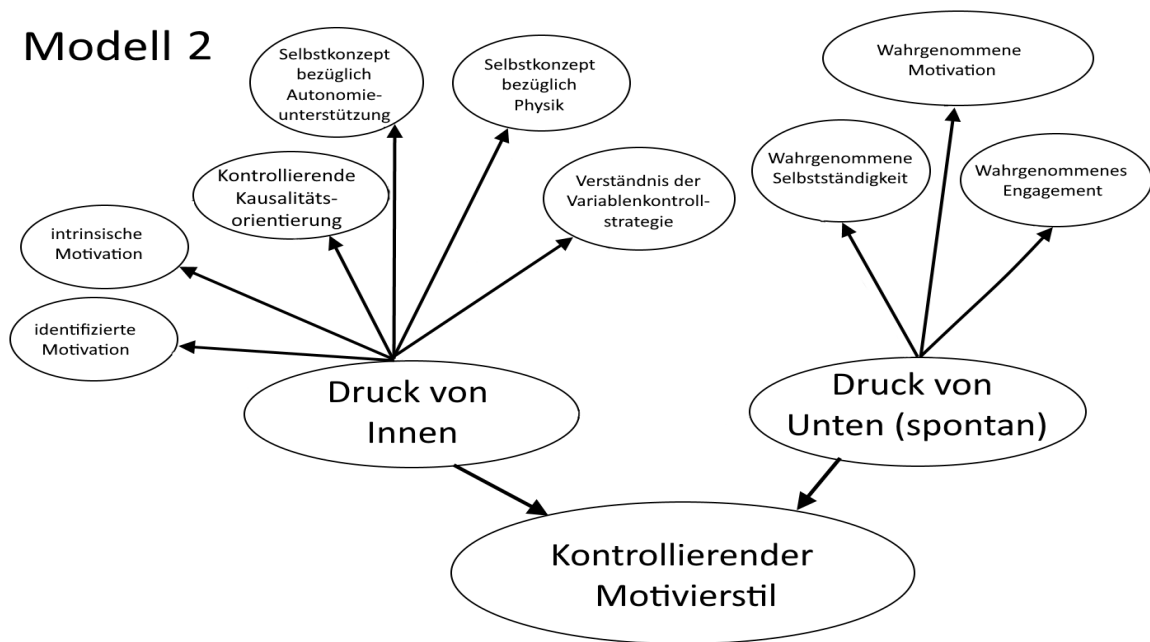


Abbildung 10.4: Modell 2; Zweites vermutetes Modell der Wirkbeziehung der Aspekte des Drucks von Innen und der Aspekte des Drucks von Unten (spontan) auf den beobachteten kontrollierenden Motivierstil.

Modell 3

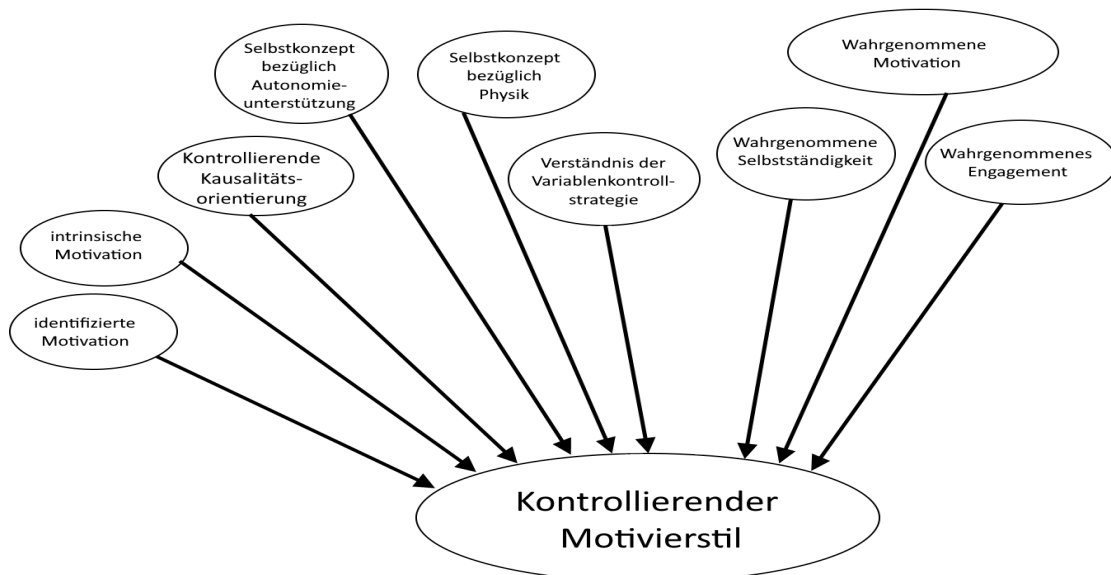


Abbildung 10.5: Modell 3; Drittes vermutetes Modell der Wirkbeziehung der Aspekte des Drucks von Innen und der Aspekte des Drucks von Unten (spontan) auf den beobachteten kontrollierenden Motivierstil; Alle Prädiktoren können miteinander korrelieren, was zum Erhalt der Übersichtlichkeit nicht dargestellt wurde.

Zur Analyse wurden lediglich die Daten der 116 Tutoren mit vollständigen Datensätzen genutzt (siehe Tabelle 9.12). Als Schätzer wurde der robuste MLM-Schätzer (maximum likelihood estimator with robust standard errors and a Satorra-Bentler scaled test statistic; Satorra und Bentler, 1994) genutzt, da die Daten nicht normalverteilt sind (siehe Kapitel 9.3). In den folgenden Modellen werden die betrachteten latenten Konstrukte aus den in Kapitel 9.2 untersuchten Items geschätzt. Crossladungen zwischen den Items und den latenten Konstrukten sind auf null fixiert. Die Kovarianzen zwischen Itemfehlern sind ebenfalls auf null fixiert. Da der MLM-Schätzer genutzt wurde, sind in Tabelle 10.7 sowohl die robusten als auch die skalierten Fitindizes aufgeführt (Satorra & Bentler, 1994).

Modell	df	N	χ^2	p(χ^2)	CFI	TLI	RMSEA	SRMR
					(rob./sca.)	(rob./sca.)	(rob./sca.)	
1	848	116	1213.11	<.001	.829/.825	.817/.813	.064/.061	.104
2	849	116	1218.27	<.001	.826/.823	.815/.811	.064/.061	.114
3	815	116	1139.96	<.001	.848/.844	.832/.827	.061/.059	.085

Tabelle 10.7: Fitwerte der zu vergleichenden Modelle 1 - 3

Die Passung der Strukturgleichungsmodelle auf die Daten wird anhand der Fitwerte CFI (comparative fit index), TLI (Tucker-Lewis index), dem RMSEA (root mean square error of approximation) und dem SRMR (Standardized Root Mean Square Residual), dem Verhältnis des χ^2 zur Anzahl der Freiheitsgrade sowie der Signifikanz des χ^2 -Test bewertet. Ein Modell beschreibt die Daten akzeptabel, wenn der CFI und der TLI über 0.9, der RMSEA und der SRMR unter 0.08 und das Verhältnis des χ^2 zur Anzahl der Freiheitsgrade unter 5 ist (Byrne, 2016; Kyriazos, 2018). Ein Modell beschreibt die Daten gut, wenn der CFI und der TLI über .95 sind, der RMSEA unter 0.05 ist, der SRMR unter .07 ist, das Verhältnis des χ^2 zur Anzahl der Freiheitsgrade unter 5 ist und der χ^2 -Test nicht mit $p < .05$ signifikant ist (Byrne, 2016; Kyriazos, 2018). Modelle werden ab einem Δ CFI von .002 als besser bezeichnet (Meade et al., 2008).

Die Modelle haben zwar keine tolerierbaren Fitwerte, sie sind jedoch nicht völlig ungenügend. Da Modell 3 deutlich die besten Fitwerte hat, wird die Ursachenanalyse der schlechten Fitwerte anhand von Modell 3 durchgeführt. Tabelle 10.8 listet die Fitwerte der Modelle auf, in denen jeweils eines der latenten Konstrukte und ihre manifesten Indikatoren entfernt sind.

Zwar hat keines der Modelle ideale Fitwerte, doch ist das Modell 3.1 ohne Berücksichtigung der Kausalitätsorientierung den anderen Modellen deutlich überlegen.

Modell 3 ohne	df	N	χ^2	p(χ^2)	CFI	TLI	RMSEA	SRMR
					(rob./sca.)	(rob./sca.)	(rob./sca.)	
3.1 Kau_K_Post	398	116	484.4	.002	.945/.942	.935/.932	.046/.043	.074
3.2 VKS	704	116	995.2	<.001	.854/.851	.839/.835	.063/.060	.085
3.3 SRQ_intr	741	116	2582.3	<.001	.847/.845	.830/.827	.064/.061	.085
3.4 SRQ_ident	704	116	989.9	<.001	.859/.853	.844/.838	.062/.059	.084
3.5 FSK_A	629	116	922.3	<.001	.843/.839	.824/.820	.066/.063	.084
3.6 FSK_P	629	116	908.7	<.001	.833/.828	.814/.808	.065/.062	.087
3.7 TW_En	743	116	1044.8	<.001	.850/.847	.835/.831	.062/.059	.085
3.8 TW_IM	704	116	991.7	<.001	.848/.845	.832/.828	.062/.059	.087
3.9 TW_Se	743	116	1027.6	<.001	.854/.850	.839/.834	.060/.057	.087

Tabelle 10.8: Fitwerte der zu vergleichenden Modelle 3.1 - 3.9; Kau_K_Post = (post-interventionelle) kontrollierende Kausalitätsorientierung; VKS = Verständnis der Variablen-Kontroll-Strategie; SRQ_intr = Intrinsische Motivation; SRQ_ident = Identifizierte Regulation der extrinsischen Motivation; FSK_A/P = Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung/Physik; TW_En/IM/Se = Von den Tutoren wahrgenommenes Engagement/wahrgenommene Motivation/wahrgenommene Selbstständigkeit der Tutees.

Modelle werden ab einem ΔCFI von .002 als besser bezeichnet (Meade et al., 2008). Das Modell ohne die kontrollierende Kausalitätsorientierung hat ein ΔCFI von mindestens .086 zu jedem der anderen Modelle. Deshalb wird die kontrollierende Kausalitätsorientierung im Folgenden nicht mehr berücksichtigt.

Da ein Item der identifizierten Regulation der extrinsischen Motivation in Modell 3.1 eine negative Varianz hat, wird das Modell 3.1.0 gegen das Modell 3.1.1 ohne das latente Konstrukt der identifizierten Regulation der extrinsischen Motivation verglichen.

Modell	df	N	χ^2	p(χ^2)	CFI (rob./sca.)	TLI (rob./sca.)	RMSEA (rob./sca.)	SRMR
3.1.0	398	116	484.4	.002	.945/.942	.935/.932	.046/.043	.074
3.1.1	322	116	387.6	.007	.956/.952	.948/.944	.044/.042	.068

Tabelle 10.9: Fitwerte der zu vergleichenden Modelle 3.1.0 und 3.1.1.

Das Modell ohne Berücksichtigung des latenten Konstrukts der identifizierten Motivation weist keine negative Itemvarianz mehr auf und hat zudem akzeptable Fitwerte. Die CFI, die RMSEA und der SRMR sind alle im guten Bereich. Lediglich die Signifikanz des χ^2 -Test weist noch auf Probleme des Fits hin.

Damit eines der ursprünglich unpassenden Gesamtmodelle die Daten besser beschreibt, mussten die kontrollierende Kausalitätsorientierung und die identifizierte Tutormotivation ausgeschlossen werden. Aufbauend auf diesen Kürzungen werden in Tabelle 10.10 und 10.11 erneut die drei Modelle verglichen, in denen die Konstrukte über den Druck von Innen und Unten auf den kontrollierenden Motivierstil wirken (Modell 5), der Druck von Innen zusätzlich den Druck von Unten beeinflusst (Modell 4) beziehungsweise das Modell, in dem alle Prädiktoren direkt den kontrollierenden Motivierstil vorhersagen (Modell 6, welches Modell 3.1.1 entspricht).

Modell	df	N	χ^2	p(χ^2)	CFI (rob./sca.)	TLI (rob./sca.)	RMSEA (rob./sca.)	SRMR
4	340	116	438.9	<.001	.933/.927	.926/.919	.053/.050	.095
5	341	116	447.0	<.001	.928/.922	.920/.914	.054/.052	.116
6	322	116	387.6	.007	.956/.952	.948/.944	.044/.042	.068

Tabelle 10.10: Fitwerte der zu vergleichenden Modelle 4 - 6.

Einerseits weist Modell 4 negative Itemvarianzen und Modell 5 negative Itemvarianzen und negative Varianzen eines latenten Konstrukts auf. Weiterhin hat Modell 6 die deutlich besten Fitwerte, was durch die ersten beiden Chi-Quadrat-Tests in Tabelle 10.11 bestätigt wird. Die Chi-Quadrat-Tests zeigen, dass das Modell ohne latente Zwischenkonstrukte die Daten signifikant besser abbildet. Dies bestätigt den Vergleich aus Tabelle 10.7.

Modell	df	AIC	BIC	χ^2	p
6 gegen	322	6327.3	6558.6	428.93	
4	340	6347.6	6529.4	485.28	<.001
5	341	6355.1	6534.1	494.75	<.001
4 gegen	340	6347.6	6529.4	485.28	
5	341	6355.1	6534.1	494.75	.012

Tabelle 10.11: Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).

Zusammengenommen wird Modell 6 zur Auswertung der Zusammenhänge der verschiedenen Aspekte des Drucks von Unten beziehungsweise Innen mit dem kontrollierenden Motivierstil genutzt. Entsprechend den Tabellen 10.7 und 10.11 scheinen diese Begriffe jedoch eher Kategorien von Konstrukten und weniger latente Konstrukte in den Daten zu sein. Abbildung 10.6 stellt eine Veranschaulichung des Modells 6 dar und Tabelle 10.12 listet die standardisierten Regressionsgewichte der latent geschätzten Konstrukte bezüglich des beobachteten kontrollierenden Motivierstils sowie die Signifikanzniveaus auf. Die Befunde werden in Kapitel 11.3 diskutiert.

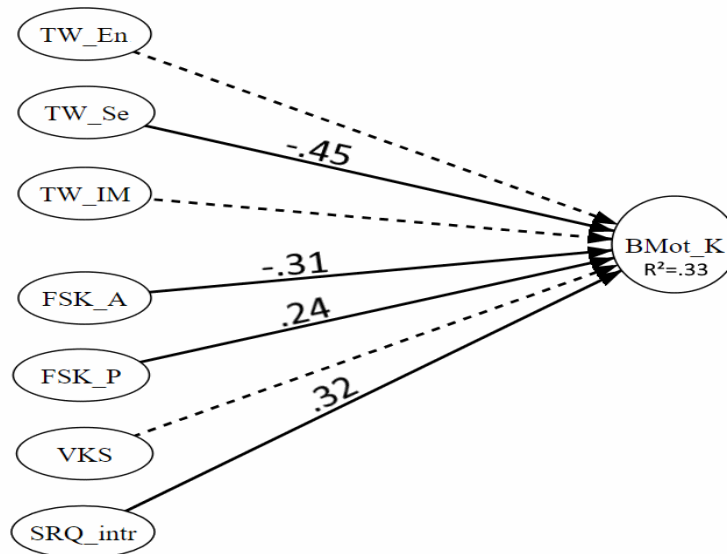


Abbildung 10.6: Strukturgleichungsmodell 6; Signifikante ($p < .05$) standardisierte Regressionsgewichte der Aspekte des Drucks von Innen und Unten (spontan) der Regression bezüglich des beobachteten kontrollierenden Motivierstils; Die Korrelationen der Prädiktoren untereinander wurden zum Erhalt der Übersichtlichkeit nicht dargestellt. Nicht signifikante ($p > .05$) Pfade sind gestrichelt dargestellt. BMot_K = beobachteter kontrollierender Motivierstil; TW_En/Se/IM = wahrgenommenes Engagement/ wahrgenommene Selbstständigkeit/ wahrgenommene Motivation; FSK_A/P = Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung/ Physik; VKS = Verständnis der Variablen-Kontroll-Strategie; SRQ_intr = Intrinsische Motivation bezüglich Tutorentätigkeit.

Latent geschätzte Konstrukte	BMot_K ($R^2 = .328$)	
	β	p
<i>Druck von Innen</i>		
Intrinsische Motivation bezüglich Tutorentätigkeit	.317	.028
Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung	-.308	.038
Fähigkeitsselbstkonzept bezüglich Physik	.238	.020
Verständnis der Variablen-Kontroll-Strategie	-.178	.138
<i>Druck von Unten</i>		
Wahrgenommenes Engagement	.154	.200
Wahrgenommene Motivation	-.236	.099
Wahrgenommene Selbstständigkeit	-.445	<.001

Tabelle 10.12: Regressionsgewichte der latent geschätzten Konstrukte bezüglich des beobachteten kontrollierenden Motivierstils. BMot_K = beobachteter kontrollierender Motivierstil

In Kapitel 9.2 wurde dargelegt, dass für das autonomieunterstützende Tutorverhalten keine Komponente extrahiert werden konnte. Um trotzdem ein Indiz für das autonomieunterstützende Tutorverhalten zu gewinnen, werden in Anlehnung an Reeve und Jang (2006) die Auswertungskategorien mit dem wahrgenommenen autonomieunterstützenden Motivierstil korreliert. Im Sinne einer diskriminanten Validität sollten die kontrollierenden Verhaltensweisen keine oder eine negative Korrelation zum wahrgenommenen autonomieunterstützenden Motivierstil aufweisen (siehe Kapitel 8.4.6 und 10.1). Da Tabelle 10.13 elf Korrelationen aufweist, wurde entsprechend der Bonferonni-Korrektur ein alphafehlerkorrigiertes Signifikanzniveau von $.05/11 = .0045$ angesetzt (Armstrong, 2014), sodass der testweite kumulierte Alphafehler $.05$ beträgt. Die übliche Signifikanzgrenze von $.001$ verschiebt sich somit zu $.001/11 = .0009$. Da einige der Auswertungskategorien nicht normalverteilt sind, listet Tabelle 10.13 Korrelationen nach Spearman auf.

	Vermutet autonomieunterstützend						Vermutet kontrollierend				
	TW1 _A	TW2 _A	E1	E2	E3	E4	TW1 _K	TW2 _K	S1	S2	S3
LCQ	-.06	.13	.02	-.02	.34***	-.03	.18	.12	-.00	.10	.14

Tabelle 10.13: Spearman-Korrelationen der auf die Videolänge und relative Ausprägung normierten vermutet autonomieunterstützenden Verhaltensweisen mit dem wahrgenommenen autonomieunterstützenden Motivierstil (LCQ); * $p < .0045$; *** $p < .00009$.

Eine Regression mit $F(6, 124) = 3.17, p < .01 (R^2 = .13)$ bestätigt, dass von den vermutet autonomieunterstützenden Auswertungskategorien lediglich E3_{soz} signifikant den wahrgenommenen autonomieunterstützenden Motivierstil vorhersagt. Dabei bestätigen sowohl das übliche als auch das durch ein Bootstrapping-Verfahren geschätzte Signifikanzniveau die ausschließliche Relevanz von E3_{soz}. Eine Regression unter zusätzlicher Berücksichtigung der kontrollierenden Verhaltensweisen bestätigt erneut mit $F(11, 119) = 1.87, p = .04 (R^2 = .15)$ die ausschließliche Relevanz von E3_{soz}.

Aufbauend auf den Modellvergleich in Tabelle 10.11, werden erneut drei Modelle verglichen. In den ersten beiden Vergleichsmodellen ergeben die Prädiktoren den Druck von Innen und Unten. In Modell 8 wirken die beiden Drücke nur direkt auf die Motivierstile, in Modell 7 wirkt der Druck von Innen auch auf den Druck von Unten. In Modell 9 wirken alle Prädiktoren direkt auf beide Motivierstile.

Modell	df	N	χ^2	$p(\chi^2)$	CFI (rob./sca.)	TLI (rob./sca.)	RMSEA (rob./sca.)	SRMR
7	365	116	464.0	<.001	.935/.930	.928/.922	.051/.048	.092
8	366	116	474.1	<.001	.929/.923	.921/.915	.053/.050	.113
9	342	116	410.9	.006	.955/.951	.946/.942	.044/.042	.067

Tabelle 10.14: Fitwerte der zu vergleichenden Modelle 7 - 9.

Erneut hat das Modell mit den direkten Wirkungen der Prädiktoren auf die Motivierstile die besten Fitwerte. Wie die Chi-Quadrat-Tests in Tabelle 10.15 bestätigen, beschreibt Modell 9 die Daten signifikant besser als die Modelle 7 und 8.

Modell	df	AIC	BIC	χ^2	p
9 gegen	342	6232.2	6488.2	451.98	
7	365	6243.7	6436.5	509.53	<.001
8	366	6253.2	6443.2	520.98	<.001
7 gegen	365	6243.7	6436.5	509.53	
8	366	6253.2	6443.2	520.98	.004

Tabelle 10.15: Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).

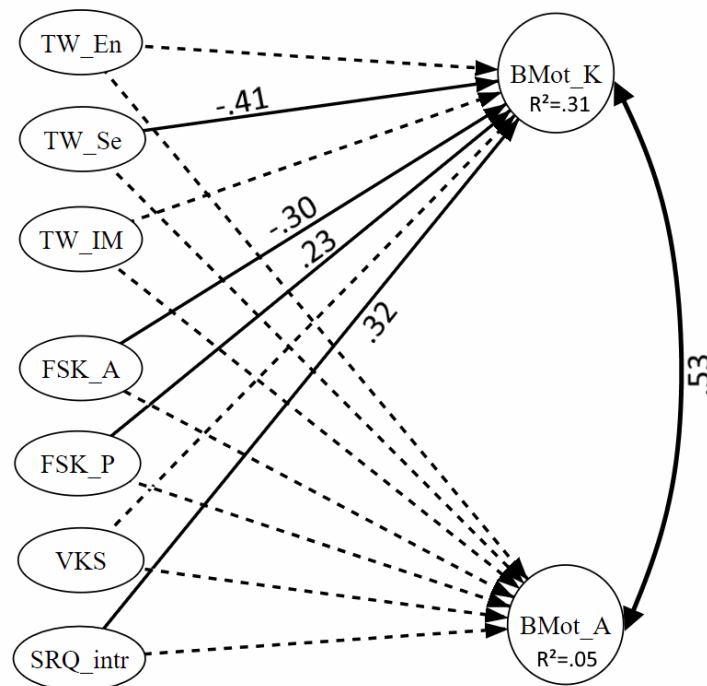


Abbildung 10.7: Strukturgleichungsmodell 9; Signifikante ($p < .05$) standardisierte Regressionsgewichte der Aspekte des Drucks von Innen und Unten (spontan) der Regressionen bezüglich des beobachteten kontrollierenden und autonomieunterstützenden Motivierstils; Die Korrelationen der Prädiktoren untereinander wurden zum Erhalt der Übersichtlichkeit nicht dargestellt. Nicht signifikante ($p > .05$) Pfade sind gestrichelt dargestellt. BMot_A/K = beobachteter autonomieunterstützender/ kontrollierender Motivierstil; TW_En/Se/IM = wahrgenommenes Engagement/ wahrgenommene Selbstständigkeit/ wahrgenommene Motivation; FSK_A/P = Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung/ Physik; VKS = Verständnis der Variablen-Kontroll-Strategie; SRQ_intr = Intrinsische Motivation bezüglich Tutorentätigkeit.

Der Vergleich der Modelle in Tabelle 10.15 bestätigt erneut, dass Modell 9 die Daten am besten beschreibt und somit, dass die Begriffe des Drucks von Unten und Innen eher Kategorien von Konstrukten und weniger als latente Konstrukte in den Daten zu verstehen sind. Abbildung 10.7 veranschaulicht Modell 9. Tabelle 10.16 listet die standardisierten Regressionsgewichte der latent geschätzten Konstrukte bezüglich der beiden Motivierstile sowie die Signifikanzniveaus auf. Die Befunde werden in Kapitel 11.3 diskutiert.

Latent geschätzte Konstrukte	BMot_A ($R^2 = .052$)		BMot_K ($R^2 = .305$)	
	β	p	β	p
<i>Druck von Innen</i>				
Intrinsische Motivation bezüglich Tutorentätigkeit	.112	.372	.322	.023
Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung	-.124	.397	-.299	.038
Fähigkeitsselbstkonzept bezüglich Physik	.172	.149	.227	.025
Verständnis der Variablen-Kontroll-Strategie	-.152	.241	-.168	.152
<i>Druck von Unten</i>				
Wahrgenommenes Engagement	.111	.225	.146	.205
Wahrgenommene Motivation	-.009	.945	-.248	.072
Wahrgenommene Selbstständigkeit	-.119	.331	-.412	.001

Tabelle 10.16: Regressionsgewichte der latent geschätzten Konstrukte bezüglich des beobachteten autonomieunterstützenden und kontrollierenden Motivierstils.

Von den ursprünglich 131 brauchbaren Videos wurden bisher nur 116 genutzt, da einzelne Werte in den manifesten Variablen fehlten. Acht der 15 entfernten Probanden hatten einzelne fehlende Werte bei einem der drei Items zur Erfassung des Verständnisses der Variablenkontrollstrategie. Werden die fehlenden Werte entsprechend den Mittelwerten des Items bezüglich der anderen Probanden ergänzt, liegen 131 vollständige Datensätze zur Auswertung vor. Die Fitwerte sind bei den 131 Probanden etwas besser als bei den 116 (siehe Anhang Kapitel 12.3.4). Die Signifikanzen der Prädiktoren ändern sich durch die Imputation nicht (siehe Anhang Kapitel 12.3.4).

Ergänzende Rechnungen zum Zusammenhang der beobachteten Motivierstile

Ein weiterer Befund des Modells 9 ist die hohe positive Korrelation zwischen dem beobachteten autonomieunterstützenden und kontrollierenden Motivierstil. Für die 116 ursprünglich vollständigen beziehungsweise 131 teilweise imputierten Datensätze betragen die Parameterschätzungen der Korrelationen $.529, p < .001$ und $.553, p < .001$. Dies widerspricht der Vermutung, dass die Motivierstile schwach negativ oder gar nicht miteinander zusammenhängen (siehe Kapitel 10.1 sowie Reeve et al., 2014). Ergänzende Rechnungen zur Untersuchung dieses vermeintlichen Widerspruchs werden ausführlich im Anhang (siehe Kapitel 12.3.5) dargestellt und zusammengefasst in Kapitel 11.1 inhaltlich diskutiert. Zusammenfassend ergibt die Rechnung, dass der beobachtete autonomieunterstützende und der kontrollierende Motivierstil (bezüglich der auf die Anzahl der Events normierten Verhaltensweisen) mit $-.314, p = .074$ erwartungskonform nicht mehr signifikant miteinander korrelieren.

10.4 Ergebnisse der Forschungsfrage 4

10.4.1 Zusammenhänge der Konstrukte auf der Tuteeebene

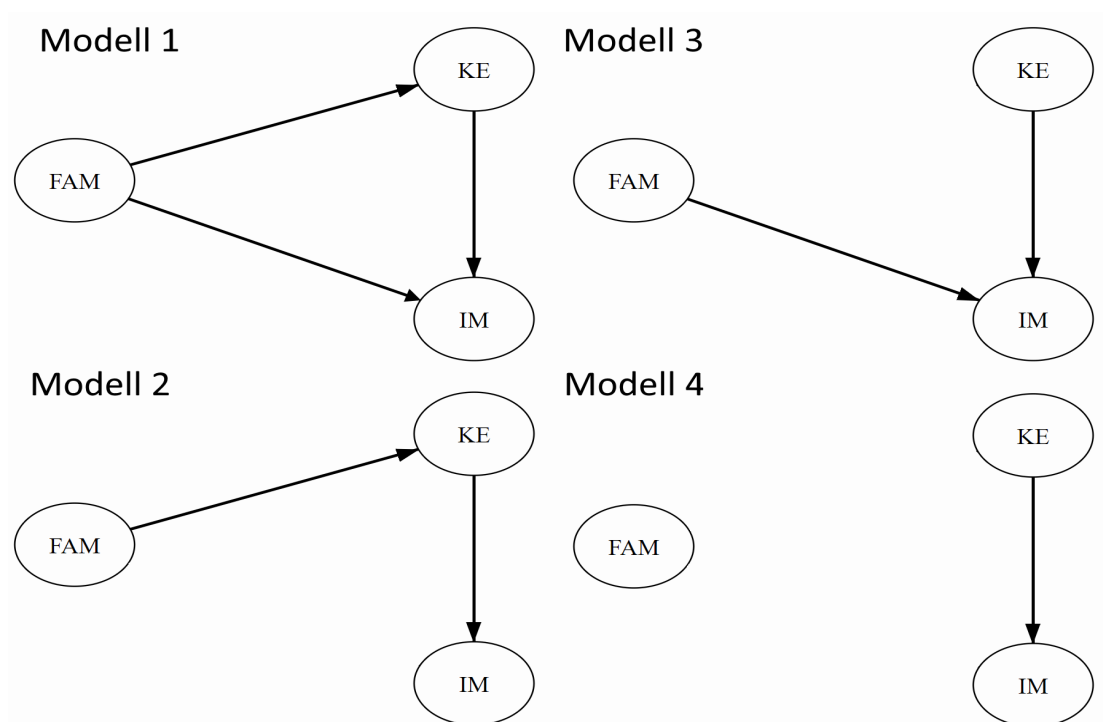


Abbildung 10.8: Die vier untersuchten Strukturgleichungsmodelle der latenten Konstrukte der intrinsischen Motivation (IM), des Kompetenzerlebens (KE) und der intrinsischen Motivation vor der Durchführung des Experiments (FAM).

Um die Struktur des wahrgenommenen autonomieunterstützenden Motivierstils, des Kompetenzerlebens und der intrinsischen Motivation zu untersuchen, werden verschiedene kovarianzbasierte Strukturgleichungsmodelle miteinander verglichen. Um die Motivation vor

der Durchführung des Experimentes angemessen berücksichtigen zu können, wird im ersten Schritt der Zusammenhang des Kompetenzerlebens, der intrinsischen Motivation und der Motivation vor der Durchführung des Experimentes untersucht. Entsprechend Kapitel 5.2 beeinflusst das Kompetenzerleben vermutlich die intrinsische Motivation. Die Motivation vor der Durchführung des Experimentes kann entweder sowohl das Kompetenzerleben als auch die intrinsische Motivation beeinflussen (Modell 1), jeweils nur eines der beiden beeinflussen (Modell 2 und 3) oder mit keinem der beiden zusammenhängen (Modell 4, siehe auch Abbildung 10.8).

Zur Analyse wurden lediglich die Daten der 133 Tutees mit vollständigen Datensätzen genutzt. Als Schätzer wurde der robuste MLM-Schätzer (maximum likelihood estimator with robust standard errors and a Satorra-Bentler scaled test statistic; Satorra und Bentler, 1994) genutzt, da die Daten nicht normalverteilt sind (siehe Kapitel 9.3). Entsprechend Kapitel 9.2 wird die intrinsische Motivation durch sechs, das Kompetenzerleben durch fünf, der wahrgenommene autonomieunterstützende Motivierstil durch sechs und die intrinsische Motivation vor Durchführung des Experimentes durch fünf Items repräsentiert. In den Modellen ergeben diese Items jeweils eine latente Variable. Crossladungen zwischen den Items und den anderen latenten Konstrukten sind auf null fixiert. Die Korrelationen von Messfehlern zwischen Items verschiedener latenter Konstrukte sind ebenfalls auf null fixiert. Werden zusätzlich die Korrelationen von Messfehlern von Items gleicher latenter Konstrukte auf null fixiert, weisen alle Modelle unzureichende Fitwerte auf (siehe Tabelle 10.17).

Modell	df	N	χ^2	p(χ^2)	CFI (rob./sca.)	TLI (rob./sca.)	RMSEA (rob./sca.)	SRMR
Modelle mit auf null fixierten korrelierten Messfehlern								
1	101	133	153.97	.001	.924/.864	.910/.839	.074/.063	.067
2	102	133	162.94	<.001	.912/.844	.896/.817	.079/.067	.088
3	102	133	175.50	<.001	.896/.812	.877/.779	.086/.074	.160
4	103	133	184.03	<.001	.880/.793	.861/.758	.092/.077	.205
Modelle mit frei geschätzten korrelierten Messfehlern								
1	94	133	101.58	.279	.990/.981	.987/.975	.028/.025	.054
2	95	133	111.19	.123	.978/.959	.972/.948	.041/.036	.074
3	95	133	125.27	.020	.959/.923	.948/.902	.056/.049	.155
4	96	133	135.58	.005	.944/.899	.930/.873	.065/.056	.200

Tabelle 10.17: Fitwerte der Modelle mit und ohne vollständig auf null fixierte korrelierte Messfehler.

Da der MLM-Schätzer genutzt wurde, sind in Tabelle 10.17 sowohl die robusten als auch die skalierten Fitindizes aufgeführt (Satorra & Bentler, 1994). Durch Zulassen der frei geschätzten korrelierten Messfehler von Items gleicher latenter Konstrukte haben die ersten beiden Modelle gute Fitwerte. Modell 3 hat teils akzeptable und teils problematische Fitwerte. Modell 4 hat mangelhafte Fitwerte (Byrne, 2016; Kyriazos, 2018). Inwiefern die Berücksichtigung dieser korrelierten Messfehler die Ergebnisse und Interpretierbarkeit beeinflusst, wird in Kapitel 11.5 diskutiert. Entsprechend Rheinberg et al. (2001) kann die Vermutung aufgestellt werden, dass die Motivation vor der Durchführung des Experiments sowohl das Kompetenzerleben als auch die intrinsische Motivation beeinflusst. Folglich sollte Modell 1 am besten auf die Daten passen. Tabelle 10.18 enthält den Vergleich der drei übrigen Modelle gegen das erste Modell.

Modell	df	AIC	BIC	χ^2	p
1 gegen	94	5938.9	6060.3	133.57	
2	95	5951.6	6070.1	148.18	.016
3	95	5966.6	6085.2	163.27	<.001
4	96	5985.7	6101.3	184.35	<.001

Tabelle 10.18: Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).

Entsprechend den Modellvergleichen hat das erste Modell erwartungskonform mit $\chi^2(94, N = 133) = 101.58, p = .279, CFI = .990/.981, TLI = .987/.975, RMSEA = .028/.025, SRMR = .054$ einen sehr guten Modellfit (Byrne, 2016; Kyriazos, 2018). Insbesondere fittet das erste Modell die Daten signifikant besser als die übrigen drei Modelle. Abbildung 10.9 stellt Modell 1 dar.

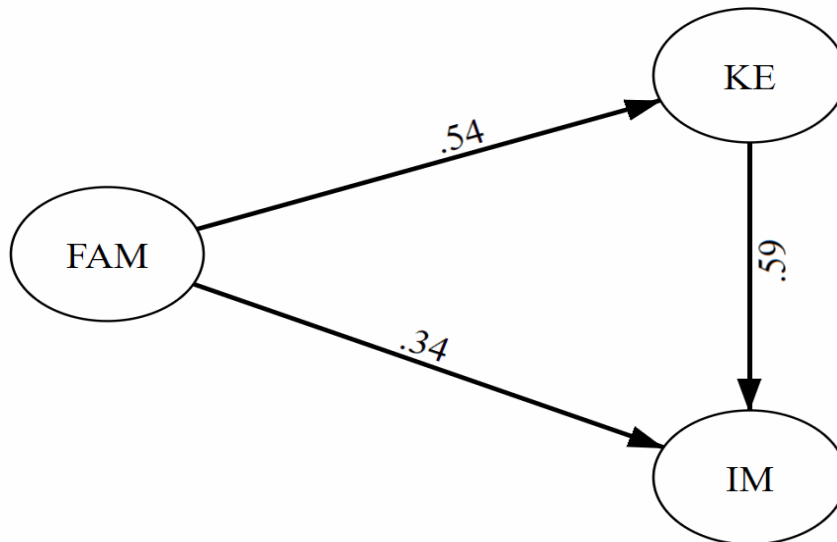


Abbildung 10.9: Strukturgleichungsmodell 1 der latenten Konstrukte der intrinsischen Motivation (IM), des Kompetenzerlebens (KE) und der intrinsischen Motivation vor der Durchführung des Experiments (FAM).

Da die Struktur zwischen der Motivation vor der Durchführung des Experiments, dem Kompetenzerleben und der intrinsischen Motivation erwartungskonform bestätigt werden konnte, kann darauf aufbauend die Struktur nach Hinzunahme des wahrgenommenen autonomieunterstützenden Motivierstils untersucht werden. Abbildung 10.10 stellt die vier verglichenen Modelle dar.

Das erste Modell stellt ein partiell mediiertes Modell dar, in dem der wahrgenommene autonomieunterstützende Motivierstil einerseits direkt und andererseits mediiert über das Kompetenzerleben auf die intrinsische Motivation wirkt. Modell 2 stellt eine vollständige Mediation des wahrgenommenen autonomieunterstützenden Motivierstils über das Kompetenzerleben auf die intrinsische Motivation dar. Entsprechend dem dritten Modell beeinflussen zwar sowohl das Kompetenzerleben als auch der wahrgenommene autonomieunterstützende Motivierstil die intrinsische Motivation, jedoch hängen die beiden nicht

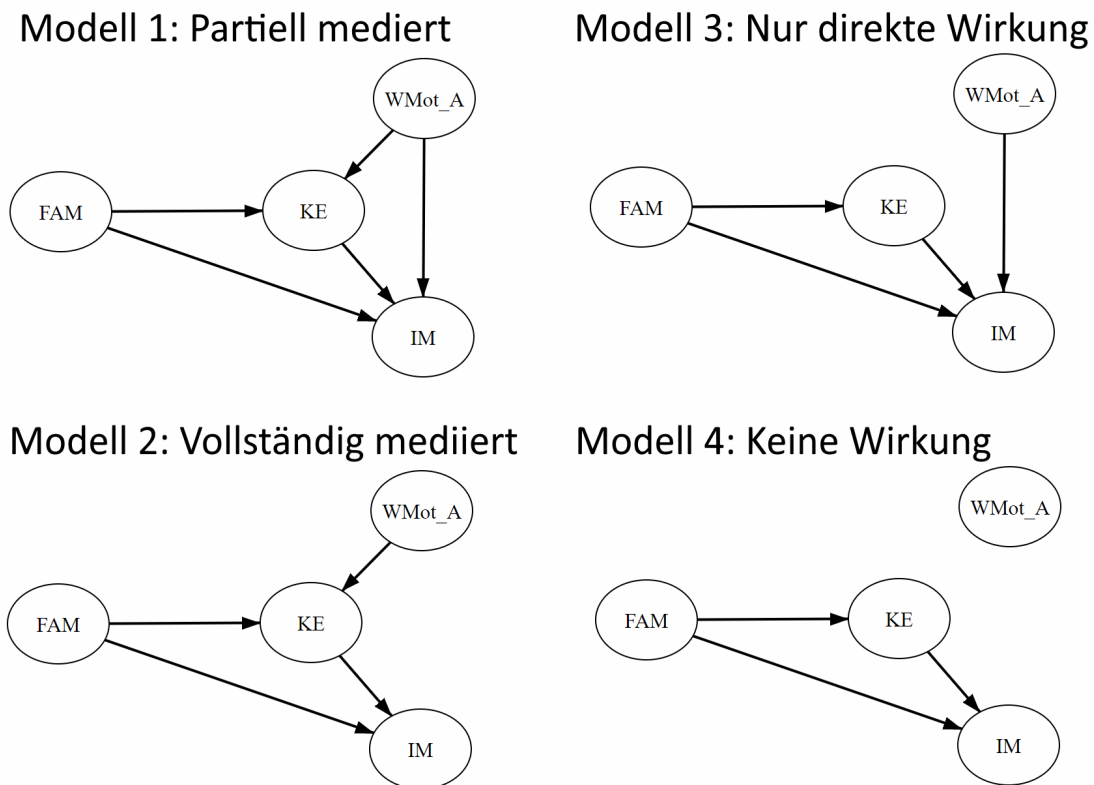


Abbildung 10.10: Die vier untersuchten Strukturgleichungsmodelle der latenten Konstrukte der intrinsischen Motivation (IM), des Kompetenzerlebens (KE), der intrinsischen Motivation vor der Durchführung des Experimentes (FAM) und des wahrgenommenen autonomieunterstützenden Motivierstils (WMot_A).

miteinander zusammen. Das Modell stellt also einen ausschließlich direkten und nicht mediierten Effekt des wahrgenommenen autonomieunterstützenden Motivierstils auf die intrinsische Motivation dar. Möglicherweise ist der wahrgenommene autonomieunterstützende Motivierstil völlig irrelevant für das Kompetenzerleben und die intrinsische Motivation. Sollte dies der Fall sein, müsste das vierte Modell die Daten am besten beschreiben. Das zweite, dritte und vierte Modell sind im ersten Modell genestet. Das vierte Modell ist im dritten Modell genestet. Deshalb zeigen ein Vergleich des zweiten und ersten Modells sowie der Vergleich des vierten und dritten Modells, ob der direkte Pfad vom latenten Konstrukt des wahrgenommenen autonomieunterstützenden Motivierstils auf das latente Konstrukt der intrinsischen Motivation signifikant mehr Varianz bei der Vorhersage der intrinsischen Motivation aufklärt. Weiterhin ist das vierte Modell im zweiten Modell genestet. Ein Vergleich des zweiten und vierten Modells sowie der Vergleich des ersten und dritten Modells zeigen, ob der Pfad vom latenten Konstrukt des wahrgenommenen autonomieunterstützenden Motivierstils auf das latente Konstrukt des Kompetenzerlebens und somit des mediierten Einflusses auf das latente Konstrukt der intrinsischen Motivation signifikant mehr Varianz bei der Vorhersage des Kompetenzerlebens und der intrinsischen Motivation aufklärt. Modelle, in denen die Motivation vor der Durchführung des Experimentes mit dem wahrgenommenen autonomieunterstützenden Motivierstil zusammenhängt, wurden nicht untersucht, da sie inhaltlich nicht gerechtfertigt werden können. Entsprechend Hypothese H4a sollte das zweite Modell die Daten signifikant besser beschreiben als die

übrigen drei Modelle. Liegt eine partielle Mediation vor, fittet das erste Modell am besten. Da der MLM-Schätzer genutzt wurde, sind in Tabelle 10.19 sowohl die robusten als auch die skalierten Fitindizes der vier Modelle aufgeführt (Satorra & Bentler, 1994).

Modell	df	N	χ^2	p(χ^2)	CFI	TLI	RMSEA	SRMR
					(rob./sca.)	(rob./sca.)	(rob./sca.)	
1	196	133	224.85	.077	.968/.952	.962/.944	.037/.033	.078
2	197	133	232.26	.043	.961/.941	.954/.931	.041/.037	.089
3	197	133	232.06	.044	.961/.942	.955/.932	.041/.037	.099
4	198	133	239.56	.023	.954/.931	.946/.920	.045/.040	.116

Tabelle 10.19: Fitwerte der Modelle 1 - 4.

Modell 1 hat als einziges durchgehend gute Fitwerte. Alle weiteren Modelle haben problematisch hohe SRMR und einen signifikanten χ^2 -Test. Die Fitwerte von Modell 2 und 3 können noch als grenzwertig betrachtet werden, wobei die Fitwerte von Modell 4 insgesamt als problematisch zu bewerten sind (Byrne, 2016; Kyriazos, 2018).

Relevanz des direkten Einflusses des latenten Konstruktes des wahrgenommenen autonomieunterstützenden Motivierstils auf das latente Konstrukt der intrinsischen Motivation

Tabelle 10.20 enthält die Vergleiche des ersten und zweiten Modells sowie des dritten und vierten Modells.

Modell	df	AIC	BIC	χ^2	p
1 gegen	196	8769.9	8934.7	282.5	
2	197	8778.4	8940.3	293.0	.006
3 gegen	197	8775.2	8937.0	289.8	
4	198	8784.3	8943.3	300.9	.003

Tabelle 10.20: Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).

Beide Modellvergleiche zeigen, dass der direkte Pfad vom latenten Konstrukt des wahrgenommenen autonomieunterstützenden Motivierstils auf das latente Konstrukt der intrinsischen Motivation signifikant mehr Varianz bei der Vorhersage der intrinsischen Motivation aufklärt.

Relevanz des medierten Einflusses des latenten Konstruktes des wahrgenommenen autonomieunterstützenden Motivierstils auf das latente Konstrukt der intrinsischen Motivation

Tabelle 10.21 enthält die Vergleiche des ersten und dritten Modells sowie des zweiten und vierten Modells.

Modell	df	AIC	BIC	χ^2	p
1 gegen	196	8769.9	8934.7	282.5	
3	197	8775.2	8937.0	289.8	.003
2 gegen	197	8778.4	8940.3	293.0	
4	198	8784.3	8943.3	300.9	.002

Tabelle 10.21: Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).

Beide Modellvergleiche zeigen, dass der Pfad vom latenten Konstrukt des wahrgenommenen autonomieunterstützenden Motivierstils auf das latente Konstrukt des Kompetenzerlebens und somit der medierte Einfluss auf das latente Konstrukt der intrinsischen Motivation signifikant mehr Varianz bei der Vorhersage des Kompetenzerlebens und der intrinsischen Motivation aufklärt. Scheinbar ist das partiell medierte Modell das Beste der vier Modelle. Dies wird durch die folgenden drei paarweisen Modellvergleiche in Tabelle 10.22 bestätigt. Abbildung 10.11 stellt Modell 1 dar.

Modell	df	AIC	BIC	χ^2	p
1 gegen	196	8769.9	8934.7	282.5	
2	197	8778.4	8940.3	293.0	.006
3	197	8775.2	8937.0	289.8	.003
4	198	8784.3	8943.3	300.9	<.001

Tabelle 10.22: Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).

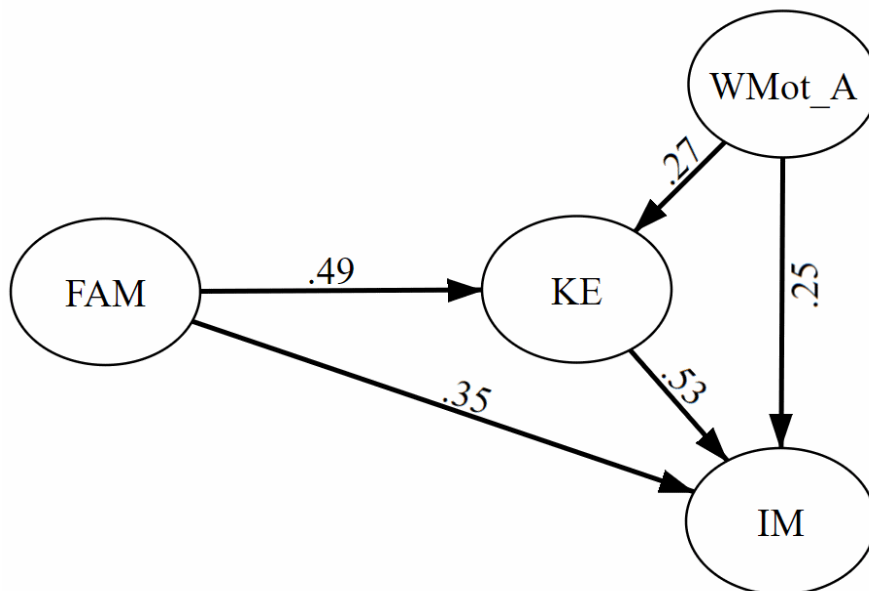


Abbildung 10.11: Strukturgleichungsmodell 1 der latenten Konstrukte der intrinsischen Motivation (IM), des Kompetenzerlebens (KE) und der intrinsischen Motivation vor der Durchführung des Experiments (FAM) und dem wahrgenommenen autonomieunterstützenden Motivierstil (WMot_A).

Neben den dargestellten Mediationsanalysen mittels Strukturgleichungsmodellen können dem Anhang (siehe Kapitel 12.3.6) Mediationsanalysen nach Baron und Kenny (1986), nach Hayes (2018) mit Hilfe seines SPSS-Makros *PROCESS* sowie nach Tingley et al. (2014) mit Hilfe des zugehörigen R-Pakets *mediation* entnommen werden. Sowohl aus den Strukturgleichungsmodellen als auch den Mediationsanalysen nach Baron und Kenny (1986), Hayes (2018) und Tingley et al. (2014) kann geschlussfolgert werden, dass der wahrgenommene autonomieunterstützende Motivierstil über das Kompetenzerleben partiell mediert auf die intrinsische Motivation wirkt.

10.4.2 Relevanz der Geschlechterkomposition

Als Voruntersuchung wurde überprüft, dass sich weibliche und männliche Tutees weder in ihrem angegebenen Kompetenzerleben, ihrer intrinsischen Motivation noch in ihrer anfänglichen Motivation unterscheiden. Weiterhin geben sie auch keinen signifikant unterschiedlichen wahrgenommenen autonomieunterstützenden Motivierstil an. Um weiter ausschließen zu können, dass sich weibliche oder männliche Tutoren aus Sicht der Tutees grundsätzlich autonomieunterstützender verhalten, das Kompetenzerleben mehr befriedigen oder stärker motivieren, wurde ein entsprechender Gruppenvergleich gerechnet. Weder die T-Tests noch die Mann-Whitney-U-Tests zeigen Unterschiede.

Entsprechend Hypothese H4b wird untersucht, ob die Tutees mit gleichgeschlechtlichem Tutor einen signifikant höheren wahrgenommenen autonomieunterstützenden Motivierstil, eine höhere Befriedigung des Kompetenzerlebens sowie eine höhere intrinsische Motivation im Vergleich zu Tutees mit einem Tutor des anderen Geschlechts angeben. Dazu wurde eine Gruppenvariable berechnet, welche kodiert, ob die Tutoren und Tutees das gleiche oder ein unterschiedliches Geschlecht haben. Tabelle 10.23 listet die Geschlechterverteilung in den Gruppen auf.

	Gleiches Geschlecht	Unterschiedliches Geschlecht
Tutor weiblich	47	22
Tutor männlich	27	39
Summe Tutor-Tutee-Paare	74	61

Tabelle 10.23: Geschlechterverteilung der Tutoren und Tutees

Das Kompetenzerleben, die intrinsische Motivation vor der Durchführung des Experiments und der wahrgenommene autonomieunterstützende Motivierstil sind in den Einzelgruppen näherungsweise normalverteilt. Deshalb können für diese drei Variablen T-Tests berechnet werden, wobei die Ergebnisse durch Bootstrapping-Verfahren mit 10000 Stichproben abgesichert werden. Für die intrinsische Motivation wird lediglich ein Mann-Whitney-U-Test berichtet. Um die Vergleichbarkeit der Gruppen miteinander sicherzustellen, werden etwaige Gruppenunterschiede bezüglich der Motivation vor der Durchführung des Experiments herangezogen.

	M ₁ (SD)	M ₂ (SD)	Levene-Test	d _{Cohen}	PT-Test/ P _{Bootstrapping}	Z _{Wilcoxon}	r	PMWU
Überprüfen der Gruppenvergleichbarkeit								
FAM	.716(.189)	.727(.146)	.226	-.067	.732/.725	-0.040	-.003	.968
Gruppenvergleich								
LCQ	.684(.171)	.740(.165)	.899	-.338	.053/.054	-1.982	-.171	.048
PC	.784(.178)	.791(.160)	.410	-.047	.788/.794	-0.055	-.005	.956
IE	.861(.190)	.885(.133)	—	—	—	-0.311	-.027	.756

Tabelle 10.24: T-Tests und Mann-Whitney-U-Tests für gleich- vs. getrenntgeschlechtliche Gruppen; FAM = Intrinsische Motivation vor der Durchführung des Experiments; LCQ = Wahrgenommener autonomieunterstützender Motivierstil; PC = Kompetenzerleben; IE = Intrinsische Motivation.

Da sich die Motivation vor der Durchführung des Experiments in den Gruppen nicht signifikant unterscheidet, werden die Gruppen vor der Tutoring-Sitzung als vergleichbar angesehen (siehe Tabelle 10.24). Die Gruppenvergleiche zeigen, dass sich die Gruppen hinsichtlich ihres Kompetenzerlebens und ihrer intrinsischen Motivation im Anschluss an die

Tutoring-Sitzung nicht unterscheiden. Bezüglich des wahrgenommenen autonomieunterstützenden Motivierstils wird der T-Test mit einer Effektstärke von $|d_{Cohen}|=0.34$ und einem Signifikanzniveau von .053 knapp nicht signifikant, wobei das Ergebnis durch ein Bootstrapping-Verfahren mit 10000 Stichproben und einem resultierenden Signifikanzniveau von $p = .054$ abgesichert ist. Der zugehörige Mann-Whitney-U-Test wird mit einem Korrelationskoeffizient von $|r| = .17$ und einem Signifikanzniveau von .048 knapp signifikant. Die Effektstärken sind im kleinen Bereich (J. Cohen, 1988).

11 Diskussion der Ergebnisse

Tabelle 11.1 und 11.2 fassen als vorausgehenden Überblick zusammen, welche Hypothesen bestätigt oder abgelehnt wurden. Den nachfolgenden Kapiteln 11.1 bis 11.4 können die zugehörigen detaillierten Diskussionen der Befunde aus Kapitel 10 entnommen werden.

Kürzel	Hypothese	Verweis	Zeichen
Zusammenhang der Perspektiven der Motivierstile			
H1a	Der selbstberichtete autonomieunterstützende Motivierstil der Tutoren, der beobachtete autonomieunterstützende Motivierstil und der von den Tutoren wahrgenommene autonomieunterstützende Motivierstil hängen positiv in Höhe von $r = .4$ miteinander zusammen.	S. 83 - 84	✓
H1b	Der selbstberichtete kontrollierende Motivierstil der Tutoren und der beobachtete kontrollierende Motivierstil hängen positiv in Höhe von $r = .4$ miteinander zusammen.	S. 83 - 84	✓
Prä-Post-Mittelwertsunterschiede			
H2a ₁	Eine Intervention kann den selbstberichteten autonomieunterstützenden Motivierstil von Tutoren mit einer Effektstärke von etwa $d = 0.8$ fördern.	S. 84 - 85	✓
H2b ₁	Eine Intervention kann den selbstberichteten kontrollierenden Motivierstil von Tutoren mit einer Effektstärke von etwa $d = 0.8$ reduzieren.	S. 84 - 85	✓
Prädiktoren der Prä-Post-Mittelwertsunterschiede			
H2a _{2,I}	Je stärker die autonome Kausalitätsorientierung der Tutoren ausgeprägt ist, desto mehr steigt der selbstberichtete autonomieunterstützende Motivierstil durch eine entsprechende Intervention.	S. 85 - 89	✗
H2b _{2,I}	Je stärker die kontrollierende Kausalitätsorientierung der Tutoren ausgeprägt ist, desto weniger wird der selbstberichtete kontrollierende Motivierstil durch eine entsprechende Intervention reduziert.	S. 85 - 89	✗
H2a _{2,II} / H2b _{2,II}	Je stärker die Tutoren den Interventionsinhalten zustimmen, desto mehr steigt/ sinkt der selbstberichtete autonomieunterstützende/kontrollierende Motivierstil ($\beta = .3$) durch eine entsprechende Intervention.	S. 85 - 89	✓
H2a _{2,III} / H2b _{2,III}	Je mehr die Tutoren die Interventionsinhalte bereits vor der Intervention kannten, desto weniger steigt/sinkt ihr präinterventioneller selbstberichteter autonomieunterstützender/ kontrollierender Motivierstil durch eine entsprechende Intervention.	S. 85 - 89	✓/(✓)

Tabelle 11.1: Zusammenfassung über die Annahme und die Ablehnung der aufgestellten Hypothesen - Teil 1.

11 Diskussion der Ergebnisse

Kürzel	Hypothese	Verweis	Zeichen
H2a _{2,IV} / H2b _{2,IV}	Der Zuwachs/ die Reduktion des autonomieunterstützenden/ kontrollierenden Motivierstils ist unabhängig von der vergange- nen Zeit zwischen der Intervention und der Posterhebung, also dem zweiten Messzeitpunkt.	S. 85 - 89	✓/✗
H2a _{2,V} / H2b _{2,V}	Je besser die Deutschnote ist, desto größer sind die Interventi- onseffekte.	S. 85 - 89	✗
H2a _{2,VI} / H2b _{2,VI}	Weibliche Tutoren haben aufgrund ihrer präinterventionellen Motivierstile einen geringeren Interventionserfolg.	S. 85 - 89	✗
Differentielle Effekte			
H2a _{3,I}	Eine Intervention erhöht den selbstberichteten autonomieun- terstützenden Motivierstil von Teilnehmenden mit einer rela- tiv stark ausgeprägten autonomen Kausalitätsorientierung un- abhängig von ihrer Zustimmung zu den Interventionsinhalten.	S. 89 - 90	✗
H2a _{3,II}	Eine Intervention reduziert den selbstberichteten kontrollieren- den Motivierstil von Teilnehmenden mit einer relativ schwach ausgeprägten kontrollierenden Kausalitätsorientierung unabhän- gig von ihrer Zustimmung zu den Interventionsinhalten.	S. 89 - 90	✗
H2b _{3,I}	Eine Intervention erhöht den selbstberichteten autonomieun- terstützenden Motivierstil von Teilnehmenden mit einer rela- tiv schwach ausgeprägten autonomen Kausalitätsorientierung in starker Abhängigkeit ($\beta > .6$) von ihrer Zustimmung zu den Interventionsinhalten.	S. 89 - 90	✗
H2b _{3,II}	Eine Intervention reduziert den selbstberichteten kontrollieren- den Motivierstil von Teilnehmenden mit einer relativ stark aus- geprägten kontrollierenden Kausalitätsorientierung in starker Abhängigkeit ($\beta > .6$) von ihrer Zustimmung zu den Interventi- onsinhalten.	S. 89 - 90	✗
Prädiktoren des gezeigten Tutorverhaltens			
H3a	Die Aspekte des Drucks von Innen und die Aspekte des Drucks von Unten (spontan) sagen den beobachteten autonomieunter- stützenden Motivierstil vorher.	S. 90 - 99	(✗)
H3b	Die Aspekte des Drucks von Innen und die Aspekte des Drucks von Unten (spontan) sagen den beobachteten kontrollierenden Motivierstil vorher.	S. 90 - 99	✓
H3c	Der Druck von Innen sagt die Motivierstile partiell mediiert über den Druck von Unten (spontan) vorher.	S. 90 - 99	✗
Tuteeebene			
H4a	Unter Kontrolle der intrinsischen Motivation vor der Durchfüh- rung des Experiments, sagt der von den Tutees wahrgenommene autonomieunterstützende Motivierstil der Tutoren ihre intrinsi- sche Motivation vorher. Dieser Zusammenhang ist nicht direkt, sondern wird vollständig über die Befriedigung des Kompeten- zerlebens der Tutees mediiert.	S. 99 - 105	(✗)
H4b	Bei gleichgeschlechtlichem Tutor geben Tutees einen signifikant ($p < .05$) höheren wahrgenommenen autonomieunterstützenden Motivierstil, eine höhere intrinsische Motivation und eine höhere Befriedigung des Kompetenzerlebens an im Vergleich zu Tutees mit einem Tutor eines anderen Geschlechts.	S. 105 - 106	(✗)

Tabelle 11.2: Zusammenfassung über die Annahme und die Ablehnung der aufgestellten Hypothesen - Teil 2.

11.1 Diskussion der Struktur der Motivierstile (Forschungsfrage 1)

Entsprechend den Hypothesen H_{FF1a} und H_{FF1b} wurde vermutet, dass die verschiedenen Messperspektiven der Motivierstile ungefähr in Höhe von $r = .4$ miteinander zusammenhängen. Dies konnte für den selbstberichteten und beobachteten kontrollierenden Motivierstil bestätigt werden. Weiterhin trifft dies auch näherungsweise für den beobachteten und von den Tutees wahrgenommenen autonomieunterstützenden Motivierstil zu. Wie bei Aelterman et al. (2014) konnten die übrigen Zusammenhänge der Messperspektiven nicht gefunden werden. Somit treffen die Hypothesen H_{FF1a} und H_{FF1b} eingeschränkt zu. Die fehlenden Zusammenhänge legen eine Mediation über oder eine Moderation von nicht berücksichtigten Variablen nahe. Beispielsweise korrelieren einzelne vermutete autonomieunterstützende Verhaltensweisen der Tutoren nicht mit einer entsprechenden Wahrnehmung durch die Tutees. Die Kausalitätsorientierungen der Tutees könnten Moderatoren dieser nicht gefundenen Zusammenhänge sein (Baard et al., 2004; Deci & Ryan, 1985). Die in Kapitel 6.1 aufgestellte Vermutung, dass die Messperspektiven der Motivierstile unterschiedliche Konstrukte sind, könnte eine weitere Erklärung der fehlenden Korrelationen sein. Die Vermutung wird durch die Daten der vorliegenden Studie gestützt, da der beobachtete kontrollierende Motivierstil nicht nur durch den selbstberichteten kontrollierenden Motivierstil vorhergesagt wird, sondern die wahrgenommene Selbstständigkeit der Tutees 6.6% zusätzliche Varianz aufklären (siehe Anhang Kapitel 12.4.1).

Die Untersuchung einer vermeintlich erwartungswidrigen Korrelation des Strukturgleichungsmodells (siehe Abbildung 10.7) ergab, dass Tutoren mit einem ausgeprägten kontrollierenden Motivierstil so viel mehr geredet haben, dass sie absolut häufiger auf die Tutees eingegangen sind als die Tutoren, die sehr selten etwas gesagt haben, dann aber relativ häufiger auf ihre Tutees eingegangen sind (siehe Kapitel 10.3 und Anhang Kapitel 12.3.5). Dieses Missverständnis der Tutoren bezüglich der Intervention könnte einen weiteren Grund für die oben genannte fehlende Korrelation der Kategorie E3 zum selbstberichteten autonomieunterstützenden Motivierstil darstellen. Entsprechend den Daten sollten die Tutoren grundsätzlich darauf trainiert werden, mehr mit den Tutees zu reden. Dabei sollten sie allerdings keine inhaltlichen Hinweise geben oder ins Experiment eingreifen. Vielmehr sollten die Tutoren ermutigend auf die Aussagen, Pläne und Handlungen der Tutees eingehen, ihnen ein Gefühl von Sicherheit für das Erproben eigener Ideen vermitteln und inhaltliche Hinweise ausschließlich dann geben, wenn ihre Tutees Probleme haben. Die Tutoren sollten motivierende oder emotional unterstützende Aussagen nicht nur bei Problemen der Tutees, sondern regelmäßig äußern.

11.2 Diskussion der Intervention (Forschungsfrage 2)

Die Diskussion der Ergebnisse der zweiten Forschungsfrage ist in drei Blöcke geteilt. Einerseits werden die Prä-Post-Mittelwertsunterschiede diskutiert. Im zweiten Block werden die Prädiktoren dieser Unterschiede thematisiert. Abschließend werden etwaige differentielle Effekte der Kausalitätsorientierungen in Bezug auf die zuvor untersuchten Prädiktoren besprochen.

11.2.1 Diskussion der Prä-Post-Mittelwertsunterschiede

Aufbauend auf den Befunden von Berger et al. (2017), Hänze et al. (2018) und der Metaanalyse von Su und Reeve (2011) wurden die Hypothesen H2a₁ und H2b₁ aufgestellt, dass eine Intervention den (selbstberichteten) autonomieunterstützenden Motivierstil von Tutoren mit einer Effektstärke von etwa $d = 0.8$ fördern und den selbstberichteten kontrollierenden Motivierstil von Tutoren mit einer Effektstärke von etwa $d = 0.8$ reduzieren kann. Die in Kapitel 10.2.1 vorgestellten Prä-Post-Mittelwertvergleiche weisen große Effektstärken auf ($d_{\Delta SMot_A} = 1.23, p < .001$; $d_{\Delta SMot_K} = -1.58, p < .001$). Die gefundenen Effektstärken sind also über den erwarteten Werten. Somit wäre die in Kapitel 8.3 vorgestellte Intervention stärker als erwartet. Dies könnte einerseits dadurch begründet werden, dass sich die in Kapitel 8.3 vorgestellte Intervention an verschiedenen Empfehlungen orientiert hat (Berger et al., 2017; Reeve, 1998; Su & Reeve, 2011). Andererseits könnten die Motivierstile von Tutoren im Vergleich zu Lehrkräften leichter veränderlich sein. Dass die zeitlich stabile (Deci & Ryan, 1985) kontrollierende Kausalitätsorientierung der Tutoren mit einer mittleren Effektstärke durch die Intervention geändert wurde (siehe Tabelle 10.2), stützt weiter die Vermutung, dass sich motivationale Orientierungen der Tutoren verhältnismäßig leicht ändern lassen. Insgesamt können die Hypothesen H2a₁ und H2b₁ angenommen werden. Im Folgenden werden Prädiktoren dieser Änderungen der Motivierstile diskutiert.

11.2.2 Diskussion der Prädiktoren der Prä-Post-Mittelwertsunterschiede

Entsprechend den Hypothesen H2a_{2.I-VI} und H2b_{2.I-VI} wurde vermutet, dass die Mittelwertsunterschiede der Motivierstile von den jeweiligen Kausalitätsorientierungen, der Zustimmung der Teilnehmenden zu den Interventionsinhalten, der vorherigen Bekanntheit dieser Inhalte, der Zeit zwischen der Intervention und Messzeitpunkt 2 sowie der Deutschnote und dem Geschlecht der Tutoren vorhergesagt werden können. Es wurde vermutet, dass der selbstberichtete autonomieunterstützende Motivierstil der Tutoren umso mehr zunimmt, je stärker ihre autonome Kausalitätsorientierung ausgeprägt ist (H2a_{2.I}), je mehr sie den Interventionsinhalten zustimmen (H2a_{2.II}), je weniger ihnen diese Inhalte vorher bekannt waren (H2a_{2.III}) und je besser die Deutschnote ist (H2a_{2.V}). Diese Zunahme sollte unabhängig von der Zeit zwischen der Intervention und Messzeitpunkt 2 sein (H2a_{2.IV}). Weiterhin wurde vermutet, dass der autonomieunterstützende Motivierstil von weiblichen Teilnehmenden aufgrund ihres möglicherweise ohnehin höheren präinterventionellen autonomieunterstützenden Motivierstils (H2a_{2.VI}) weniger steigen wird. Bezüglich des kontrollierenden Motivierstils wurde umso mehr eine Reduktion vermutet, je schwächer die kontrollierende Kausalitätsorientierung der Tutoren ausgeprägt ist (H2b_{2.I}), je mehr sie den Interventionsinhalten zustimmen (H2b_{2.II}), je weniger ihnen diese Inhalte vorher bekannt waren (H2b_{2.III}) und je besser die Deutschnote ist (H2b_{2.V}). Diese Reduktion sollte unabhängig von der Zeit zwischen der Intervention und Messzeitpunkt 2 sein (H2b_{2.IV}). Weiterhin wurde vermutet, dass der kontrollierende Motivierstil von weiblichen Teilnehmenden aufgrund ihres möglicherweise ohnehin niedrigeren präinterventionellen kontrollierenden Motivierstils (H2b_{2.VI}) weniger reduziert wird. Um die Hypothesen zu überprüfen, wurden hierarchische lineare Regressionen gerechnet. Als abhängige Variablen wurden die Änderungen der selbstberichteten Motivierstile gewählt. Die unabhängigen Variablen wurden gemäß der Hypothesen H2a_{2.I-VI} und H2b_{2.I-VI} gewählt. Die Regressionsmodelle sind bezüglich der Änderungen beider Motivierstile und sowohl mit allen unabhängigen Variablen als auch nur mit den signifikanten Prädiktoren insgesamt signifikant geworden

(siehe Kapitel 10.2.2). Tabelle 11.3 stellt die Ergebnisse der Rechnungen aus Kapitel 10.2.2 verkürzt dar. Im Folgenden werden die Hypothesen H2a_{2.I-VI} und H2b_{2.I-VI} diskutiert.

Hypothese	Variable	Δ SMot_A	Δ SMot_K
H2a _{2.I} / H2b _{2.I}	Kau_A/K_Prä	n.s.	n.s.
H2a _{2.II} / H2b _{2.II}	Zustimmung	.300	-.379
H2a _{2.III} / H2b _{2.III}	Bekanntheit	-.326	.134
H2a _{2.IV} / H2b _{2.IV}	Zeit	n.s.	.235
H2a _{2.V} / H2b _{2.V}	Deutschnote	n.s.	n.s.
H2a _{2.VI} / H2b _{2.VI}	Geschlecht	n.s.	n.s.

Tabelle 11.3: Standardisierte Regressionsgewichte der signifikanten Prädiktoren ($p < .05$) der Regressionen zu den Änderungen des autonomieunterstützenden und des kontrollierenden Motivierstils aus Tabelle 10.3; Δ SMot_A/K = Änderung des autonomieunterstützenden/ kontrollierenden Motivierstils; Kau_A/K_Prä = präinterventionelle autonome/ kontrollierende Kausalitätsorientierung; Zustimmung = Zustimmung zu den Interventionsinhalten; Bekanntheit = Vorherige Bekanntheit der Interventionsinhalte; Zeit = Zeit zwischen Intervention und der Posterhebung.

Diskussion der Kausalitätsorientierungen

Entsprechend Tabelle 11.3 sagt die autonome Kausalitätsorientierung nicht den Zuwachs des autonomieunterstützenden Motivierstils vorher. Ebenso sagt die kontrollierende Kausalitätsorientierung nicht die Reduktion des kontrollierenden Motivierstils vorher. Es kann ausgeschlossen werden, dass sich die Kausalitätsorientierungen Varianz mit den signifikanten Prädiktoren teilen (siehe Anhang 12.4.2). Zusammengenommen sagen die Kausalitätsorientierungen nicht die Änderungen der Motivierstile vorher.

Die Ergebnisse passen zu den Befunden von Reeve (1998), der keine Interaktion des Interventionseffektes (Änderung des bilanzierten Motivierstils) mit der bilanzierten Kausalitätsorientierung nachweisen konnte. Wie in Kapitel 7.2.2 dargestellt, haben erst wenige Studien die Relevanz der Kausalitätsorientierungen für die Änderung der Motivierstile untersucht. Diese Studien zeigen inkonsistente Ergebnisse. Reeve et al. (2018) konnten zeigen, dass die jeweiligen Kausalitätsorientierungen die entsprechenden postinterventionellen Motivierstile unter Kontrolle der Präwerte vorhersagen. Zwar konnte dies auch in der vorliegenden Studie bestätigt werden (siehe Anhang 12.4.3), allerdings sind die standardisierten Regressionsgewichte mit $\beta_{Kau_A_Prä \rightarrow SMot_A_Post} = .22$ und $\beta_{Kau_K_Prä \rightarrow SMot_K_Post} = .21$ deutlich kleiner als bei Reeve et al. (2018), der Regressionsgewichte in Höhe von $\beta_{Kau_A_Prä \rightarrow SMot_A_Post} = .40$ und $\beta_{Kau_K_Prä \rightarrow SMot_K_Post} = .32$ gefunden hat. Eine mögliche Ursache für die unterschiedlichen Befunde zwischen Reeve et al. (2018) und der vorliegenden Studie könnten die verwendeten Messinstrumente sein. Zwar sind die Instrumente zur Erfassung der Kausalitätsorientierungen bis auf die Übersetzung miteinander vergleichbar, die Instrumente zur Erfassung der Motivierstile jedoch nicht. Insbesondere wurde in Kapitel 8.4.1 argumentiert, weshalb das Instrument zur Erfassung der Motivierstile von Tutoren tutorspezifische Beschreibungen enthalten sollte. Darüber hinaus unterscheiden sich 1:1-Tutorsettings deutlich von regulärem Unterricht. Somit ist es durchaus möglich, dass die Motivierstile in Abhängigkeit der betrachteten Situationen von

unterschiedlichen Prädiktoren beziehungsweise unterschiedlich stark vorhergesagt werden (Reeve et al., 2018). Zusammengenommen werden die Hypothese $H2a_{2.I}$ und $H2a_{2.II}$ mit der Anmerkung zurückgewiesen, dass die Kausalitätsorientierungen von Tutoren für ihre Motivierstile relevant sind.

Diskussion der Zustimmung zu den Interventionsinhalten

Entsprechend Tabelle 11.3 war die Zustimmung der Teilnehmenden zu den Interventionsinhalten in allen betrachteten hierarchischen linearen Regressionen ein signifikanter Prädiktor. Je stärker die Teilnehmenden den Inhalten der Intervention zustimmten, desto größer war der Zuwachs des autonomieunterstützenden Motivierstils ($\beta = .300, p < .001$) und desto größer war die Reduktion ihres kontrollierenden Motivierstils ($\beta = -.379, p < .001$). Folglich werden die Hypothesen $H2a_{2.II}$ und $H2b_{2.II}$ angenommen. Die Befunde passen zu den Befunden von Reeve (1998). Ob die Zustimmung zu den Interventionsinhalten für alle Teilnehmenden relevant ist oder eher für Teilnehmende mit einer ausgeprägten kontrollierenden Kausalitätsorientierung, wie die Befunde von Reeve (1998) nahelegen, wird in Kapitel 11.2.3 präsentiert.

Diskussion der vorherigen Bekanntheit der Interventionsinhalte

Entsprechend Tabelle 11.3 war die vorherige Bekanntheit der Interventionsinhalte in der hierarchischen linearen Regression bezüglich der Änderung des autonomieunterstützenden Motivierstils ein signifikanter Prädiktor. Je bekannter die Interventionsinhalte schon vor der Intervention waren, desto kleiner war der Zuwachs des autonomieunterstützenden Motivierstils ($\beta = -.326, p < .001$). Eine Erklärung der Befunde könnte sein, dass einige Teilnehmende tatsächlich autonomieunterstützende Verhaltensweisen als „gute Betreuung“ kannten. Es ist aufgrund des Alters nicht davon auszugehen, dass sie Autonomieunterstützung als psychologisch-didaktisches Konzept kannten. Dieses Kenntnis bzw. Wahrnehmung als „gute Betreuung“ könnte einen ausgeprägteren präinterventionellen autonomieunterstützenden Motivierstil zur Folge haben. Aufgrund dessen könnten sie ein geringeres Änderungspotential haben. Eine nachträgliche Rechnung bestätigt, dass Teilnehmende mit einem stark ausgeprägten präinterventionellen autonomieunterstützenden Motivierstil nach der Intervention erwartungskonform eine höhere vorherige Bekanntheit der Inhalte angeben (siehe Anhang 12.4.4). Weiterhin scheinen diese Teilnehmenden deutlich weniger von der Intervention zu profitieren. Diese Befunde erklären zum Teil, warum die Änderung des autonomieunterstützenden Motivierstils mit einem negativen Regressionsgewicht von der vorherigen Bekanntheit der Inhalte und dem damit einhergehenden geringeren Änderungspotential vorhergesagt wird. Somit scheint die Intervention eher für Teilnehmende mit einem schwach oder mäßig ausgeprägten autonomieunterstützenden Motivierstil geeignet zu sein. Insgesamt kann Hypothese $H2a_{2.III}$ angenommen werden, wobei die Intervention diesbezüglich überarbeitet werden sollte. Konkrete Vorschläge zur Überarbeitung der Intervention werden in Kapitel 11.6.1 gegeben.

Die Relevanz der Bekanntheit für die Änderung des kontrollierenden Motivierstils ist grenzwertig. Je mehr die Probanden die Interventionsinhalte schon vor der Intervention kannten, desto geringer war die Reduktion des kontrollierenden Motivierstils ($\beta = .134, p_{Bootstrap} = .037$). Die vorherige Bekanntheit der Interventionsinhalte scheint eher mit der Änderung des autonomieunterstützenden Motivierstils und weniger mit der Änderung des kontrollierenden Motivierstils zusammenzuhängen. Aus der genauen Betrachtung des Items zur Erfassung der Bekanntheit ergibt sich ein möglicher systematischer Fehler. Durch die Itemformulierung *Wie bekannt bzw. vertraut waren dir die Inhalte des Seminartages (im Besonderen die Betreuungsregeln) VOR dem Seminartag?* wird der Fokus auf

die Betreuungsregeln gelegt. Entsprechend Kapitel 8.3 sind die Betreuungsregeln vor allem Zusammenfassungen der vier berücksichtigten autonomieunterstützenden Verhaltensgruppen, welche in Kapitel 4 vorgestellt wurden. Aufgrund der Itemformulierung ist nur eine vorläufige Aussage zu Hypothese $H2b_{2.III}$ möglich. Die aus dieser Fehlkonstruktion resultierende Implikation für zukünftige Forschung können Kapitel 11.6.2 entnommen werden. Zwar bleibt das Problem der Itemformulierung bestehen, doch könnten auch unterschiedliche Entwicklungspotentiale der Motivierstile die unterschiedlichen Regressionsgewichte erklären. Der präinterventionelle autonomieunterstützende Motivierstil hat einen Mittelwert von .626 und somit ein Wachstumspotential von höchstens .374 (siehe Kapitel 9.3). Der präinterventionelle kontrollierende Motivierstil hat einen Mittelwert und somit ein Reduktionspotential von .602, also mehr als das Eineinhalbfache. Somit könnten vorwissensbedingte Deckeneffekte aufgrund des größeren Änderungspotentials des kontrollierenden Motivierstils eine untergeordnete Rolle spielen. Insgesamt kann Hypothese $H2b_{2.III}$ unter Vorbehalt angenommen werden.

Diskussion der Zeit zwischen Intervention und Messzeitpunkt 2

Wie in Kapitel 8.1 beschrieben, konnte die Zeit zwischen der Intervention und Messzeitpunkt 2 aus schulorganisatorischen Gründen nicht schulübergreifend konstant gehalten werden. Deshalb wurde diese Zeit erfasst und als weitere Kontrollvariable für Forschungsfrage 2 berücksichtigt. Entsprechend Tabelle 11.3 war diese Zeit ein Prädiktor der Änderung des kontrollierenden Motivierstils und kein Prädiktor der Änderung des autonomieunterstützenden Motivierstils. Der Befund bezüglich des autonomieunterstützenden Motivierstils ist erwartungskonform und im Einklang mit Reeve (1998). Auf Basis der Befunde kann nicht ausgeschlossen werden, dass der Interventionseffekt über längere Zeiträume, wie beispielsweise einem Jahr, nachlässt. Ein kurzfristiger Effekt der Zeit scheint allerdings nicht vorzuliegen. Deshalb wird die Hypothese $H2a_{2.IV}$ angenommen.

Je mehr Zeit zwischen der Intervention und dem zweiten Messzeitpunkt vergangen ist, desto geringer war die Reduktion des kontrollierenden Motivierstils ($\beta = .235, p = .001$, siehe Tabelle 11.3). Da dieses Ergebnis durch ein Bootstrapping-Verfahren abgesichert wurde, ist dieser Befund vermutlich kein statistisches Artefakt der Verteilung der Zeit zwischen Intervention und Messzeitpunkt 2 (siehe Kapitel 9.3). Die Ergebnisse könnten mit Vergessen bzw. einer Rückentwicklung der Motivierstile begründet werden. Reeve et al. (2018) konnten zeigen, dass die Motivierstile von Lehrkräften nur schwer veränderlich sind. Bei Tutoren scheinen die Motivierstile vergleichsweise leicht veränderlich zu sein (siehe Kapitel 11.2.1). Diese Veränderlichkeit hat möglicherweise die Kehrseite, dass sich die Motivierstile auch leichter zurückentwickeln. Insgesamt muss $H2b_{2.IV}$ abgelehnt werden und daraus resultierende Implikationen für die Intervention können Kapitel 11.6.1 entnommen werden.

Diskussion der Deutschnote und des Geschlechts der Tutoren

Die Deutschnote sowie das Geschlecht der Tutoren wurden in den Rechnungen in Kapitel 10.2 als Kontrollvariablen berücksichtigt. In keiner der hierarchischen linearen Regressionen waren sie Prädiktoren für die Änderungen der Motivierstile. Das Geschlecht und die Deutschnote scheinen also keinen Einfluss auf den Interventionserfolg zu haben. Entsprechend werden die Hypothesen $H2a_{2.V}$, $H2a_{2.VI}$, $H2b_{2.V}$ und $H2b_{2.VI}$ zurückgewiesen.

11.2.3 Diskussion differentieller Effekte

Tabelle 11.4 stellt die Ergebnisse der Rechnungen aus Kapitel 10.2.3 verkürzt dar. Im Folgenden werden die Hypothesen $H2a_{3,I-II}$ und $H2b_{3,I-II}$ diskutiert.

Hypothesen	$\Delta SMot_A$			Hypothesen	$\Delta SMot_K$		
	Gesamt	▲	▼		Gesamt	▲	▼
$H2a_2/ H2a_{3,I-II}$.300	.263	.275	$H2b_2/H2b_{3,I-II}$	-.379	-.302	-.419

Tabelle 11.4: Signifikante ($p < .05$) standardisierte Regressionsgewichte der Zustimmung zu den Interventionsinhalten in den Regressionen bezüglich der Änderung des autonomieunterstützenden und des kontrollierenden Motivierstils aus den Tabellen 10.3, 10.5 und 10.6; $\Delta SMot_A/K$ = Änderung des autonomieunterstützenden/ kontrollierenden Motivierstils.

Bezüglich der Änderung des autonomieunterstützenden Motivierstils sind die Regressionsgewichte der Zustimmung zu den Interventionsinhalten in ähnlicher Höhe. Somit scheinen die Tutoren unabhängig von ihrer autonomen Kausalitätsorientierung in Abhängigkeit ihrer Zustimmung zu den Interventionsinhalten von der Intervention zu profitieren. Folglich können die Hypothesen $H2a_{3,I}$ und $H2a_{3,II}$ abgelehnt werden. Bezüglich des kontrollierenden Motivierstils zeigen die Regressionen leicht unterschiedliche Regressionsgewichte (siehe Tabelle 11.4). Da die Zustimmung zu den Interventionsinhalten für Teilnehmende mit einer schwach ausgeprägten kontrollierenden Kausalitätsorientierung ein signifikanter Prädiktor der Reduktion des kontrollierenden Motivierstils ist, muss Hypothese $H2b_{3,I}$ abgelehnt werden. Entsprechend Reeve (1998) sowie $H2b_{3,I}$ und $H2b_{3,II}$ hätte man erwartet, dass die Zustimmung zu den Interventionsinhalten für die Gruppe mit der ausgeprägteren kontrollierenden Kausalitätsorientierung relevanter ist. Da es sogar umgekehrt ist und das Regressionsgewicht betragsmäßig unter .6 ist, wird auch Hypothese $H2b_{3,II}$ abgelehnt. Die Befunde passen somit nicht zu den Befunden von Reeve (1998). Reeve (1998) interpretiert seine Ergebnisse so, dass Teilnehmende mit einer relativ kontrollierenden bilanzierten Kausalitätsorientierung einen Konzeptwechsel durchlaufen müssen, um von der Intervention profitieren zu können. Weiterhin vermutet Reeve (1998), dass die Teilnehmenden mit einer relativ autonomen bilanzierten Kausalitätsorientierung den Inhalten schon vor der Intervention zugeneigt sind, deshalb keinen Konzeptwechsel durchlaufen müssen und folglich unabhängig von der Zustimmung zu den Interventionsinhalten profitieren. Entgegen der Erklärung von Reeve (1998) scheint es in der vorliegenden Studie nicht so zu sein, dass eine Gruppe einen Konzeptwechsel durchlaufen muss und die andere nicht. Vielmehr scheint die Zustimmung zu den Interventionsinhalten für alle Teilnehmenden einen ähnlichen Einfluss auf die Änderungen der Motivierstile zu haben.

11.3 Diskussion der Verhaltensprädiktoren (Forschungsfrage 3)

Im Folgenden werden die Ergebnisse aus Kapitel 10.3 diskutiert. Die Fitwerte der ausgewerteten Modelle aus den Abbildungen 10.6 und 10.7 (siehe Tabelle 10.10 und Tabelle 10.14) können als hinreichend angesehen werden, da der CFI über 0.9, der RMSEA und der SRMR unter 0.08 und das Verhältnis des χ^2 zur Anzahl der Freiheitsgrade unter 5 ist (Byrne, 2016). Da der CFI in allen ausgewerteten Modellen über .95, der RMSEA unter 0.05 und der SRMR unter .07 ist, erfüllen die Modelle sogar die strengeren Grenzwerte für

einen guten Modellfit (Byrne, 2016). Lediglich die χ^2 -Tests sind trotz des angemessenen Verhältnisses in allen Fällen signifikant geworden. Da die Signifikanzniveaus zwischen .005 und .05 lagen und die Verhältnisse der χ^2 zur Anzahl der Freiheitsgrade unter 5 sind, können alle ausgewerteten Modelle als hinreichend fittend angesehen werden (Byrne, 2016; Kyriazos, 2018). Neben den vorgestellten und verglichenen Modellen sind noch alternative Modelle denkbar. Beispielsweise könnte das Modell untersucht werden, bei dem alle Konstrukte nur auf die Motivierstile wirken und nicht miteinander zusammenhängen. Diese alternativen Modelle fitten die Daten im Vergleich zu den Modellen aus den Abbildungen 10.6 und 10.7 mit $p < .001$, $\Delta CFI > .08$ schlechter.

Die Generalfaktoren des Drucks von Innen, Oben und Unten (spontan)

Die Modellvergleiche in den Tabellen 10.11 und 10.14 sprechen gegen die statistische Gruppierung der Konstrukte in einen Druck von Unten (spontan) und einen Druck von Innen. Somit kann Hypothese H3c, dass der Druck von Innen die Motivierstile partiell mediiert über den Druck von Unten (spontan) vorhersagt, nicht beantwortet werden. Zwar hängen die Aspekte des Drucks von Innen und Unten (spontan) miteinander zusammen, doch konnten keine latenten Konstrukte des Drucks von Unten und des Drucks von Innen bestätigt werden. Gemäß der Hypothese H3c könnte weiter untersucht werden, ob die Aspekte des Drucks von Innen die Motivierstile über die Aspekte des Drucks von Unten (spontan) partiell oder vollständig mediiert vorhersagen. Diese alternativen Modelle wurden untersucht und beschreiben die Daten signifikant ($p < .001$) schlechter. Detaillierte Ergebnisse finden sich im Anhang (siehe Kapitel 12.4.5). Zusammengenommen muss Hypothese H3c abgelehnt werden.

Nichtsdestotrotz dienen die Begriffe der verbalen Kategorisierung, da sich die Aspekte des Drucks von Unten direkt auf die Tutorwahrnehmungen der Tutees beziehen. Die Aspekte des Drucks von Innen sind nicht von einer bestimmten Betreuungssituation oder Tutees abhängig, sondern ergeben sich aus allgemeinen Einstellungen, Überzeugungen oder Wissen. Bei Pelletier et al. (2002) sowie Rocchi et al. (2013) bestand der Druck von Unten ausschließlich aus der wahrgenommenen Motivation der Betreuten. Es wurden in Anlehnung an die Selbstbestimmungstheorie (siehe Kapitel 3) zwar verschiedene Dimensionen der wahrgenommenen Motivation erhoben, doch sollten diese Dimensionen die bilanzierte intrinsische vs. extrinsische Motivation der Betreuten abbilden. Somit war in Anlehnung an die Selbstbestimmungstheorie bei Pelletier et al. (2002) sowie Rocchi et al. (2013) zu erwarten, dass sich ein Generalfaktor ergibt (siehe Kapitel 3). In der vorliegenden Studie wurden neben der wahrgenommenen Motivation auch das wahrgenommene Engagement und die wahrgenommene Selbstständigkeit erhoben. Auch wenn die beiden letztgenannten konzeptuell zum Druck von Unten gehören, scheinen sie statistisch nicht mit der wahrgenommenen Motivation über einen Generalfaktor auf die Motivierstile zu wirken. Der Begriff des Drucks von Innen geht nicht auf Pelletier et al. (2002) oder Rocchi et al. (2013) zurück, sondern wurde von Reeve (2009) geprägt. Reeve (2009) hat in Anlehnung an Pelletier et al. (2002) verschiedene Prädiktoren der Motivierstile kategorisiert und in Abgrenzung zu äußeren Rahmenbedingungen (Druck von Oben) und dem Einfluss durch die Wahrnehmung der Betreuten (Druck von Unten) alle Einflüsse der Überzeugungen, Werte und persönlicher Dispositionen der Betreuten (Druck von Innen) abgegrenzt. Bei Reeve (2009) dienen die Kategorien allerdings ausschließlich der Einordnung der Konstrukte und stellen keine vermuteten statistischen Generalfaktoren dar. Somit widersprechen die Befunde weder der Studie von Pelletier et al. (2002), Reeve (2009) noch Rocchi et al. (2013). Deshalb wird im Folgenden zwar von Aspekten des Drucks von Unten und Innen gesprochen, allerdings bezieht sich die nachfolgende Diskussion auf die Modelle aus den Abbildungen 10.6 und

10.7, also die Modelle mit einem direkten Bezug aller Konstrukte auf die beobachteten Motivierstile.

Die Prädiktoren des kontrollierenden Verhaltens

Die in Tabellen 10.12 und 10.16 dargestellten Regressionen der Strukturgleichungsmodelle 6 und 9 (siehe Abbildungen 10.6 und 10.7) bestätigen die Hypothese H3b, dass sowohl Aspekte des Drucks von Innen als auch Aspekte des Drucks von Unten den beobachteten kontrollierenden Motivierstil vorhersagen. Deshalb kann die Hypothese H3b insgesamt angenommen werden. Die Befunde werden im Folgenden differenzierter diskutiert.

Das Verständnis der Variablenkontrollstrategie, das wahrgenommene Tuteeengagement und die wahrgenommene Tuteemotivation scheinen irrelevant für den kontrollierenden Motivierstil der Tutoren zu sein. Von den drei Aspekten des Drucks von Unten ist die wahrgenommene Tuteeselbstständigkeit ein signifikanter Prädiktor des kontrollierenden Verhaltens. Erwartungskonform verhalten sich Tutoren umso kontrollierender, je weniger selbstständig sie ihre Tutees wahrnehmen (siehe Tabellen 10.12 und 10.16; siehe auch Reeve, 2009). Man könnte argumentieren, dass Tutees mit einer geringeren Selbstständigkeit auch tatsächlich mehr Hilfe benötigen und Tutoren folglich mehr eingreifen sollten. Entsprechend Kapitel 8.4.2 berücksichtigen die Auswertungskategorien dies. Die Tutoren haben sich also über die nötige Hilfe hinaus umso kontrollierender verhalten, je unselbstständiger sie ihre Tutees wahrgenommen haben. Das wahrgenommene Tuteeengagement und die wahrgenommene Tuteemotivation sagen das kontrollierende Verhalten nicht signifikant vorher. Die Irrelevanz dieser beiden Konstrukte kann einerseits dadurch bedingt sein, dass sie nicht mit dem kontrollierenden Verhalten zusammenhängen oder sie sich mit der wahrgenommenen Tuteeselbstständigkeit Varianz teilen (Field, 2009). Weder das wahrgenommene Engagement ($\rho = .004, p = .961$) noch die wahrgenommene Motivation ($\rho = -.138, p = .117$) hängen mit dem beobachteten kontrollierenden Motivierstil zusammen. Insgesamt können die Befunde bezüglich des Drucks von Unten als zu Pelletier et al. (2002) und Rocchi et al. (2013) passend bezeichnet werden, da diese Aspekte den beobachteten kontrollierenden Motivierstil vorhersagen.

Von den vier einbezogenen Aspekten des Drucks von Innen sagen drei den beobachteten kontrollierenden Motivierstil signifikant vorher. Probanden mit einem ausgeprägten Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung verhalten sich erwartungskonform weniger kontrollierend. Das Fähigkeitsselbstkonzept bezüglich Physik hängt positiv mit dem kontrollierenden Tutorverhalten zusammen. Dies wurde in den Kapiteln 5.2.1 und 7.2.2 anders vermutet. Der positive Zusammenhang des Selbstkonzeptes bezüglich Physik und des kontrollierenden Verhaltens könnte so begründet werden, dass die Tutoren aufgrund ihres ausgeprägteren Selbstkonzeptes bezüglich Physik ungeduldiger sind, wenn Tutees fachliche Fehler machen. Möglicherweise haben Tutoren ein bestimmtes Vorgehen im Kopf, welches sie den Tutees aufzeigen wollen und dadurch die Perspektive der Tutees übergehen. Weiterhin scheinen sich die Tutoren umso kontrollierender zu verhalten, je intrinsisch motivierter ihre Gründe für ihre Tutorentätigkeit sind (siehe Tabelle 10.12 und 10.16). Im Kapitel 5.2.1 wurde die Erwartungshaltung dargelegt, dass die Tutoren sich umso weniger kontrollierend verhalten, je intrinsischer motiviert sie sind. Diese Erwartung basierte auf der Annahme, dass intrinsisch motivierte Tutoren ihre Tutees eher intrinsisch unterstützen (z. B. durch Berücksichtigung der Tuteeinteressen). Möglicherweise haben Tutoren mit einer ausgeprägten intrinsischen Motivation ein stärkeres Bedürfnis zu helfen und laufen deshalb Gefahr, mehr zu helfen als nötig. Zusammengenommen scheinen Tutoren mit einem ausgeprägten Selbstkonzept bezüglich Physik und einer eher intrinsischen Motivation einen Drang zum Helfen zu haben, welcher zu unnötigen Eingriffen ins Experi-

ment führt. Insgesamt können die Befunde bezüglich des Drucks von Innen als zu Pelletier et al. (2002) und Rocchi et al. (2013) passend bezeichnet werden, da einzelne Aspekte des Drucks von Innen den beobachteten kontrollierenden Motivierstil vorhersagen.

Das Verständnis der Variablenkontrollstrategie sagt nicht den beobachteten kontrollierenden Motivierstil vorher. Die ursprüngliche Vermutung war, dass Tutoren mit einem geringen Verständnis der Variablenkontrollstrategie sich unsicher bei der Betreuung fühlen und deshalb eher die Situation kontrollieren wollen und sich nicht zutrauen, inhaltlich angemessen auf Ansätze der Tutees eingehen zu können. Da die Tutoren alle vielfach mit dem Experiment konfrontiert waren, könnten sich alle Tutoren unabhängig vom tatsächlichen Verständnis der Variablenkontrollstrategie hinreichend kompetent gefühlt haben. Somit kann die Teildimension UN des Verständnisses der Variablenkontrollstrategie als Prädiktor kontrollierenden Verhaltens ausgeschlossen werden. Möglicherweise spielen die übrigen Dimensionen IN und IS für das grundsätzliche Verständnis und damit die Vorhersage des kontrollierenden Verhaltens eine größere Rolle. Alternativ könnte die fehlende Relevanz des Verständnisses der Variablenkontrollstrategie auch durch das Tutorarbeitenblatt mit Hinweisen erklärt werden (siehe Kapitel 8.2). Da sie alle Lösungen für etwaige inhaltliche Probleme hatten, waren die Tutoren weniger auf ihr eigenes Verständnis angewiesen. Zusammengenommen deutet die Irrelevanz des Verständnisses der Variablenkontrollstrategie für das kontrollierende Tutorverhalten darauf hin, dass alle Tutoren hinreichend vorbereitet wurden, sodass individuelle Unterschiede im Verständnis das Verhalten der Tutoren nicht weiter beeinflussen. Die kontrollierende Kausalitätsorientierung musste aufgrund schlechter Fitwerte der Strukturgleichungsmodelle entfernt werden. Allerdings korreliert die kontrollierende Kausalitätsorientierung mit dem kontrollierenden Tutorverhalten ($\rho = .194, p = .027$). Ob dieser Zusammenhang unter Berücksichtigung der anderen vermuteten Prädiktoren erhalten bleibt, konnte in der vorliegenden Studie nicht entschieden werden. Die postinterventionelle kontrollierende Kausalitätsorientierung hatte mit $\alpha_{Kau_K_Post} = .856$ eine gute interne Konsistenz. Folglich scheinen keine umfangreichen Erhebungsprobleme bei der kontrollierenden Kausalitätsorientierung vorzuliegen (Field, 2009). Die Fitwerte eines Strukturgleichungsmodells können sehr empfindlich auf zu kleine Stichproben reagieren (Kyriazos, 2018). Durch das Entfernen der kontrollierenden Kausalitätsorientierung wurde die Anzahl der Freiheitsgrade von 815 auf 398 und damit etwa um die Hälfte reduziert. Das Entfernen der kontrollierenden Kausalitätsorientierung reduziert also deutlich die Modellkomplexität, sodass die 116 bzw. 131 berücksichtigten Probanden ausreichend waren, um ein Modell mit akzeptablen Fitwerten zu schätzen (Kyriazos, 2018). Zusammengenommen musste die kontrollierende Kausalitätsorientierung weder wegen Reliabilitäts- oder Validitätsproblemen des Fragebogens entfernt werden, sondern aufgrund der Stichprobengröße (Kyriazos, 2018).

Die Prädiktoren des autonomieunterstützenden Verhaltens

Für den beobachteten autonomieunterstützenden Motivierstil konnte keine gemeinsame Komponente aus den Verhaltensweisen extrahiert werden. Deshalb wurden die vermutet autonomieunterstützenden Verhaltensweisen mit dem wahrgenommenen autonomieunterstützenden Motivierstil korreliert, um zu untersuchen, welche der vermutet autonomieunterstützenden Verhaltensweisen auch als solche wahrgenommen werden. Das Eingehen der Tutoren auf Aussagen der Tutees (Kategorie E3) war die einzige Kategorie, welche als autonomieunterstützend wahrgenommen wurde. Deshalb wurde E3 als Ersatz für den beobachteten autonomieunterstützenden Motivierstil im Strukturgleichungsmodell 9 berücksichtigt (siehe Abbildung 10.7). Keiner der vermuteten Prädiktoren kann signifikant Varianz des Auftretens der Kategorie E3 aufklären (siehe Tabelle 10.16). Aufgrund der

abgesicherten Korrelation zum wahrgenommenen autonomieunterstützenden Motivierstil kann die Validität von E3 nur schwierig angezweifelt werden (siehe Kapitel 10.3, 11.1 sowie 12.3.5). Keines der sieben untersuchten Konstrukte sagte den beobachteten autonomieunterstützenden Motivierstil im Sinne von E3 voraus. Ein Grund für die scheinbare Irrelevanz der Prädiktoren könnte sein, dass E3 das Konstrukt der Autonomieunterstützung nicht vollständig abbildet. Insgesamt kann aufgrund der Daten Hypothese H3a nur abgelehnt werden, wobei dieser Befund aufgrund der unvollständigen Erfassung des beobachteten autonomieunterstützenden Motivierstils als sehr vorläufig anzusehen ist. Somit passen die Befunde nicht zu Pelletier et al. (2002) und Rocchi et al. (2013), was ebenfalls auf die unvollständige Erfassung des beobachteten autonomieunterstützenden Motivierstils zurückgeführt werden könnte.

11.4 Diskussion der Tuteeebene (Forschungsfrage 4)

Im Folgenden werden einerseits die Wirkzusammenhänge der Konstrukte auf der Tuteeebene (H4a) und andererseits eine mögliche Relevanz der Geschlechterkomposition (H4b) diskutiert.

11.4.1 Zusammenhänge der Konstrukte auf der Tuteeebene

Entsprechend Hypothese H4a wurde untersucht, inwiefern der wahrgenommene autonomieunterstützende Motivierstil, die Befriedigung des Kompetenzerlebens, sowie die intrinsische Motivation der Tutees nach dem Tutoring miteinander zusammenhängen. In einem ersten Schritt wurden dazu Strukturgleichungsmodelle aufgestellt und miteinander verglichen, welche die intrinsische Motivation der Tutees vor der Durchführung des Experiments berücksichtigen. Entsprechend Tabelle 10.18 konnte bestätigt werden, dass diese anfängliche intrinsische Motivation sowohl das Kompetenzerleben als auch die intrinsische Motivation nach der Durchführung des Experiments beeinflusst. Anschließend wurden Strukturgleichungsmodelle aufgestellt und miteinander verglichen, welche verschiedene mögliche Einflüsse des wahrgenommenen autonomieunterstützenden Motivierstils auf die intrinsische Motivation abgebildet haben. Weiterhin wurden auch Mediationsanalysen nach Baron und Kenny (1986) sowie mit Hilfe des SPSS-Makros PROCESS (Hayes, 2018) und des R Pakets *mediation* (Tingley et al., 2014) gerechnet (siehe Anhang Kapitel 12.3.6). Alle oben genannten Analysen bestätigen, dass der wahrgenommene autonomieunterstützende Motivierstil sowohl mit dem Kompetenzerleben als auch mit der intrinsischen Motivation zusammenhängt. Dabei scheint der wahrgenommene autonomieunterstützende Motivierstil die intrinsische Motivation partiell mediiert über das Kompetenzerleben zu beeinflussen. Unter Berücksichtigung der in Kapitel 11.5 aufgeführten Limitationen können die Befunde wie folgt interpretiert werden:

In Kapitel 5.1 wurde aus bestehender Forschungsliteratur das Modell abgeleitet, dass die Motivierstile von Betreuenden die psychologischen Grundbedürfnisse ihrer Betreuten beeinflussen und sich diese Grundbedürfnisse wiederum auf verschiedene Zielvariablen auswirken. Die Befunde der vorliegenden Studie bestätigen dies für das psychologische Grundbedürfnis des Kompetenzerlebens und die Zielvariable der intrinsischen Motivation von Tutees, welche von Cross-Age-Tutoren bei der Durchführung physikalischer Experimente betreut wurden. Je mehr sich die Tutees in ihrem Autonomieerleben unterstützt gefühlt haben, desto kompetenter erleben sie sich und desto stärker sind sie intrinsisch motiviert.

Aufgrund mehrerer experimenteller Studien kann eher davon ausgegangen werden, dass die wahrgenommene Autonomieunterstützung das Kompetenzerleben beeinflusst und der Einfluss nicht umgekehrt ist (z. B. Jang et al., 2016b; Jang et al., 2009). Vergleicht man das untersuchte Cross-Age-Tutoring-Setting mit regulärem Schulunterricht (Jang et al., 2009), scheint die Befriedigung des Kompetenzerlebens für die intrinsische Motivation vergleichbar wichtig zu sein: Entsprechend dem Strukturgleichungsmodell (siehe Abbildung 10.6) beträgt der standardisierte geschätzte Parameter des Prädiktors Kompetenzerleben bezüglich der intrinsischen Motivation $.53, p < .001$. Dazu passend haben Jang et al. (2009) Werte von $.47$ und $.55$ gefunden.

Entgegen der Hypothese H4a scheint der wahrgenommene autonomieunterstützende Motivierstil die Zielvariable intrinsische Motivation nicht nur indirekt über die psychologischen Grundbedürfnisse, sondern auch direkt zu beeinflussen. Im Gegensatz zu Jang et al. (2009) wurde in der vorliegenden Studie nicht die Befriedigung des Autonomieerlebens erhoben. Also könnte der wahrgenommene autonomieunterstützende Motivierstil die intrinsische Motivation vollständig mediiert über das Autonomie- und Kompetenzerleben beeinflussen. Bei Jang et al. (2009) hatte der wahrgenommene autonomieunterstützende Motivierstil auf das Autonomieerleben standardisierte Parameterschätzungen von $.47$ beziehungsweise $.52$. Die standardisierten Parameterschätzungen der Pfade von der Befriedigung des Autonomieerlebens zur intrinsischen Motivation betragen $.32$ beziehungsweise $.41$. Die standardisierten Parameterschätzungen der mediierten Pfade betragen somit $.15$ beziehungsweise $.21$. Der wahrgenommene autonomieunterstützende Motivierstil hat in der vorliegenden Studie bezüglich der intrinsischen Motivation eine standardisierte Parameterschätzung von $.25, p = .008$. Somit könnte der direkte Pfad von der wahrgenommenen Autonomieunterstützung zur intrinsischen Motivation in der vorliegenden Studie dem Pfad von der wahrgenommenen Autonomieunterstützung über die Befriedigung des Autonomieerlebens zur intrinsischen Motivation bei Jang et al. (2009) entsprechen. Zwar muss Hypothese H4a abgelehnt werden, doch legen die Befunde sowie der Vergleich zu Jang et al. (2009) die neue Hypothese nahe, dass der Einfluss des wahrgenommenen autonomieunterstützenden Motivierstils auf die intrinsische Motivation vollständig über das Kompetenz- und Autonomieerleben mediiert wird.

11.4.2 Relevanz der Geschlechterkomposition

Entsprechend Hypothese H4b wurde untersucht, ob die Tutees mit gleichgeschlechtlichem Tutor einen signifikant höheren wahrgenommenen autonomieunterstützenden Motivierstil, eine höhere Befriedigung des Kompetenzerlebens sowie eine höhere intrinsische Motivation im Vergleich zu Tutees mit einem Tutor des anderen Geschlechts angeben. Zur Untersuchung wurden die Tutoren und Tutees für die Analyse in zwei Gruppen aufgeteilt. Die Gruppen sind miteinander vergleichbar (siehe Kapitel 10.4.2). Durch die Mittelwertvergleiche können Unterschiede in der intrinsischen Motivation und dem Kompetenzerleben abgelehnt werden. Bezüglich des wahrgenommenen autonomieunterstützenden Motivierstils ist der T-Test abgesichert knapp nicht signifikant geworden. Der Mann-Whitney-U-Test ist knapp signifikant geworden. Folglich ist ein klares Urteil schwierig und es können lediglich mittlere bis große Zusammenhänge der Geschlechterkomposition mit dem wahrgenommenen autonomieunterstützenden Motivierstil ausgeschlossen werden.

Einerseits könnten die Befunde so interpretiert werden, dass die Signifikanz des Mann-Whitney-U-Tests ein statistischer Fehler ist (z. B. kumulierter Alphafehler). Dies würde

die Organisation von Tutoringsitzungen einfacher gestalten, da bezüglich der Tuteemotivation bei der Zuteilung nicht auf die Geschlechter geachtet werden müsste. Weiterhin kann dies als Chancengleichheit bezüglich der Geschlechter gesehen werden. Liegt kein Zusammenhang vor, wird kein Geschlecht gegenüber dem anderen bevorzugt. Falls die Geschlechterzusammensetzung tatsächlich mit dem wahrgenommenen autonomieunterstützenden Motivierstil zusammenhängt, hat dieser Zusammenhang eine kleine Größe (Field, 2009). Man könnte aufgrund der kleinen Größe des Zusammenhangs annehmen, dass nicht alle Tutees einen gleichgeschlechtlichen Tutor bevorzugen (Field, 2009). Statt den Tutees aufgrund dieser Befunde immer einen Tutor des gleichen Geschlechts zuzuweisen, könnte man die Tutees ihre Tutoren aussuchen lassen. Diese Auswahl könnte praktisch umgesetzt werden, indem die Tutoren und Tutees vor der Betreuung ein kurzes Gespräch im Plenum haben, sodass alle einander oberflächlich kennenlernen. Wenn die Tutees wollen, dürfen sie sich einen Tutor aussuchen. Andernfalls wird ihnen angesichts der Befunde von Leung (2015) ein gleichgeschlechtlicher Tutor zugewiesen.

Laut Leung (2015) profitieren Tutees bezüglich ihrer Lernleistungen deutlich von einem gleichgeschlechtlichen Tutor. Auf Basis der Befunde der vorliegenden Studie können negative Effekte von Tutoren des gleichen Geschlechts auf die intrinsische Motivation der Tutees ausgeschlossen werden. Zusammengenommen kann man den Tutees entweder die eigene Auswahl des Tutors gestatten oder sollte ihnen einen gleichgeschlechtlichen Tutor zuweisen.

11.5 Limitationen

Im Folgenden werden die Limitationen der vorliegenden Forschungsarbeit diskutiert.

Das Fehlen einer Kontrollgruppe

Die erste Limitation ist das Fehlen einer Kontrollgruppe. Entsprechend den großen Effektstärken der Prä-Post-Mittelwertvergleiche wurden die Hypothesen H1a und H1b angenommen. Da keine Kontrollgruppe vorhanden ist, können die großen Effektstärken der Prä-Post-Mittelwertvergleiche (siehe Kapitel 11.2.1) nicht als Effekte im kausalen Sinne interpretiert werden. Beispielsweise kann nicht ausgeschlossen werden, dass sich die Mittelwerte der Motivierstile der Probanden durch mehrfaches Ausfüllen der Fragebögen ändern. Aufgrund der zahlreichen Studien mit Kontrollgruppen ist ein Effekt der Intervention sehr wahrscheinlich (Wirksamkeit solcher Intervention für Tutoren: Berger et al., 2017; Wirksamkeit solcher Interventionen für verschiedene Personengruppen: Su und Reeve, 2011). Strukturgleichungsmodelle können rechnerisch keine kausalen Hypothesen auf Basis von korrelativ erhobenen Daten überprüfen. Somit können auch die Befunde aus Kapitel 10.3 und 10.4 ebenfalls nur korrelativ und nicht kausal interpretiert werden. Allerdings passen die auf diese angenommenen kausalen Wirkungen aufbauenden Modelle gut zu den Daten der vorliegenden Studie (siehe Kapitel 10.3 und 10.4). Diese Einschränkung ist insofern unproblematisch, da es nicht das Ziel der vorliegenden Studie war, die mehrfach gezeigten kausalen Wirkungen zu reproduzieren (z. B. Berger et al., 2017; Hänze et al., 2018), sondern Prädiktoren des Tutorverhaltens (siehe Kapitel 10.3) sowie eine detaillierte Analyse der Zusammenhänge der Konstrukte auf Tuteeebene unter Berücksichtigung ihrer psychologischen Grundbedürfnisse durchzuführen (siehe Kapitel 10.4).

Die Auswahl und Größe der Stichprobe

Die zweite, dritte und vierte Limitation resultieren aus der Stichprobe. Einerseits wurde in Kapitel 8.4.5 dargelegt, dass lediglich eine Gelegenheitsstichprobe vorliegt. Somit kann eine positive Auswahl nicht ausgeschlossen werden. Beispielsweise könnten besonders motivierte Schülerinnen und Schüler ausgewählt worden sein. Möglicherweise können die Befunde deshalb nicht auf durchschnittliche Schülerinnen und Schüler generalisiert werden. Dem entgegen steht, dass alle Schülerinnen und Schüler teilnehmen durften, die wollten. Zumindest für die betrachteten Schulen und bezüglich Forschungsfrage 2 könnte somit von einer Vollerhebung gesprochen werden (Field, 2009). Da die betrachteten Schulen Standortfaktoren von zwei bis vier hatten, wurden unterschiedliche Zusammensetzungen der Schülerschaft sowie regional unterschiedliche Strukturen berücksichtigt. Da keine der Schulen den Standortfaktor eins oder fünf hatte, können die Befunde nicht auf solche Schulen generalisiert werden.

Neben der Limitation durch die Auswahl der Probanden ergibt die Stichprobengröße eine weitere Einschränkung. Für die Beantwortung der zweiten Forschungsfrage war die Stichprobengröße ausreichend. Allerdings konnten aufgrund der Stichprobengröße keine kleinen Effekte beziehungsweise ausgeschlossen werden (Field, 2009). Nach Field (2009) benötigt man mindestens 400 Probanden, um kleine Effekte zuverlässig ausschließen zu können. Praktisch bedeutet dies, dass weniger relevante Prädiktoren der Interventionserfolge möglicherweise übersehen wurden. Im Folgenden wird dargelegt, dass die Stichprobengröße je nach betrachteter Quelle ausreichend oder für die Beantwortung der Forschungsfragen FF3 und FF4a knapp nicht ausreichend war. Die Stichprobengröße kann aus zwei Gründen für die Strukturgleichungsmodelle in Kapitel 10.3 und 10.4 problematisch sein. Einerseits können die Fitwerte eines Strukturgleichungsmodells sehr empfindlich auf zu kleine Stichproben reagieren (Kyriazos, 2018). Dass die Kausalitätsorientierungen mit jeweils 12 Items die Anzahl der Freiheitsgrade verdoppelt haben (siehe Kapitel 11.3), erklärt vermutlich die schlechten Fitwerte (Kyriazos, 2018). Somit kann keine Aussage zur Relevanz der Kausalitätsorientierungen und der identifizierten Motivation für die beobachteten Motivierstile getroffen werden. Weiterhin sollten auch die Gesamtmodellgüten der Strukturgleichungsmodelle kritisch diskutiert werden. Es gibt verschiedene Empfehlungen für benötigte Stichprobengrößen. Beispielsweise wird eine Stichprobe von mindestens 100 verlangt (Gorsuch, 2013), was in der vorliegenden Studie erfüllt ist. Andere Empfehlungen verlangen mindestens 250 (Hu & Bentler, 1999). Nach Westland (2010) ergibt sich die benötigte Stichprobengröße aus der Anzahl modellierter latenter Konstrukte k und der Anzahl eingesetzter Items gemäß der Formel:

$$N \geq 50 \left(\frac{p}{k} \right)^2 - 450 \left(\frac{p}{k} \right) + 1100$$

Modell	Anzahl latenter Konstrukte k	Anzahl eingesetzter Items p	Benötigte Stichprobe
FF3: Abbildung 10.6	8	28	138
FF3: Abbildung 10.7	9	29	172
FF4a: Abbildung 10.9	3	16	122
FF4a: Abbildung 10.11	4	22	138

Tabelle 11.5: Benötigte Stichproben für die berichteten Strukturgleichungsmodelle.

Entsprechend der Empfehlung von Westland (2010) fehlen für das Strukturgleichungsmodell aus Abbildung 10.6 einige Personen. Diese nach Westland (2010) berechnete Grenze basiert auf verschiedenen Simulationsstudien und darf nicht als Schwellenwert verstanden

werden. 138 Probanden sind demnach nicht deutlich besser als 131 Probanden. Die Schätzungen des Modells aus Abbildung 10.6 sind aufgrund der nahezu ausreichenden Personenanzahl zwar akzeptabel, jedoch sind diese Schätzungen nicht durch eine sehr große Stichprobe abgesichert. Folglich sollten die Befunde lediglich als erste Indizien interpretiert werden. Gleiches gilt für das Strukturgleichungsmodell aus Abbildung 10.11. Gemäß der Formel von Westland (2010) ist die Personenanzahl für die Schätzung des Modells aus Abbildung 10.7 nicht ausreichend, da mindestens 24% mehr Probanden nötig sind. Aus der zu kleinen Stichprobengröße kann allerdings nicht geschlussfolgert werden, dass die Modelle unzulässig sind. Solange die Modelle konvergieren und keine unzulässigen Lösungen produziert werden, wie beispielsweise *Heywood cases* (Kolenikov & Bollen, 2012), sind die Parameterschätzungen im Wesentlichen unbiased (Kyriazos, 2018). Da die Fitwerte empfindlich auf zu kleine Stichproben reagieren können und die Fitwerte der oben genannten vier Modelle akzeptabel bis gut waren, wurden die Ergebnisse berichtet und interpretiert.

Die aufgeklärten Varianzen sind insgesamt im kleinen bis mittleren Bereich (Field, 2009). Möglicherweise wurden relevante Prädiktoren nicht berücksichtigt (Field, 2009). Wie im letzten Absatz erläutert, konnten bezüglich Forschungsfrage 3 und 4 aufgrund der Stichprobengröße ohnehin keine weiteren Konstrukte berücksichtigt werden. Insgesamt sollten künftige Studien deutlich größere Stichproben akquirieren, um so die Möglichkeit zur Berücksichtigung weiterer Prädiktoren zu bekommen. Vorschläge, welche Konstrukte zukünftige Studien bei hinreichender Stichprobengröße berücksichtigen sollten, sind Kapitel 11.6.2 zu entnehmen.

Die Güte der eingesetzten Testinstrumente

Eine weitere Limitation bezieht sich auf die Gütekriterien der eingesetzten Skalen. Aufgrund der Voruntersuchungen (siehe Kapitel 8.4 und 9.2.6) schienen die meisten Skalen hinreichende Reliabilität und Validität zu besitzen. In den Kapiteln 11.1 bis 11.4 konnte dies für die meisten Skalen bestätigt werden. Einzelne Skalen haben keine hinreichende interne Konsistenz aufgewiesen und andere haben möglicherweise Validitätsprobleme. Beispielsweise hängt der selbstberichtete autonomieunterstützende Motivierstil weder mit Tuteevariablen noch mit den vermutet autonomieunterstützenden Verhaltensweisen zusammen. Allerdings hängt dieser Motivierstil erwartungskonform mit weiteren Variablen der Tutoren zusammen. Insgesamt sollte die prädiktive Validität des selbstberichteten autonomieunterstützenden Motivierstils von Tutoren weiter untersucht werden. Da der selbstberichtete kontrollierende Motivierstil erwartungskonform mit vermutet kontrollierenden Verhaltensweisen und einzelne autonomieunterstützende Verhaltensweisen mit dem wahrgenommenen autonomieunterstützenden Motivierstil zusammenhängen (siehe Kapitel 11.1), konnten die Forschungsfragen trotz der benannten Probleme hinreichend untersucht werden. Weiterhin ist die Erhebung der vorherigen Bekanntheit der Interventionsinhalte zu nennen. Einerseits wurde in Anlehnung an Reeve (1998) nur ein Item eingesetzt, andererseits schienen die Probanden das Item aufgrund der Formulierung teilweise missverstanden zu haben, weshalb die Befunde bezüglich des kontrollierenden Motivierstils nur als erste Anhaltspunkte zu interpretieren sind (siehe Kapitel 11.2.2).

Die soziale Erwünschtheit von Antworten und Angaben

Weiterhin kann nicht ausgeschlossen werden, dass alle Probanden die vermeintlich sozial erwünschten Antworten gegeben haben. Die Tutees wurden darauf hingewiesen, dass sie ihre Tutoren nicht bewerten und die Tutoren die Antworten nicht sehen. Wie bei Reeve (1998) wurden die Probanden mehrfach gebeten, ausschließlich entsprechend ihrer eigenen Meinung anzukreuzen. Allerdings könnten trotzdem einige Tutees tendenziell positivere

Antworten als tatsächlich zutreffend angegeben haben. Somit könnten die meisten Tutoren im Anschluss an die Intervention autonomieunterstützendes Verhalten als angemessen angegeben haben, selbst wenn sie dem ehrlich nicht zustimmen. Dies könnte eine mögliche Erklärung der fehlenden Korrelation zwischen dem selbstberichteten und beobachteten autonomieunterstützenden Motivierstil sein (siehe Kapitel 11.1). Ebenso kann nicht ausgeschlossen werden, dass sich die Tutoren in den videografierten Videos möglichst autonomieunterstützend und in anschließenden nicht beobachteten Tutorsitzungen deutlich kontrollierender Verhalten haben.

Statistische Limitationen und Probleme

Eine weitere Limitation bezieht sich auf die durchgeführten Mediansplits (siehe Kapitel 10.2.3 und 10.4.2). Mediansplits führen zu einem Informations- und Powerverlust (J. Cohen, 1983; Humphreys, 1978; Lagakos, 1988). Beispielsweise wurden die Probanden für die Überprüfung der Hypothesen $H2a_{3,I-II}$ und $H2b_{3,I-II}$ mit Hilfe von Mediansplits jeweils auf zwei Gruppen aufgeteilt. Ein Proband mit einer kontrollierenden Kausalitätsorientierung knapp unter dem Median ist in der selben Gruppe, wie ein Proband, der eine sehr schwach ausgeprägte kontrollierende Kausalitätsorientierung hat. Somit resultiert der Informationsverlust aus der Dichotomisierung der näherungsweise kontinuierlich verteilten Kausalitätsorientierungen und dem damit verbundenen Varianzverlust.

Die signifikanten korrelierten Messfehler beziehungsweise die korrelierten Residualvarianzen der in Kapitel 10.4.1 berichteten Strukturgleichungsmodelle könnten als Problem verstanden werden. Die korrelierten Messfehler wurden zwischen Items verschiedener Konstrukte auf null fixiert, zwischen Items desselben Konstrukts nicht. Fixiert man die Letzgenannten ebenfalls auf null, verschlechtern sich die Fitwerte des finalen Modells aus Abbildung 10.11 signifikant. Bezüglich der latenten Konstrukte und ihren Zusammenhängen ändern sich die Parameterschätzungen wenig und die Signifikanzen gar nicht. Somit sind die berichteten Zusammenhänge und insbesondere die gefundene partielle Mediation robust gegenüber verschiedenen Modellierungen, was die Belastbarkeit der berichteten Ergebnisse erhöht. Eine detaillierte Analyse kann dem Anhang (siehe Kapitel 12.4.6) entnommen werden.

11.6 Ausblick und Implikationen

Aus den Diskussionen der einzelnen Befunde in den Kapiteln 11.1 bis 11.4 ergeben sich einige Implikationen für die Intervention, für künftige Forschung und für die Schulpraxis.

11.6.1 Implikationen für die Intervention

Da der wahrgenommene autonomieunterstützende Motivierstil sowohl das Kompetenzerleben als auch die intrinsische Motivation zu beeinflussen scheint (siehe Kapitel 11.4.1), sollten Tutoren ihre Tutees autonomieunterstützend betreuen. Dafür sollten Tutoren umfangreiche Hilfe erhalten, da dies keine Selbstverständlichkeit ist, die von Tutoren ohne Training verlangt werden kann (Berger et al., 2017). Da kein durchschlagender Effekt des intervenierten selbstberichteten autonomieunterstützenden Motivierstils auf die Tuteeebene gefunden werden konnte (siehe Kapitel 11.1), sollte die Intervention weiterentwickelt werden. Da der beobachtete mit dem wahrgenommenen autonomieunterstützenden Motivierstil zusammenhängt, ist eine Überarbeitung der Intervention vermutlich ausreichend.

Kann durch die Überarbeitung der durchschlagende Effekt erreicht werden, stellt die überarbeitete Intervention aufgrund der großen Effektstärke eine hinreichende Hilfe für die Tutoren dar (siehe Kapitel 11.2.1). Aus den Befunden (siehe Kapitel 10) beziehungsweise deren Diskussion (siehe Kapitel 11.1 bis 11.4) resultieren einige Hinweise, inwiefern die Intervention überarbeitet werden sollte.

Inhaltliche Überarbeitung vorhandener Bausteine - Teil 1: Weniger Zurückhalten und mehr autonomieunterstützende Kommunikation im Sinne der Kategorien E1 bis E4

Die Untersuchung einer vermeintlich erwartungswidrigen Korrelation eines Strukturgleichungsmodells (siehe Abbildung 10.7) ergab, dass Tutoren mit einem ausgeprägten kontrollierenden Motivierstil so viel mehr geredet haben, dass sie absolut häufiger auf die Tutees eingegangen sind als die Tutoren, die sehr selten etwas gesagt haben, dann aber relativ häufiger auf ihre Tutees eingegangen sind (siehe Kapitel 11.1 und Anhang Kapitel 12.3.5). Entsprechend sollten die Tutoren grundsätzlich darauf trainiert werden, mehr mit den Tutees zu reden. Dabei sollten sie allerdings keine inhaltlichen Hinweise geben oder ins Experiment eingreifen. Vielmehr sollten die Tutoren ermutigend auf die Aussagen, Pläne und Handlungen der Tutees eingehen und somit ein Gefühl von Sicherheit für das Erproben eigener Ideen vermitteln. Dieser Änderungsvorschlag könnte auf zwei Arten in die Intervention integriert werden. Einerseits könnten an der Station der Betreuungsregel *Hilf nur so viel, dass dein Schüler selber weiterarbeiten kann*. zwei weitere prototypische Videosequenzen ergänzt werden. In einem der Videos würde man ein sehr zurückhaltendes Tutorverhalten sehen, welches man als Vernachlässigung verstehen könnte. Im anderen Video würde der Tutor häufig ermutigend auf die Aussagen, Pläne und Handlungen des Tutees eingehen und ihnen ein Gefühl von Sicherheit für das Erproben eigener Ideen vermitteln. Alternativ könnte auch die zweite Übungsphase (Rollenspiele mit Beobachtern, siehe Kapitel 8.3) überarbeitet werden. Nach den standardisierten Betreuungssituationen könnten einzelne Interventionsteilnehmende über einen Zeitraum von mehreren Minuten die Seminarleitung betreuen. Anschließend könnte im Plenum diskutiert werden, ob im Rollenspiel häufig genug auf die Tutees im Sinne der Kategorien E eingegangen wurde.

Inhaltliche Überarbeitung vorhandener Bausteine - Teil 2: Überarbeitung der Kategorien E1 bis E4

Mit dem Problem des letzten Absatzes verbunden scheint die Intervention bezüglich der Kategorien E1 bis E4 den größten Überarbeitungsbedarf zu haben. E3 als einzige tatsächlich autonomieunterstützende Verhaltensweise wurde absolut häufiger von Tutoren mit einem ausgeprägten kontrollierenden Motivierstil gezeigt (siehe letzter Absatz). Hinzu kommt, dass die Kategorien E1 und E2 sehr selten beobachtet wurden. Da die Tutoren fast alle Tuteeschwierigkeiten auch als solche erkannt haben (siehe Kapitel 9.1), wäre es vermutlich ausreichend, wenn die Tutoren bei aufgetretenen Schwierigkeiten konsequent ihre Tutees ermutigen und Verständnis zeigen würden. Weiterhin könnte die Intervention insofern überarbeitet werden, dass die Tutoren ihre Tutees regelmäßig ermutigen und auf sie eingehen, ohne dabei zwingend auf Tuteeschwierigkeiten zu warten. Beispielsweise könnte die Betreuungsregel *Gib Sicherheit und ermutige bei Schwierigkeiten* zu *Ermutigen, Verständnis zeigen und auf den Tutee eingehen* abgeändert werden. Dabei sollte betont werden, dass die Tutoren ihren Tutees bei erkannten Schwierigkeiten neben nötiger inhaltlicher Hilfe auch zusätzlich mit Aussagen der Kategorien E1, E2 und E3 motivational helfen. Dadurch könnte die absolute Anzahl an Ermutigungen und eingehenden Aussagen im Sinne von E1 und E3 steigen, was entsprechend den Befunden der vorliegenden Studie und Reeve und Jang (2006) zu einem erhöhten wahrgenommenen autonomieunterstützen-

den Motivierstil führen sollte. Auch Kategorie E4 ist selten vorgekommen. Die Pilotierung hat gezeigt, dass es für die Tutoren schwierig ist, ihre Tutees angemessen häufig zu loben, ohne dabei übertrieben und unaufrichtig zu wirken (siehe Kapitel 8.3). Entsprechend könnte die zweite Übungsphase (Rollenspiele mit Beobachtern) um eine Standardsituation ergänzt werden, in der ein Lob angebracht ist, um anhand dieses Beispiels die Angemessenheit und Ernsthaftigkeit des Lobs zu diskutieren.

Inhaltliche Überarbeitung vorhandener Bausteine - Teil 3: Detaillierteres Verständnis fördern, wann eine inhaltliche Hilfe wirklich nötig ist

Entsprechend den deskriptiven Videodaten erkennen Tutoren über 92% der Tuteeprobleme als solche. Neben diesen angemessenen Eingriffen haben sie über vier Mal so häufig unangemessen eingegriffen (siehe Kapitel 9.1). Somit müssen Tutoren lernen, wann eine inhaltliche Hilfe wirklich angemessen ist. Konkret ergab die Diskussion der Befunde zu Forschungsfrage 3, dass Tutoren ihren Tutees mehr als nötig inhaltlich helfen, wenn sie ihre Tutees als wenig selbstständig wahrnehmen (siehe Kapitel 11.3). Einerseits kann im Vortrag zur Einführung des Konzeptes der Autonomieunterstützung betont werden, dass die Tutees durch ungewolltes Eingreifen im Anschluss nicht selbstständiger werden (Deci & Ryan, 2002). Weiterhin rechtfertigt die wahrgenommene Unselbstständigkeit nicht das Verwenden von druckerzeugenden Modalverben und Imperativen. In der Phase der Rollenspiele mit Beobachtern können zwei Fälle gegenübergestellt werden, in denen die Tutees im gleichen Maße Hilfe benötigen aber unterschiedlich selbstständig erscheinen. Dieser Vergleich hilft den angehenden Tutoren einerseits, Fälle nötiger Hilfe zuverlässiger erkennen zu lernen. Andererseits kann der Vergleich die Grundlage einer Diskussion sein, bei der die Seminarleitung darauf hinweist, dass die wahrgenommene Unselbstständigkeit per se nicht mehr Hilfe nötig macht.

Weiterhin haben Tutoren mit einer ausgeprägten intrinsischen Motivation und einem ausgeprägten Fähigkeitsselbstkonzept bezüglich Physik möglicherweise ein stärkeres Bedürfnis zu helfen und laufen deshalb Gefahr, mehr zu helfen als nötig (siehe Kapitel 11.3). Zur Entgegenwirkung dieses Zusammenhangs bieten sich die Rollenspiele mit Beobachtern an. In dieser Phase können fachliche Ungenauigkeiten des Tutees eingebaut werden, welche keinen Eingriff nötig machen. Diese könnten gegen Überforderungen oder fachliche Fehler des Tutees abgegrenzt werden, die einen Eingriff des Tutors erfordern. Obwohl keine konkrete Notwendigkeit für Hilfe vorliegt, können Tutoren dem Drang zu helfen durch Ermutigungen (E1), gezeigtem Verständnis für Schwierigkeiten (E2) und durch Eingehen (E3) nachkommen. In der Reflexionsphase kann die Tendenz der Tutoren diskutiert werden, mehr helfen zu wollen, wenn sie besonders motiviert sind, ein ausgeprägtes Fähigkeitsselbstkonzept bezüglich Physik haben oder wenn sie ihre Tutees als weniger selbstständig wahrnehmen.

Inhaltliche Überarbeitung vorhandener Bausteine - Teil 4: Explizite Förderung der Zustimmung zu den Interventionsinhalten

Da die Zustimmung zu den Interventionsinhalten für alle Teilnehmenden relevant zu sein scheint (siehe Kapitel 11.2.3), könnten die Teilnehmenden am Ende jeder Station ihre Zustimmung zu den betrachteten autonomieunterstützenden und kontrollierenden Verhaltensweisen diskutieren. Die bereits vorhandene Reflexionsphase könnte um eine Diskussion über die Zustimmung zu den Interventionsinhalten ergänzt werden. So können möglicherweise Vorbehalte gegen Autonomieunterstützung abgebaut und die Zustimmung erhöht werden.

Hinzufügen weiterer Interventionsinhalte zur Vermeidung von Deckeneffekten

Ein weiterer Hinweis zur Weiterentwicklung der Intervention resultiert aus den Befunden bezüglich der vorherigen Bekanntheit der Interventionsinhalte. Entsprechend diesen Befunden scheint die Intervention eher für Teilnehmende mit einem schwach oder mäßig ausgeprägten autonomieunterstützenden Motivierstil geeignet zu sein (siehe Kapitel 11.2.2 und Anhang 12.4.4). Folglich sollte die Intervention auch hinsichtlich der Inhalte erweitert werden. Um eine einfache Zugänglichkeit für alle Teilnehmenden zu gewährleisten, wurden alle Inhalte so einfach wie möglich formuliert (siehe Kapitel 8.3). Um die Komplexität möglichst niedrig zu halten, wurden schwierige Inhalte ausgeschlossen und zwei der sechs Verhaltensgruppen autonomieunterstützenden Verhaltens weggelassen (siehe Kapitel 8.3). Eine Implikation der Befunde für die Weiterentwicklung der Intervention ist somit, die fehlenden Verhaltensgruppen ebenfalls als Inhalte anzubieten. Um die Intervention für andere Teilnehmende einfach zugänglich zu halten, könnten diese Inhalte zusätzlich für Teilnehmende angeboten werden, die aufgrund ihrer Vorkenntnisse die Stationen schneller bearbeiten konnten als die übrigen Gruppen. Diese Form der Binnendifferenzierung erleichtert auch die praktische Umsetzung der Intervention, da sich die betroffenen Teilnehmenden ohne diese Zusatzinhalte entweder langweilen oder sich mit interventionsfremden Inhalten beschäftigen. Konkret könnte die Intervention um zwei optionale Stationen und Betreuungsregeln erweitert werden, welche die fehlenden Verhaltensgruppen behandeln. Dies könnte entsprechend Deci et al. (1994) zu einem signifikant ausgeprägteren autonomieunterstützenden Motivierstil führen. Weiterhin könnten Tutoren mit einem ausgeprägten präinterventionellen autonomieunterstützenden Motivierstil eher das praktische Ausführen autonomieunterstützender Verhaltensweisen üben (Su & Reeve, 2011). Während dieser Zeit sollte die Intervention an die vermutlich vorhandenen Bedenken von Teilnehmenden anknüpfen, die einen ausgeprägten präinterventionellen kontrollierenden Motivierstil haben. Beispielsweise könnten in Diskussionen Gründe erörtert werden, weshalb sie kontrollierende Verhaltensweisen für motivations- und lernförderlicher halten. Weiterhin können sie sich beispielsweise auch Interviews von Tutoren ansehen, die kontrollierend betreut wurden und ihre Unzufriedenheit diesbezüglich ausdrücken (Reeve et al., 2018).

Ergänzend zu weiteren Inhalten zu Autonomieunterstützung könnten auch weitere lern- und motivationsförderliche Inhalte thematisiert werden. Eine gelungene Struktur hängt mit gelungener Autonomieunterstützung zusammen (Aelterman et al., 2018). Wissen über Struktur führt möglicherweise zu einem höheren Interventionseffekt (Cheon et al., 2020). Laut Cheon et al. (2020) sind das klare Kommunizieren der Erwartungen, freiwillige gestufte Lernhilfen und Feedback die drei Aspekte autonomieunterstützender Struktur. Konkret könnten die hinzuzunehmenden Inhalte an der Intervention von Cheon et al. (2020) orientiert werden. Weiterhin könnten die Tutoren auch explizit kompetenzunterstützendes Verhalten (Rocchi et al., 2017) lernen. Dieses kompetenzunterstützende Verhalten der Tutoren könnte mediiert über die Grundbedürfnisse erwünschte Zielvariablen wie die intrinsische Motivation oder den Lernerfolg beeinflussen (Jang et al., 2016; Jang et al., 2009). Da Su und Reeve (2011) die größten Effektstärken bei Interventionen im Umfang mehrerer Stunden finden konnten, ist unklar, ob eine Hinzunahme weiterer Konstrukte wie kompetenzunterstützendes Verhalten die Intervention überfrachtet.

Auffrischen der Interventionsinhalte im Anschluss an die Intervention

Die Relevanz der Zeit zwischen Intervention und Posterhebung liefert einen dritten Hinweis für die Weiterentwicklung der Intervention. Im Vergleich zu Erwachsenen scheinen die Motivierstile von Tutoren leichter veränderlich zu sein (siehe Kapitel 10.2.1 und 10.2.2). Dies kann zu den gefundenen großen Effektstärken führen, aber zur vergleichsweise schnel-

len Rückbildung der Motivierstile. Deshalb sollten die Interventionsinhalte im Anschluss an die Intervention regelmäßig aufgefrischt werden. Ob umfangreiche Treffen von mehreren Stunden nötig sind oder eine kurze Diskussion der Regeln im Umfang weniger Minuten ausreicht, müsste experimentell überprüft werden. In der Praxis könnte diese Auffrischung durch kurze Diskussionsrunden der Tutoren einer Schule mit der hauptverantwortlichen Lehrkraft oder durch Handouts realisiert werden, welche sie vor tatsächlichen Betreuungssituationen durchlesen können und somit an die zentralen Inhalte der Intervention erinnert werden.

11.6.2 Implikationen für künftige Forschung

Im Folgenden werden Implikationen für künftige Forschung dargelegt, welche sich aus den Befunden und Limitationen der vorliegenden Studie ergeben.

Absichern der Befunde

Aufgrund der unterschiedlichen Effektstärken und Signifikanzniveaus können einige Befunde als statistisch abgesichert angesehen werden (Beispielsweise die Relevanz der wahrgenommenen Selbstständigkeit für den beobachteten kontrollierenden Motivierstil, siehe Kapitel 11.3; Field, 2009; Kyriazos, 2018). Andere Befunde sind aufgrund der Nähe des Signifikanzniveaus zur Ausschlussgrenze von .05 und der geringen Stichprobengröße fragwürdig (beispielsweise die Signifikanzen der Fähigkeitsselbstkonzepte und der intrinsischen Motivation als Prädiktoren des beobachteten kontrollierenden Motivierstils, siehe Kapitel 11.3; Field, 2009; Kyriazos, 2018). Deswegen sollten die Befunde der vorliegenden Forschungsarbeit von künftigen Studien mit größeren Stichproben repliziert und damit abgesichert werden. Beispielsweise könnten diese Studien die widersprüchlichen Befunde des parametrischen und nicht parametrischen Tests bezüglich Forschungsfrage 4b aufklären (siehe Kapitel 11.4.2; Field, 2009).

Überarbeitung genutzter Skalen

Wenn künftige Studien die eingesetzten Instrumente adaptieren möchten, sollten einige der Skalen zuvor überarbeitet werden. Beispielsweise wurde das Item zur Erfassung der vorherigen Bekanntheit der Interventionsinhalte scheinbar nur bezüglich Autonomieunterstützung verstanden (siehe Kapitel 11.2.2). Um sicherzustellen, dass das Item tatsächlich falsch verstanden wurde, könnte man Interviews mit den Teilnehmenden durchführen, in denen sie erklären, was sie sich unter den Inhalten vorstellen und kodieren, ob sie eher autonomieunterstützende oder kontrollierende Inhalte nennen. Aufgrund des vermuteten Fokus dürften vermutlich wenige Interviews ausreichen, um ein signifikantes Ergebnis zu erhalten (Field, 2009). Bestätigt sich die Vermutung, können zukünftige Studien dieses Problem lösen, indem sie entweder konkreter auf die Inhalte eingehen oder die Begriffe Autonomieunterstützung und Kontrolle nutzen, da die Teilnehmenden die Begriffe während der Intervention kennenlernen. Weiterhin sollten künftige Studien mehrere Items zur Erfassung der vorherigen Bekanntheit nutzen, durch eine Hauptkomponentenanalyse die eindimensionale Struktur bestätigen und diese Hauptkomponente als die Bekanntheit in den Rechnungen berücksichtigen.

Weiterhin hatten die Skalen des Fachwissens sowie der introjizierten beziehungsweise externalen Regulation der extrinsischen Tutormotivation keine hinreichenden internen Konsistenzen (siehe Kapitel 9.2). Zur Überarbeitung können weitere ähnliche Items konstruiert und gemeinsam mit den Bestehenden pilotiert werden, um intern konsistente und valide

Skalen zu finden. Das Kodiermanual lieferte ebenfalls keinen intern konsistenten beobachteten autonomieunterstützenden Motivierstil. Alle Beobachtungs- und Auswertungskategorien weisen eine gute Urteilsübereinstimmung auf (siehe Kapitel 8.4.2). Verhalten entsprechend der Kategorie E3 konnte als autonomieunterstützend nachgewiesen werden (siehe Kapitel 10.3). E3 klärt allerdings nur ungefähr 11,6% des wahrgenommenen autonomieunterstützenden Motivierstils auf. Die übrigen 88,4% resultieren vermutlich aus Tuteeigenschaften und weiteren Tutorverhaltensweisen (Flunger et al., 2019). Die Tutoren haben die Tutees so selten ermutigt (E1) oder ihnen Verständnis für Schwierigkeiten gezeigt (E2), dass die Relevanz nicht zuverlässig geschätzt werden konnte. Neben diesen sozialen Aspekten der Autonomieunterstützung könnte durch eine Interviewstudie untersucht werden, weshalb die inhaltlichen Aspekte der Autonomieunterstützung (TW1_A und TW2_A) nicht signifikant mit dem wahrgenommenen autonomieunterstützenden Motivierstil zusammenhängen. In Kapitel 11.6.1 wurde die Empfehlung gegeben, dass die Intervention künftig um die zwei fehlenden Gruppen autonomieunterstützenden Verhaltens (siehe Kapitel 8.3) erweitert werden sollte. Werden all diese Aspekte in künftigen Studien berücksichtigt und die Erfassung der Verhaltensweisen entsprechend angepasst, könnte sich eine intern konsistente Skala des beobachteten autonomieunterstützenden Motivierstils ergeben, welche mit Hilfe des wahrgenommenen autonomieunterstützenden Motivierstils validiert werden könnte.

Die Auswertung der 131 Videos mit einer durchschnittlichen Länge von ungefähr zwanzig Minuten hat viele Ressourcen in Anspruch genommen. Somit wäre eine Reduktion des Zeitaufwandes wünschenswert. Einerseits könnten die Motivierstile wie bei Berger et al. (2017) auf Likert-Skalen für das ganze Video oder wie bei Kaur et al. (2015) pro fünf Minuten geratet werden. Alternativ könnten entsprechend dem *thin-slices-Ansatz* von Begrich et al. (2021) lediglich Ausschnitte der Videos ausgewertet werden. Es ist zu prüfen, ob die alternativen Verfahren zu ähnlichen Ergebnissen kommen und inwiefern die detaillierte Erfassung des Verhaltens nötig bleibt.

Berücksichtigung weiterer Konstrukte

Im Rahmen der Limitation der vorliegenden Studie wurden die kleinen bis mittleren aufgeklärten Varianzen diskutiert (siehe Kapitel 11.5). Möglicherweise wurden relevante Prädiktoren nicht berücksichtigt (Field, 2009). Haben zukünftige Studien hinreichend große Stichproben akquiriert, können sie folgende weitere Prädiktoren berücksichtigen. Wie in Kapitel 5.2 dargestellt, wurden in der vorliegenden Forschungsarbeit die Überzeugungen der Tutoren bezüglich der Effektivität und der Einfachheit der Umsetzung der Motivierstile nicht berücksichtigt. Reeve und Cheon (2016) konnten die Relevanz dieser Überzeugungen für die Motivierstile und deren Änderung nachweisen. Wird neben autonomieunterstützendem auch strukturunterstützendes Verhalten interveniert, sollten zukünftige Studien auch strukturunterstützendes und chaosförderndes Verhalten erfassen.

Wird bei den Tutees der wahrgenommene autonomieunterstützende Motivierstil erhoben, sollte auch der wahrgenommene kontrollierende Motivierstil erhoben werden. Dafür könnten künftige Studien beispielsweise den *Controlling Teacher Scale* (Jang et al., 2016) einsetzen. Wenn eines der psychologischen Grundbedürfnisse erhoben wird, sollten auch die übrigen zwei erhoben werden. Mindestens sollten sowohl das Kompetenz- als auch das Autonomieerleben erhoben werden. Dies wurde in der vorliegenden Forschungsarbeit bei der Beantwortung der vierten Forschungsfrage klar, da nicht entschieden werden konnte, ob der wahrgenommene autonomieunterstützende Motivierstil partiell oder vollständig mediiert über das Kompetenz- und Autonomieerleben auf die intrinsische Motivation wirkt (siehe Kapitel 11.4.1). Als Zielvariable auf Tuteebene wurde lediglich die intrinsische Moti-

vation berücksichtigt. Künftige Studien sollten sowohl Leistungsvariablen als auch affektive Variablen erfassen. Beispielsweise könnten der Zuwachs des Fachwissens der Tutees sowie geänderte Selbstkonzepte bezüglich des Fachs erfasst werden. Sollen die Tutees durch das Tutoring auf Prüfungen vorbereitet werden, würde sich ebenfalls die Erhebung negativ assoziierter Zielvariablen (z. B. Prüfungsangst, Arten extrinsischer Motivation oder Amotivation) anbieten. Werden negativ assoziierte Zielvariablen erhoben, sollte auch die Frustration der psychologischen Grundbedürfnisse erhoben werden (siehe Kapitel 5.1). Vermutlich wird die Befriedigung der psychologischen Grundbedürfnisse (wie in der vorliegenden Studie) positiv assoziierte Zielvariablen vorhersagen. Die Frustration der psychologischen Grundbedürfnisse wird vermutlich negativ assoziierte Zielvariablen vorhersagen, wie beispielsweise Lernschwierigkeiten (Jang et al., 2016).

Erfassen zukünftige Studien alle Konstrukte der vorliegenden Studie und die vorgeschlagenen weiteren Konstrukte, können die Befunde der vorliegenden Forschungsarbeit bestätigt werden. Weiterhin können vermutete Wirkzusammenhänge zwischen der Struktur und der Autonomieunterstützung beziehungsweise zwischen dem kontrollierenden und dem chaosunterstützenden Verhalten bestätigt werden (Aelterman et al., 2018). Aufgrund der erwarteten Universalität der Selbstbestimmungstheorie sollten die im Tutoring behandelten Themen keine Prädiktoren für das Tutorverhalten oder die Zielvariablen der Tutees sein (Ryan & Deci, 2017). Auf der Tuteeebene wird sich vermutlich ergeben, dass die positiv assoziierten Zielvariablen umso mehr gefördert wurden, je autonomieunterstützender und je strukturunterstützender sich die Tutoren verhalten haben. Dieser Effekt wird vermutlich vollständig über das Autonomie- und Kompetenzerleben der Tutees mediiert (siehe Kapitel 11.4). Betreuen die Tutoren mehr als einen Tutee, wird der Effekt möglicherweise auch über die Befriedigung der sozialen Eingebundenheit mediiert.

Ausweitung des Geltungsbereichs der Befunde

Die Tutees haben lediglich das in Kapitel 8.2 vorgestellte Experiment durchgeführt, bei der sie unterschiedliche Helligkeiten einer Glühlampe in Abhängigkeit der Dicke, der Länge und des Materials der Kabel zur Batterie untersuchen (Schwichow et al., 2015). Die Tutees haben in der Studie von Berger et al. (2017) vier Experimente zur Umwandlung verschiedener Energieformen zu elektrischer Energie durchgeführt, beispielsweise das Nutzen eines Kurbelgenerators, um eine Glühlampe zu betreiben. Zusammengenommen konnten sowohl Berger et al. (2017) als auch die vorliegende Studie zeigen, dass Autonomieunterstützung bei der Durchführung von Experimenten im Bereich Elektrizitätslehre erwünschte Zielvariablen positiv beeinflusst oder mit diesen zusammenhängt. Zukünftige Studien sollten die Befunde für andere Inhaltsgebiete der Physik (z. B. Optik), andere Naturwissenschaften und andere Tutor-Tutee-Kombinationen bestätigen, um gemeinsam mit der vorliegenden Forschungsarbeit Indizien für die universelle positive Relevanz von Autonomieunterstützung (Ryan & Deci, 2017) zu sammeln.

Ein konkreter Vorschlag für ein zukünftiges Forschungsvorhaben

Folgendes Beispiel veranschaulicht, wie eine Studie aussehen könnte, welche die oben genannten Empfehlungen berücksichtigt. In der Vorbereitungsphase der Studie werden die Intervention (siehe Kapitel 8.3) entsprechend den Empfehlungen aus Kapitel 11.6.1 und die Instrumente entsprechend Kapitel 11.6.2 überarbeitet. Da mittlere bis große Effekte erwartet werden können (siehe Kapitel 11.2.1 und 11.2.2), sollten ungefähr 150 Teilnehmende für die Intervention rekrutiert werden. Sollen Wirkmechanismen durch kovarianzbasierte Strukturgleichungsmodelle aufgedeckt und abgesichert werden, sollten mindestens 250 Probanden in der Interventionsgruppe sein (Hu & Bentler, 1999; Kyriazos, 2018). Als

Konstrukte sollten der autonomieunterstützende und kontrollierende Motivierstil, Selbstberichte zum strukturunterstützenden und chaosunterstützenden Verhalten (Aeltermann et al., 2018), die Kausalitätsorientierungen (Reeve et al., 2018), Überzeugungen zur Einfachheit der Implementation und der Effektivität dieser Ansätze (Reeve und Cheon, 2016), die Befriedigung der psychologischen Grundbedürfnisse während der Intervention (Ryan & Deci, 2017), die vorherige Bekanntheit der Interventionsinhalte (siehe Kapitel 11.2) sowie die Fähigkeitsselbstkonzepte bezüglich Autonomieunterstützung und Physik (siehe Kapitel 11.3) erfasst werden. Dabei sollten alle Konstrukte bis auf die vorherige Bekanntheit der Interventionsinhalte prä- sowie postinterventionell erfasst werden, um durch Strukturgleichungsmodelle die Veränderungen der einzelnen Konstrukte in Bezug setzen zu können. Zur Interventionsgruppe sollte eine Kontrollgruppe in vergleichbarer Größe rekrutiert werden. Dies können angehende Tutoren sein, die im Anschluss an die Studie an der Intervention teilnehmen dürfen. So kann dazu beigetragen werden, dass die Tutoren der Kontrollgruppe ähnliche Grundeinstellungen gegenüber der Tätigkeit als Tutoren haben. Die Intervention wird vermutlich den kontrollierenden Motivierstil und die Tendenz zu chaosunterstützendem Verhalten reduzieren sowie den autonomieunterstützenden Motivierstil und die Tendenz zu strukturunterstützendem Verhalten fördern (siehe Kapitel 11.1 sowie Aeltermann et al., 2018). Die Interventionseffekte bezüglich dieser vier Konstrukte werden vermutlich von den übrigen Konstrukten vorhergesagt. Der stärkste Prädiktor wird vermutlich die Überzeugung zur Effektivität dieser Ansätze sein (Reeve und Cheon, 2016).

11.6.3 Implikationen für die Schulpraxis

Die vorliegende Forschungsarbeit liefert vielfältige konkrete Implikationen. Beispielsweise geben die Befunde bezüglich Forschungsfrage 4b Hinweise darauf, dass das Tutorgeschlecht bei der Zuteilung zu den Tutees bezüglich der Tuteemotivation nicht berücksichtigt werden muss. Da negative Effekte gleichgeschlechtlicher Tutoren ausgeschlossen werden können, kann den Tutees im Sinne einer selbstgesteuerten Betreuung erlaubt werden, ihre Tutoren selbst auszuwählen. Steht das Erwerben von Fachwissen im Vordergrund, sollten Tutees einem gleichgeschlechtlichen Tutor zugewiesen werden (Leung, 2015).

Neben diesen konkreten Implikationen liefern die einzelnen Diskussionen der Forschungsfragen allgemeine Implikationen für die Schulpraxis. Da der wahrgenommene autonomieunterstützende Motivierstil sowohl das Kompetenzerleben als auch die intrinsische Motivation zu beeinflussen scheint (siehe Kapitel 11.4.1), sollten Tutoren ihre Tutees autonomieunterstützend betreuen. Dafür sollten Tutoren umfangreiche Hilfe erhalten, da es keine Selbstverständlichkeit ist, die von den Tutoren ohne Training verlangt werden kann (siehe Kapitel 11.2; Berger et al., 2017). Diese Hilfe könnte die Teilnahme an einer Intervention zur Vorbereitung auf ihre Tutorentätigkeit sein. Die Wirksamkeit der Intervention der vorliegenden Forschungsarbeit (siehe Kapitel 8.3) konnte einerseits durch die großen Effektstärken (siehe Kapitel 11.2.1) und den Zusammenhang zum Verhalten der Tutoren (siehe Kapitel 11.1) gezeigt werden. Aus den detaillierten Untersuchungen der Prädiktoren der Interventionseffekte (siehe Kapitel 11.2.2) und der Verhaltensweisen der Tutoren (siehe Kapitel 11.3) konnten konkrete Verbesserungsvorschläge für die Intervention herausgearbeitet werden (siehe Kapitel 11.6.1). Die diskutierten Probleme der Intervention können möglicherweise durch diese Verbesserungsvorschläge behoben werden. Die überarbeitete Intervention könnte eine hinreichende Hilfe und Vorbereitung für die Tutoren darstellen. Vergleicht man die hohe Befriedigung des Kompetenzerlebens und die hohe intrinsische Motivation der Tutees mit den mittelmäßigen Werten von Schülerinnen und Schülern im

Schulalltag (siehe Anhang Kapitel 12.4.7), zeigt sich eine Stärke von Tutoring-Angeboten an Schulen. Naturwissenschaftliche Tutoring-Angebote sind für die betreuten Schülerinnen und Schüler herausfordernd und motivierend, weshalb Schulen solche Angebote anbieten sollten. Aufbauend auf den Ergebnissen der vorliegenden Forschungsarbeit kann Schulen empfohlen werden, ihre Schülerinnen und Schüler mit der überarbeiteten Intervention zu Tutoren ausbilden zu lassen (siehe Kapitel 11.6.1).

Im Rahmen des Forschungsprojekts wurden knapp 200 Schülerinnen und Schüler zu Tutoren ausgebildet, die an ihren Schulen sehr unterschiedlich eingesetzt werden. Dies unterstreicht die vielfältigen Einsatzmöglichkeiten der Tutoren. Der häufigste Einsatz der ausgebildeten Tutoren ist die Planung, Vorbereitung und beaufsichtigte Durchführung von Experimentiernachmittagen/ Experimenten an Tagen der offenen Tür und Grundschulbesuchen, bei dem die Grundschülerinnen und Grundschüler einfache Experimente durchführen und durch die Tutoren betreut werden. An einigen Schulen stehen die Tutoren als Ansprechpartner für jüngere Schülerinnen und Schüler bei Aufgaben der Wettbewerbe *Jugend forscht* oder *freestyle-physics* zur Verfügung. Andere Tutoren betreuen eine Roboter-AG oder geben Nachhilfe und können dort die Ausbildungsinhalte nutzen. Da sich die Tutoren naturwissenschaftlich mit ihrer Tutorentätigkeit engagieren, können ihre Mühen beispielsweise für das MINT-EC Zertifikat angerechnet werden. Die Intervention (siehe Kapitel 8.3 und Kapitel 11.6.1) wird seit Abschluss des Forschungsprojekts im Alfred Krupp-Schülerlabor der Ruhr-Universität Bochum weiterhin angeboten. Somit stellt das Ausbildungskonzept nicht nur eine AG an den Schulen dar, sondern ein Konzept, das an vielen Stellen für die Ausgestaltung des Schulalltags gewinnbringend genutzt werden kann.

Literatur

- Aelterman, N., Vansteenkiste, M., Haerens, L., Soenens, B., Fontaine, J. R. J. & Reeve, J. (2018). Toward an integrative and fine-grained insight in motivating and demotivating teaching styles: The merits of a circumplex approach. *Journal of Educational Psychology*. <https://doi.org/10.1037/edu0000293>
- Aelterman, N., Vansteenkiste, M., van den Berghe, L., de Meyer, J. & Haerens, L. (2014). Fostering a need-supportive teaching style: Intervention effects on physical education teachers' beliefs and teaching behaviors. *Journal of sport & exercise psychology*, *36*(6), 595–609. <https://doi.org/10.1123/jsep.2013-0229>
- Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic & physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, *34*(5), 502–508. <https://doi.org/10.1111/opo.12131>
- Baard, P. P., Deci, E. L. & Ryan, R. M. (2004). Intrinsic need satisfaction: A motivational basis of performance and well-being in two work settings. *Journal of Applied Social Psychology*, *34*(10), 2045–2068. <https://doi.org/10.1111/j.1559-1816.2004.tb02690.x>
- Baron, R. M. & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Begrich, L., Kuger, S., Klieme, E. & Kunter, M. (2021). At a first glance – How reliable and valid is the thin slices technique to assess instructional quality? *Learning and Instruction*, *74*, 101466. <https://doi.org/10.1016/j.learninstruc.2021.101466>
- Berger, R., Müller, M. & Hänze, M. (2017). Konzeption und Evaluation von Tutor-Trainings zur Förderung der intrinsischen Motivation der Tutoren und der Autonomieförderung ihrer Tutees im Cross-age Tutoring. *Zeitschrift für Didaktik der Naturwissenschaften*, *23*(1), 225–239. <https://doi.org/10.1007/s40573-017-0066-3>
- Black, A. E. & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, *84*(6), 740–756. [https://doi.org/https://doi.org/10.1002/1098-237X\(200011\)84:6<740::AID-SCE4>3.0.CO;2-3](https://doi.org/https://doi.org/10.1002/1098-237X(200011)84:6<740::AID-SCE4>3.0.CO;2-3)
- Blanz, M. (2015). *Forschungsmethoden und Statistik für die Soziale Arbeit: Grundlagen und Anwendungen*. W. Kohlhammer Verlag.

- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation*. Springer. <https://doi.org/10.1007/978-3-540-33306-7>
- Buchs, C., Butera, F. & Mugny, G. (2004). Resource interdependence, student interactions and performance in cooperative learning. *Educational Psychology, 24*(3), 291–314. <https://doi.org/10.1080/0144341042000211661>
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. Pearson Studium.
- Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (Third edition). Routledge.
- Cheon, S. H. & Reeve, J. (2013). Do the benefits from autonomy-supportive PE teacher training programs endure?: A one-year follow-up investigation. *Psychology of Sport and Exercise, 14*(4), 508–518. <https://doi.org/10.1016/j.psychsport.2013.02.002>
- Cheon, S. H., Reeve, J. & Moon, I. S. (2012). Experimentally based, longitudinally designed, teacher-focused intervention to help physical education teachers be more autonomy supportive toward their students. *Journal of Sport and Exercise Psychology, 34*(3), 365–396. <https://doi.org/10.1123/jsep.34.3.365>
- Cheon, S. H., Reeve, J. & Vansteenkiste, M. (2020). When teachers learn how to provide classroom structure in an autonomy-supportive way: Benefits to teachers and their students. *Teaching and Teacher Education, 90*. <https://doi.org/10.1016/j.tate.2019.103004>
- Cheon, S. H., Reeve, J., Yu, T. H. & Jang, H. R. (2014). The teacher benefits from giving autonomy support during physical education instruction. *Journal of sport & exercise psychology, 36*(4), 331–346. <https://doi.org/10.1123/jsep.2013-0231>
- Christoph, S. (2014). *Erfassung experimenteller Kompetenzen mit Hands-On- und Paper-Pencil-Tests im Fach Physik* (Masterarbeit bei Prof. Dr. Hendrik Härtig).
- Clark, L. A. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment, 7*(3), 309–319. <https://doi.org/10.1037/1040-3590.7.3.309>
- Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20*(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Cohen, J. (1983). The Cost of Dichotomization. *Applied Psychological Measurement, 7*(3), 249–253. <https://doi.org/10.1177/014662168300700301>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge. <https://doi.org/10.4324/9780203771587>
- Cohen, P., Kulik, J. & Kulik, C.-L. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*(2), 237–248. <https://doi.org/10.3102/00028312019002237>
- Deci, E. L., Connell, J. P. & Ryan, R. M. (1989). Self-determination in a work organization. *The Journal of applied psychology, 74*(4), 580–590. <https://doi.org/10.1037/0021-9010.74.4.580>

- Deci, E. L., Eghrari, H., Patrick, B. C. & Leone, D. R. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, *62*(1), 119–142. <https://doi.org/10.1111/j.1467-6494.1994.tb00797.x>
- Deci, E. L. & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, *19*(2), 109–134. [https://doi.org/10.1016/0092-6566\(85\)90023-6](https://doi.org/10.1016/0092-6566(85)90023-6)
- Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. *Zeitschrift für Pädagogik*, *39*(2), 223–238. www.pedocs.de/volltexte/2017/11173/
- Deci, E. L. & Ryan, R. M. (2002). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 3–33). University of Rochester Press.
- Deci, E. L., Schwartz, A. J., Sheinman, L. & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, *73*(5), 642–650. <https://doi.org/10.1037/0022-0663.73.5.642>
- Deci, E. L., Spiegel, N. H., Ryan, R. M., Koestner, R. & Kauffman, M. (1982). Effects of performance standards on teaching styles: Behavior of controlling teachers. *Journal of Educational Psychology*, *74*(6), 852–859. <https://doi.org/10.1037/0022-0663.74.6.852>
- Dickhäuser, O., Schöne, C., Spinath, B. & Stiensmeier-Pelster, J. (2002). Die Skalen zum akademischen Selbstkonzept. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *23*(4), 393–405. <https://doi.org/10.1024//0170-1789.23.4.393>
- Eagly, A., Wood, W. & Diekmann, A. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 123–174). Erlbaum.
- Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)*. Sage.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378–382. <https://doi.org/10.1037/h0031619>
- Flunger, B., Mayer, A. & Umbach, N. (2019). Beneficial for some or for everyone? Exploring the effects of an autonomy-supportive intervention in the real-life classroom. *Journal of Educational Psychology*, *111*(2), 210–234. <https://doi.org/10.1037/edu0000284>
- Fogarty, J. L. & Wang, M. C. (1982). An Investigation of the Cross-Age Peer Tutoring Process: Some Implications for Instructional Design and Motivation. *The Elementary School Journal*, *82*(5), 451–469. <https://doi.org/10.1086/461281>
- Garbarino, J. (1975). The impact of anticipated reward upon cross-age tutoring. *Journal of Personality and Social Psychology*, *32*(3), 421–428. <https://doi.org/10.1037/h0077087>

- Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U. & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology, 48*, 67–84. <https://doi.org/10.1016/j.cedpsych.2016.09.003>
- George, D. & Mallery, P. (2010). *SPSS for Windows step by step: A simple guide and reference, 17.0 update*. Allyn & Bacon.
- Gollwitzer, P. M. (2014). Weakness of the will: Is a quick fix possible? *Motivation and Emotion, 38*(3), 305–322. <https://doi.org/10.1007/s11031-014-9416-3>
- Gorsuch, R. L. (2013). *Factor Analysis*. Taylor and Francis.
- Gottfried, A. E., Fleming, J. S. & Gottfried, A. W. (2001). Continuity of academic intrinsic motivation from childhood through late adolescence: A longitudinal study. *Journal of Educational Psychology, 93*(1), 3–13. <https://doi.org/10.1037/0022-0663.93.1.3>
- Gröhlich, C. & Guill, K. (2009). Wie stabil sind Bezugsgruppeneffekte der Grundschulpflichtempfehlung für die Schulformzugehörigkeit in der Sekundarstufe? <https://doi.org/10.25656/01:4560>
- Grolnick, W. S. & Apostoleris, N. H. (2002). What makes parents controlling? In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 161–181). University of Rochester Press.
- Haerens, L., Aelterman, N., Vansteenkiste, M., Soenens, B. & van Petegem, S. (2015). Do perceived autonomy-supportive and controlling teaching relate to physical education students' motivational experiences through unique pathways? Distinguishing between the bright and dark side of motivation. *Psychology of Sport and Exercise, 16*, 26–36. <https://doi.org/10.1016/j.psychsport.2014.08.013>
- Halvari, A. E. M. & Halvari, H. (2006). Motivational Predictors of Change in Oral Health: An Experimental Test of Self-Determination Theory. *Motivation and Emotion, 30*(4), 294–305. <https://doi.org/10.1007/s11031-006-9035-8>
- Hänze, M., Müller, M. & Berger, R. (2018). Cross-age tutoring: how to promote tutees' active knowledge-building. *Educational Psychology, 38*(7), 915–926. <https://doi.org/10.1080/01443410.2018.1444734>
- Hartmann, S. (2013). *Die Rolle von Leseverständnis und Lesegeschwindigkeit beim Zustandekommen der Leistungen in schriftlichen Tests zur Erfassung naturwissenschaftlicher Kompetenz* (Dissertation).
- Hayes, A. F. (2018). Partial, conditional, and moderated mediation: Quantification, inference, and interpretation. *Communication Monographs, 85*(1), 4–40. <https://doi.org/10.1080/03637751.2017.1352100>
- Hein, V., Koka, A. & Hagger, M. S. (2015). Relationships between perceived teachers' controlling behaviour, psychological need thwarting, anger and bullying behaviour in high-school students. *Journal of adolescence, 42*, 103–114. <https://doi.org/10.1016/j.adolescence.2015.04.003>

- Hofferber, N., Basten, M., Großmann, N. & Wilde, M. (2016). The effects of autonomy-supportive and controlling teaching behaviour in biology lessons with primary and secondary experiences on students' intrinsic motivation and flow-experience. *International Journal of Science Education*, 38(13), 2114–2132. <https://doi.org/10.1080/09500693.2016.1229074>
- Horton, R. L. (1978). *The general linear model: Data analysis in the social and behavioral sciences*. New York, McGraw-Hill.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling : a multidisciplinary journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Humphreys, L. G. (1978). Doing research the hard way: Substituting analysis of variance for a problem in correlational analysis. *Journal of Educational Psychology*, 70(6), 873–876. <https://doi.org/10.1037/0022-0663.70.6.873>
- Hutcheson, G. (1999). *The Multivariate Social Scientist: Introductory Statistics Using Generalized Linear Models*. Sage Publications. <https://doi.org/10.4135/9780857028075>
- IBM Corp. (2020). *IBM SPSS Statistics for Windows, Version 27.0*. Armonk, NY: IBM Corp.
- Jang, H., Kim, E. J. & Reeve, J. (2016b). Why students become more engaged or more disengaged during the semester: A self-determination theory dual-process model. *Learning and Instruction*, 43, 27–38. <https://doi.org/10.1016/j.learninstruc.2016.01.002>
- Jang, H., Reeve, J. & Halusic, M. (2016). A new autonomy-supportive way of teaching that increases conceptual learning: Teaching in students' preferred ways. *The Journal of Experimental Education*, 84(4), 686–701. <https://doi.org/10.1080/00220973.2015.1083522>
- Jang, H., Reeve, J., Ryan, R. M. & Kim, A. (2009). Can self-determination theory explain what underlies the productive, satisfying learning experiences of collectivistically oriented Korean students? *Journal of Educational Psychology*, 101(3), 644–661. <https://doi.org/10.1037/a0014241>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. <https://doi.org/10.1007/BF02291575>
- Katz, I. & Assor, A. (2007). When choice motivates and when it does not. *Educational Psychology Review*, 19(4), 429.
- Kaur, A., Hashim, R. A. & Noman, M. (2015). Teacher autonomy support intervention as a classroom practice in a Thai school. *Journal for Multicultural Education*, 9(1), 10–27. <https://doi.org/10.1108/JME-07-2014-0033>
- Klauer, K. J. & Leutner, D. (2012). *Lehren und Lernen: Einführung in die Instruktionspsychologie*. Beltz Verlagsgruppe.
- Kolenikov, S. & Bollen, K. A. (2012). Testing Negative Error Variances. *Sociological Methods & Research*, 41(1), 124–167. <https://doi.org/10.1177/0049124112442138>

- Korner, M. (2015). *Cross-Age Peer Tutoring in Physik: Evaluation einer Unterrichtsmethode* (Dissertation). Logos Verlag Berlin.
- Krapp, A. & Ryan, R. M. (2002). Selbstwirksamkeit und Lernmotivation. Eine kritische Betrachtung der Theorie von Bandura aus der Sicht der Selbstbestimmungstheorie und der pädagogisch-psychologischen Interessentheorie. In M. Jerusalem (Hrsg.), *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen* (S. 54–82). Weinheim: Beltz.
- Kulgemeyer, C. (2017). Der Einfluss universitär vermittelter professioneller Kompetenzen auf das Handeln in unterrichtlichen Erklärsituationen. In C. Maurer (Hrsg.), *Implementation fachdidaktischer Innovation im Spiegel von Forschung und Praxis* (S. 380–383). Gesellschaft für Didaktik der Chemie und Physik. https://www.gdcp-ev.de/wp-content/tagungsbaende/GDCP_Band37.pdf
- Kulgemeyer, C. & Riese, J. (2018). From professional knowledge to professional performance: The impact of CK and PCK on teaching quality in explaining situations. *Journal of Research in Science Teaching*, *55*(10), 1393–1418. <https://doi.org/10.1002/tea.21457>
- Kulik, C.-L. C. & Kulik, J. A. (1982). Effects of Ability Grouping on Secondary School Students: A Meta-analysis of Evaluation Findings. *American Educational Research Journal*, *19*(3), 415–428. <https://doi.org/10.3102/00028312019003415>
- Kyriazos, T. A. (2018). Applied Psychometrics: Writing-Up a Factor Analysis Construct Validation Study with Examples. *Psychology*, *09*(11), 2503–2530. <https://doi.org/10.4236/psych.2018.911144>
- Lagakos, S. W. (1988). Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable. *Statistics in medicine*, *7*(1-2), 257–274. <https://doi.org/10.1002/sim.4780070126>
- Landis, J. R. & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. <https://doi.org/10.2307/2529310>
- Leung, K. C. (2015). Preliminary empirical model of crucial determinants of best practice for peer tutoring on academic achievement. *Journal of Educational Psychology*, *107*(2), 558–579. <https://doi.org/10.1037/a0037698>
- Little, T. D. (2013). *The Oxford handbook of quantitative methods. Oxford library of psychology*. Oxford: Oxford University Press.
- Mahoney, J. W., Ntoumanis, N., Gucciardi, D. F., Mallett, C. J. & Stebbings, J. (2016). Implementing an Autonomy-Supportive Intervention to Develop Mental Toughness in Adolescent Rowers. *Journal of Applied Sport Psychology*, *28*(2), 199–215. <https://doi.org/10.1080/10413200.2015.1101030>
- Mangold. (2020). *INTERACT Benutzerhandbuch*. Mangold International GmbH. www.mangold-international.com

- Markland, D. & Hardy, L. (1997). On the factorial and construct validity of the intrinsic motivation inventory: conceptual and operational concerns. *Research quarterly for exercise and sport*, 68(1), 20–32. <https://doi.org/10.1080/02701367.1997.10608863>
- Mayring, P. (2010). Qualitative Inhaltsanalyse. In G. Mey & K. Mruck (Hrsg.), *Handbuch Qualitative Forschung in der Psychologie* (S. 601–613). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92052-8_42
- Meade, A. W., Johnson, E. C. & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *The Journal of applied psychology*, 93(3), 568–592. <https://doi.org/10.1037/0021-9010.93.3.568>
- Merton, R. K. (1948). The Self-Fulfilling Prophecy. *The Antioch Review*, 8(2), 193–210. <https://doi.org/10.2307/4609267>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Pelletier, L. G., Séguin-Lévesque, C. & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology*, 94(1), 186–196. <https://doi.org/10.1037//0022-0663.94.1.186>
- Pelletier, L. G. & Vallerand, R. J. (1996). Supervisors' beliefs and subordinates' intrinsic motivation: A behavioral confirmation analysis. *Journal of Personality and Social Psychology*, 71(2), 331.
- Pintrich, P. R., Marx, R. W. & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research*, 63(2), 167–199. <https://doi.org/10.3102/00346543063002167>
- Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227. <https://doi.org/10.1002/sci.3730660207>
- Praetorius, A.-K., Pauli, C., Reusser, K., Rakoczy, K. & Klieme, E. (2014). One lesson is all you need? Stability of instructional quality across lessons. *Learning and Instruction*, 31, 2–12. <https://doi.org/10.1016/j.learninstruc.2013.12.002>
- R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Reeve, J. (1998). Autonomy Support as an Interpersonal Motivating Style: Is It Teachable? *Contemporary Educational Psychology*, 23(3), 312–330. <https://doi.org/10.1006/ceps.1997.0975>
- Reeve, J. (2002). Self-determination theory applied to educational settings. In E. L. Deci & R. M. Ryan (Hrsg.), *Handbook of self-determination research* (S. 183–203). University of Rochester Press.

- Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist*, *44*(3), 159–175. <https://doi.org/10.1080/00461520903028990>
- Reeve, J., Bolt, E. & Cai, Y. (1999). Autonomy-supportive teachers: How they teach and motivate students. *Journal of Educational Psychology*, *91*(3), 537–548. <https://doi.org/10.1037//0022-0663.91.3.537>
- Reeve, J. & Cheon, S. H. (2016). Teachers become more autonomy supportive after they believe it is easy to do. *Psychology of Sport and Exercise*, *22*, 178–189. <https://doi.org/10.1016/j.psychsport.2015.08.001>
- Reeve, J., Jang, H.-R. & Jang, H. (2018). Personality-based antecedents of teachers' autonomy-supportive and controlling motivating styles. *Learning and Individual Differences*, *62*, 12–22. <https://doi.org/10.1016/j.lindif.2018.01.001>
- Reeve, J. & Jang, H. (2006). What teachers say and do to support students' autonomy during a learning activity. *Journal of Educational Psychology*, *98*(1), 209–218. <https://doi.org/10.1037/0022-0663.98.1.209>
- Reeve, J., Jang, H., Carrell, D., Jeon, S. & Barch, J. (2004). Enhancing students' engagement by increasing teachers' autonomy support. *Motivation and Emotion*, *28*(2), 147–169. <https://doi.org/10.1023/B:MOEM.0000032312.95499.6f>
- Reeve, J., Nix, G. & Hamm, D. (2003). Testing models of the experience of self-determination in intrinsic motivation and the conundrum of choice. *Journal of Educational Psychology*, *95*(2), 375–392. <https://doi.org/10.1037/0022-0663.95.2.375>
- Reeve, J. & Tseng, C.-M. (2011). Cortisol reactivity to a teacher's motivating style: the biology of being controlled versus supporting autonomy. *Motivation and Emotion*, *35*(1), 63–74. <https://doi.org/10.1007/s11031-011-9204-2>
- Reeve, J., Vansteenkiste, M., Assor, A., Ahmad, I., Cheon, S. H., Jang, H., Kaplan, H., Moss, J. D., Olausson, B. S. & Wang, C. K. J. (2014). The beliefs that underlie autonomy-supportive and controlling teaching: A multinational investigation. *Motivation and Emotion*, *38*(1), 93–110. <https://doi.org/10.1007/s11031-013-9367-0>
- Rheinberg, F., Vollmeyer, R. & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *Diagnostica*, *47*(2), 57–66. <https://doi.org/10.1026//0012-1924.47.2.57>
- Robinson, D. R., Schofield, J. W. & Steers-Wentzell, K. L. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review*, *17*(4), 327–362. <https://doi.org/10.1007/s10648-005-8137-2>
- Rocchi, M., Pelletier, L., Cheung, S., Baxter, D. & Beaudry, S. (2017). Assessing need-supportive and need-thwarting interpersonal behaviours: The Interpersonal Behaviours Questionnaire (IBQ). *Personality and Individual Differences*, *104*, 423–433. <https://doi.org/10.1016/j.paid.2016.08.034>

- Rocchi, M., Pelletier, L. & Lauren, C. (2013). Determinants of coach motivation and autonomy supportive coaching behaviours. *Psychology of Sport and Exercise*, *14*(6), 852–859. <https://doi.org/10.1016/j.psychsport.2013.07.002>
- Rosenthal, R. & Jacobson, L. (1968). Pygmalion in the classroom. *The Urban Review*, *3*(1), 16–20. <https://doi.org/10.1007/BF02322211>
- Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology*, *43*(3), 450–461. <https://doi.org/10.1037/0022-3514.43.3.450>
- Ryan, R. M. & Deci, E. L. (2003). On assimilating identities to the self: A self-determination theory perspective on internalization and integrity within cultures. In M. R. Leary & J. P. Tangney (Hrsg.), *Handbook of self and identity* (S. 253–272). The Guilford Press.
- Ryan, R. M. & Deci, E. L. (2017). *Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness*. Guilford Press. <https://doi.org/10.1521/978.14625/28806>
- Sarrazin, P. G., Tessier, D. P., Pelletier, L. G., Trouilloud, D. O. & Chanal, J. P. (2006). The effects of teachers' expectations about students' motivation on teachers' autonomy-supportive and controlling behaviors. *International Journal of Sport and Exercise Psychology*, *4*(3), 283–301. <https://doi.org/10.1080/1612197X.2006.9671799>
- Satorra, A. & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Hrsg.), *Latent variables analysis: Applications for developmental research* (S. 399–419). Sage Publications, Inc.
- Satorra, A. & Bentler, P. M. (2010). Ensuring Positiveness of the Scaled Difference Chi-square Test Statistic. *Psychometrika*, *75*(2), 243–248. <https://doi.org/10.1007/s11336-009-9135-y>
- Schlake, T., Fischer, H. E., Härtig, H. & Krabbe, H. (2020). Materialsammlung zum externen Seminar der Laborhelferausbildung. <https://doi.org/10.17185/dupublico/71125>
- Schwanzer, A. D., Trautwein, U., Lüdtke, O. & Sydow, H. (2005). Entwicklung eines Instruments zur Erfassung des Selbstkonzepts junger Erwachsener. *Diagnostica*, *51*(4), 183–194. <https://doi.org/10.1026/0012-1924.51.4.183>
- Schwichow, M., Christoph, S., Boone, W. J. & Härtig, H. (2016). The impact of sub-skills and item content on students' skills with regard to the control-of-variables strategy. *International Journal of Science Education*, *38*(2), 216–237. <https://doi.org/10.1080/09500693.2015.1137651>
- Schwichow, M., Christoph, S. & Härtig, H. (2015). Förderung der Variablen-Kontroll-Strategie im Physikunterricht. *MNU Journal*, *68*(6), 346–350.
- Seidel, T., Prenzel, M., Duit, R. & Lehrke, M. (2003). *Technischer Bericht zur Videostudie "Lehr-Lern-Prozesse im Physikunterricht"*. IPN.

- Shapiro, S. S. & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3/4), 591. <https://doi.org/10.2307/2333709>
- Stevens, J. P. (2001). *Applied Multivariate Statistics for the Social Sciences*. Psychology Press. <https://doi.org/10.4324/9781410604491>
- Su, Y.-L. & Reeve, J. (2011). A meta-analysis of the effectiveness of intervention programs designed to support autonomy. *Educational Psychology Review*, *23*(1), 159–188. <https://doi.org/10.1007/s10648-010-9142-7>
- Tessier, D., Sarrazin, P. & Ntoumanis, N. (2008). The effects of an experimental programme to support students' autonomy on the overt behaviours of physical education teachers. *European Journal of Psychology of Education*, *23*(3), 239–253. <https://doi.org/10.1007/BF03172998>
- Tessier, D., Sarrazin, P. & Ntoumanis, N. (2010). The effect of an intervention to improve newly qualified teachers' interpersonal style, students motivation and psychological need satisfaction in sport-based physical education. *Contemporary Educational Psychology*, *35*(4), 242–253. <https://doi.org/10.1016/j.cedpsych.2010.05.005>
- Thomas, A. E. & Müller, F. H. (2016). Entwicklung und Validierung der Skalen zur motivationalen Regulation beim Lernen. *Diagnostica*, *62*(2), 74–84. <https://doi.org/10.1026/0012-1924/a000137>
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L. & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software*, *59*(5), 1–38. <http://www.jstatsoft.org/v59/i05/>
- Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, *25*(6), 631–645. <https://doi.org/10.1080/01443410500345172>
- Topping, K. J., Peter, C., Stephen, P. & Whale, M. (2004). Cross-age peer tutoring of science in the primary school: influence on scientific language and thinking. *Educational Psychology*, *24*(1), 57–75. <https://doi.org/10.1080/0144341032000146449>
- Urdan, T. C. (2011). *Statistics in Plain English*. Routledge. <https://doi.org/10.4324/9780203851173>
- van den Berghe, L., Soenens, B., Vansteenkiste, M., Aelterman, N., Cardon, G., Tallir, I. B. & Haerens, L. (2013). Observed need-supportive and need-thwarting teaching behavior in physical education: Do teachers' motivational orientations matter? *Psychology of Sport and Exercise*, *14*(5), 650–661. <https://doi.org/10.1016/j.psychsport.2013.04.006>
- Vansteenkiste, M., Sierens, E., Goossens, L., Soenens, B., Dochy, F., Mouratidis, A., Aelterman, N., Haerens, L. & Beyers, W. (2012). Identifying configurations of perceived teacher autonomy support and structure: Associations with self-regulated learning, motivation and problem behavior. *Learning and Instruction*, *22*(6), 431–439. <https://doi.org/10.1016/j.learninstruc.2012.04.002>

-
- Westland, J. C. (2010). Lower bounds on sample size in structural equation modeling. *Electronic Commerce Research and Applications*, 9(6), 476–487. <https://doi.org/10.1016/j.elerap.2010.07.003>
- White, R. W. (1959). Motivation reconsidered: the concept of competence. *Psychological review*, 66, 297–333. <https://doi.org/10.1037/h0040934>
- Wigfield, A., Eccles, J. S., Fredricks, J. A., Simpkins, S., Roeser, R. W. & Schiefele, U. (2015). Development of Achievement Motivation and Engagement. In R. M. Lerner (Hrsg.), *Handbook of Child Psychology and Developmental Science* (S. 1–44). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118963418.childpsy316>
- Wilde, B., Kovaleva, U., Kovaleva, A. & Urhahne, D. (2009). Überprüfung einer Kurzsкала intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15, 31–45.
- Williams, G. C. & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determination theory. *Journal of Personality and Social Psychology*, 70(4), 767–779. <https://doi.org/10.1037/0022-3514.70.4.767>
- Wittich, P. (2017). *Stabilität des Tutoring Verhaltens im Physikunterricht* (Bachelorarbeit bei Prof. Dr. Hendrik Härtig).
- Wittwer, J. & Renkl, A. (2008). Why Instructional Explanations Often Do Not Work: A Framework for Understanding the Effectiveness of Instructional Explanations. *Educational Psychologist*, 43(1), 49–64. <https://doi.org/10.1080/00461520701756420>

Abbildungsverzeichnis

Abbildung 5.1	Vermutetes Wirkmodell der Motivierstile auf leistungsbezogene und affektive Konstrukte	18
Abbildung 5.2	Vermutetes Wirkmodell der Motivierstile auf die intrinsische Motivation von Tutees	19
Abbildung 5.3	Allgemeines Wirkmodell der Drücke auf die Motivierstile	25
Abbildung 5.4	Angepasstes Wirkmodell der Drücke auf die Motivierstile	27
Abbildung 8.1	Studienablauf	41
Abbildung 10.1	Mittelwerte der Motivierstile und Kausalitätsorientierung vor und nach der Intervention	85
Abbildung 10.2	Probandenanzahl in den Mediansplit-Gruppen	89
Abbildung 10.3	Modell 1; Erstes vermutetes Modell der Wirkbeziehung der Aspekte des Drucks von Innen und der Aspekte des Drucks von Unten (spontan) auf den beobachteten kontrollierenden Motivierstil.	91
Abbildung 10.4	Modell 2; Zweites vermutetes Modell der Wirkbeziehung der Aspekte des Drucks von Innen und der Aspekte des Drucks von Unten (spontan) auf den beobachteten kontrollierenden Motivierstil.	92
Abbildung 10.5	Modell 3; Drittes vermutetes Modell der Wirkbeziehung der Aspekte des Drucks von Innen und der Aspekte des Drucks von Unten (spontan) auf den beobachteten kontrollierenden Motivierstil.	92
Abbildung 10.6	Strukturgleichungsmodell 6; Prädiktoren des beobachteten kontrollierenden Motivierstils	95
Abbildung 10.7	Strukturgleichungsmodell 9; Prädiktoren des beobachteten autonomieunterstützenden/ kontrollierenden Motivierstils	97
Abbildung 10.8	Die vier untersuchten Strukturgleichungsmodelle der intrinsischen Motivation, des Kompetenzerlebens und der intrinsischen Motivation vor der Durchführung des Experiments	99
Abbildung 10.9	Strukturgleichungsmodell der intrinsischen Motivation, des Kompetenzerlebens und der intrinsischen Motivation vor der Durchführung des Experiments	101
Abbildung 10.10	Die vier untersuchten Strukturgleichungsmodelle der intrinsischen Motivation, des Kompetenzerlebens, der intrinsischen Motivation vor der Durchführung des Experiments und des wahrgenommenen autonomieunterstützenden Motivierstils	102
Abbildung 10.11	Strukturgleichungsmodell der intrinsischen Motivation, des Kompetenzerlebens, der intrinsischen Motivation vor der Durchführung des Experiments und des wahrgenommenen autonomieunterstützenden Motivierstils	104

Tabellenverzeichnis

Tabelle 4.1	19 Verhaltensweisen, die autonomieunterstützende oder kontrollierende Betreuende zeigen	16
Tabelle 5.1	Übersicht über mögliche Einflüsse auf die Motivierstile von Betreuenden	23
Tabelle 8.1	Erhobene Konstrukte	43
Tabelle 8.2	Die vorhandenen Kabel beim Betreuungsexperiment der Tutoring-Sitzung	44
Tabelle 8.3	Die zu kodierenden Verhaltensweisen	52
Tabelle 8.4	Kodierbeispiel	53
Tabelle 8.5	Die aus den Beobachtungskategorien resultierenden Auswertungskategorien	54
Tabelle 8.6	Angepasste Items des Learning Climate Questionnaires	58
Tabelle 8.7	Untersuchung der konvergenten Validität	64
Tabelle 8.8	Übersicht über die Anzahl der Tutoren und Tutees aufgeschlüsselt nach Schulnummer	66
Tabelle 9.1	Absolute und durchschnittliche Häufigkeiten der Beobachtungskategorien – Teil 1	68
Tabelle 9.2	Absolute und durchschnittliche Häufigkeiten der Beobachtungskategorien - Teil 2	69
Tabelle 9.3	Absolute und durchschnittliche Häufigkeiten der Auswertungskategorien	70
Tabelle 9.4	Hauptkomponentenanalyse der Items zur Erfassung der selbstberichteten Motivierstile	72
Tabelle 9.5	Hauptkomponentenanalyse der auf die auf Videolänge und relative Ausprägung normierten Verhaltensweisen zur Erfassung der beobachteten Motivierstile	72
Tabelle 9.6	Hauptkomponentenanalyse der auf die Anzahl der Events normierten Verhaltensweisen zur Erfassung der beobachteten Motivierstile	74
Tabelle 9.7	Hauptkomponentenanalyse der Items zur Erfassung des von den Tutees wahrgenommenen autonomieunterstützenden Motivierstils	74
Tabelle 9.8	Hauptkomponentenanalyse der Items zur Erfassung der intrinsischen Motivation und des Kompetenzerlebens	75
Tabelle 9.9	Übersicht über die internen Konsistenzen der aus den Hauptkomponentenanalysen resultierenden Skalen	78
Tabelle 9.10	Übersicht über die Mittelwerte der Konstrukte und Indizien für Abweichungen von der Normalverteilung	80
Tabelle 9.11	ANOVAs bezüglich der Unterschiede verschiedener Konstrukte zwischen den neun Schulen	82
Tabelle 9.12	Stichproben zur Untersuchung der Hypothesen	82
Tabelle 10.1	Pearson und Spearman Korrelationen der Perspektiven der Motivierstile	83

Tabelle 10.2	Mittelwertunterschiede der Motivierstile und Kausalitätsorientierungen vor und nach der Intervention sowie die statistische Signifikanz dieser Unterschiede	84
Tabelle 10.3	Regressionen der Änderungen der Motivierstile auf vermutete Prädiktoren	87
Tabelle 10.4	Teststatistiken der Regressionen der Änderungen der Motivierstile auf vermutete Prädiktoren	88
Tabelle 10.5	Regressionen der Änderungen des autonomieunterstützenden Motivierstils für Teilnehmende mit unterschiedlich ausgeprägter autonomer Kausalitätsorientierung	90
Tabelle 10.6	Regressionen der Änderungen des kontrollierenden Motivierstils für Teilnehmende mit unterschiedlich ausgeprägter kontrollierender Kausalitätsorientierung	90
Tabelle 10.7	Fitwerte der zu vergleichenden Modelle 1 - 3	93
Tabelle 10.8	Fitwerte der zu vergleichenden Modelle 3.1 - 3.9	93
Tabelle 10.9	Fitwerte der zu vergleichenden Modelle 3.1.0 und 3.1.1.	94
Tabelle 10.10	Fitwerte der zu vergleichenden Modelle 4 - 6.	94
Tabelle 10.11	Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).	95
Tabelle 10.12	Regressionsgewichte bezüglich des beobachteten kontrollierenden Motivierstils	96
Tabelle 10.13	Spearman-Korrelationen der auf die Videolänge und relative Ausprägung normierten vermutet autonomieunterstützenden Verhaltensweisen mit dem wahrgenommenen autonomieunterstützenden Motivierstil	96
Tabelle 10.14	Fitwerte der zu vergleichenden Modelle 7 - 9.	97
Tabelle 10.15	Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).	97
Tabelle 10.16	Regressionsgewichte bezüglich des beobachteten autonomieunterstützenden und kontrollierenden Motivierstils	98
Tabelle 10.17	Fitwerte der Modelle mit und ohne vollständig auf null fixierte korrelierte Messfehler.	100
Tabelle 10.18	Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).	101
Tabelle 10.19	Fitwerte der Modelle 1 - 4.	103
Tabelle 10.20	Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).	103
Tabelle 10.21	Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).	103
Tabelle 10.22	Modellvergleiche mit Hilfe eines skalierten Chi-Quadrat-Differenz-Tests nach Satorra und Bentler (2010).	104
Tabelle 10.23	Geschlechterverteilung der Tutoren und Tutees	105
Tabelle 10.24	T-Tests und Mann-Whitney-U-Tests für gleich- vs. getrenntgeschlechtliche Gruppen	105
Tabelle 11.1	Zusammenfassung über die Annahme und die Ablehnung der aufgestellten Hypothesen - Teil 1.	107
Tabelle 11.2	Zusammenfassung über die Annahme und die Ablehnung der aufgestellten Hypothesen - Teil 2.	108
Tabelle 11.3	Regressionsgewichte der signifikanten Prädiktoren der Änderungen des autonomieunterstützenden und des kontrollierenden Motivierstils	111

Tabelle 11.4	Regressionsgewichte der Zustimmung zu den Interventionsinhalten in den Regressionen bezüglich der Änderung des autonomieunterstützenden und des kontrollierenden Motivierstils aufgeschlüsselt für unterschiedlich ausgeprägte Kausalitätsorientierungen	114
Tabelle 11.5	Benötigte Stichproben für die berichteten Strukturgleichungsmodelle. .	121

Abkürzungsverzeichnis

Für einigen Indizes oder Tabelleneinträge sind die vollständig ausgeschrieben Konstrukt-namen teilweise zu lang. Deshalb werden im Folgenden Abkürzungen aufgelistet, sodass einzelne Indizes und Tabelleneinträge kompakter notiert und gelesen werden können:

Konstrukte	Abkürzung
Tutoren	
Selbstberichteter autonomieunterstützender Motivierstil (präinter- ventionell)	SMot_A_Prä
Selbstberichteter autonomieunterstützender Motivierstil (postinter- ventionell)	SMot_A_Post
Selbstberichteter autonomieunterstützender Motivierstil (prä-post- Differenz)	Δ SMot_A
Selbstberichteter kontrollierender Motivierstil (präinterventionell)	SMot_K_Prä
Selbstberichteter kontrollierender Motivierstil (postinterventionell)	SMot_K_Post
Selbstberichteter kontrollierender Motivierstil (prä-post-Differenz)	Δ SMot_K
Beobachteter autonomieunterstützender Motivierstil	BMot_A
Beobachteter kontrollierender Motivierstil	BMot_K
Autonome Kausalitätsorientierung (präinterventionell)	Kau_A_Prä
Autonome Kausalitätsorientierung (postinterventionell)	Kau_A_Post
Kontrollierende Kausalitätsorientierung (präinterventionell)	Kau_K_Prä
Kontrollierende Kausalitätsorientierung (postinterventionell)	Kau_K_Post
Zustimmung zu den Interventionsinhalten	Zustimmung
Vorherige Bekanntheit der Interventionsinhalte	Bekanntheit
Zeit zwischen Intervention und Messzeitpunkt 2	Zeit
Fähigkeitsselbstkonzept bezüglich Autonomieunterstützung	FSK_A
Fähigkeitsselbstkonzept bezüglich Physik	FSK_P
Selbstbestimmtheit der Betreuungsmotivation (Intrinsisch)	SRQ_intr
Selbstbestimmtheit der Betreuungsmotivation (Identifiziert)	SRQ_ident
Selbstbestimmtheit der Betreuungsmotivation (Introjiziert)	SRQ_introj
Selbstbestimmtheit der Betreuungsmotivation (extrinsisch)	SRQ_extrin
Verständnis der Variablen-Kontroll-Strategie	VKS
Fachwissen	FW
Wahrgenommenes Engagement (Druck von Unten)	TW_En
Wahrgenommene Selbstständigkeit (Druck von Unten)	TW_Se
Wahrgenommene intrinsische Motivation (Druck von Unten)	TW_IM
Tutees	
Intrinsische Motivation vor der Durchführung des Experiments	FAM
Wahrgenommener autonomieunterstützender Motivierstil	WMot_A
Intrinsische Motivation	IM
Kompetenzerleben	KE

Überblick über die Anhänge

Informationen zum primären Anhang

Das Anhangskapitel 12 ist der Printversion beiliegenden CD zu entnehmen bzw. steht zusätzlich zum Download zur Verfügung. Dem Kapitel können vertiefende Darlegungen der Testgütern, inhaltliche Rechnungen und Diskussionen entnommen werden.

Ergänzung des Anhangs

Die Ergänzung des Anhangs kann beim Autor der vorliegenden Arbeit angefragt werden. Die Ergänzung des Anhangs enthält:

- Interventionsmaterialien
 - Intervention Tutoren
 - * Hauptstudie: Dateien der Intervention, welche in Kapitel 8.3 beschrieben wurde. Die prototypischen Videosequenzen wurden durch Videotransskripte ersetzt.
 - * Testleitermanual einer Erhebung (Messzeitpunkte 1, siehe Kapitel 8.1; siehe Kapitel 8.4)
 - * Pilotierung: Dateien der Pilotierung der Intervention, welche in Kapitel 8.3 beschrieben wurde. Die prototypischen Videosequenzen wurden durch Videotransskripte ersetzt.
 - Intervention Tutees
 - * Ablauf der Erhebungen (Messzeitpunkte 2-4, siehe Kapitel 8.1)
 - * Testleitermanuale mehrerer Erhebungen (Messzeitpunkte 2-4, siehe Kapitel 8.1; siehe Kapitel 8.4)
 - * Materialien (Betreuerarbeitsblatt und Arbeitsblatt für Tutees)
- Daten und Skripte
 - Enthält die Hauptehebungsdaten und die genutzten Skripte. Weiterhin sind die Doppelkodierungsdaten zur Berechnung der Urteilsübereinstimmung der Kodierung des Tutorverhaltens und die Skripte zur Verarbeitung und Aufbereitung der Videodaten enthalten.
- Testhefte und Kodiermanual
 - Enthält, sortiert nach Messzeitpunkt (siehe Kapitel 8.1), alle Testhefte und das Kodiermanual, welches zur Erfassung der beobachteten Motivierstile genutzt wurde.

- Veröffentlichungen
 - „Autonomieunterstützendes Cross-Age-Peer-Tutoring beim Experimentieren“ (Tagungsbandbeitrag; Nicht peer-reviewed) [Link](#)
 - „Autonomieunterstützung beim Experimentieren im Cross-Age-Tutoring“ (Tagungsbandbeitrag; Nicht peer-reviewed) [Link](#)
 - „Materialsammlung zum externen Seminar der Laborhelferausbildung“ (Nicht peer-reviewed) [Link](#)