



universität
wien

DIPLOMARBEIT / DIPLOMA THESIS

Titel der Diplomarbeit / Title of the Diploma Thesis

„Entwicklung der Skala relatedness des IMI zur
Bestimmung der intrinsischen Motivation von
SchülerInnen“

verfasst von / submitted by

Julia Pieler

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2018 / Vienna, 2018

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on
the student record sheet:

A 190 406 412

Studienrichtung lt. Studienblatt /
degree programme as it appears on
the student record sheet:

Lehramtsstudium UF Mathematik UF Physik

Betreut von / Supervisor:

Univ.-Prof. Dr. Martin Hopf

Mitbetreut von / Co-Supervisor:

Mag. Dr. Marianne Korner

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich während der Erstellung dieser Arbeit und auf dem Weg durchs Studium unterstützt haben.

Zuallererst möchte ich mich bei Univ.-Prof. Dr. Martin Hopf und Mag. Dr. Marianne Korner für die Übernahme und Betreuung dieser Diplomarbeit bedanken.

Frau Mag. Dr. Marianne Korner nahm sich immer Zeit für meine Fragen und Anliegen und ermöglichte ein sehr angenehmes Arbeitsklima, vielen Dank für die gute Betreuung.

Ich möchte mich bei den Lehrkräften bedanken, die mir ihre Zeit zur Verfügung stellten und bei den Schülerinnen und Schülern, die konzentriert mitgearbeitet haben und den Fragebogen gewissenhaft ausfüllten.

Weiters möchte ich meinen Freunden Teresa Antal und Georg Richter danken, die sich für das Korrekturlesen dieser Arbeit Zeit nahmen und mir ein sehr hilfreiches Feedback gaben.

Ich möchte mich generell bei all meinen Freunden und Studienkollegen bedanken, die mir immer zur Seite standen und mich unterstützten. Das gemeinsame Rechnen und Aufgabenlösen hat mir viel Spaß gemacht.

Der größte Dank gebührt jedoch meinen Eltern, die mich all die Jahre finanziell und emotional unterstützten und an mich geglaubt haben. Ohne euch wäre ich nicht so weit gekommen, vielen Dank für alles.

Inhaltsverzeichnis

1. Einleitung	5
2. Theoretischer Hintergrund	7
2.1. Konstruktivistische Lerntheorien	7
2.1.1. Konstruktivistische Grundlagen	7
2.1.2. Conceptual Change	8
2.1.3. Conceptual Change im konstruktivistisch orientieren Unterricht	9
2.2. Lernmotivation	12
2.3. Selbstbestimmungstheorie der Motivation	19
2.4. Intrinsic Motivation Inventory (IMI)	23
2.5. Test- & Fragebogenkonstruktion	25
2.5.1. Hauptgütekriterien	25
2.5.2. Nebengütekriterien	30
2.5.3. Exploratorische Faktorenanalyse	31
2.5.4. Konfirmatorische Faktorenanalyse	32
2.6. Messinstrumente der Motivation	33
2.6.1. Bestehende Messinstrumente	33
2.6.2. Probleme mit dem Intrinsic Motivation Inventory	35
2.6.3. Entwicklung eines Messinstruments mit den Methoden von Wilson	36
3. Forschungsdesign	40
3.1. Forschungsfragen	40
3.2. Erstellung der Construct Map und der Items	41
3.2.1. Erste Version der Construct Map und der Items	41
3.2.2. Befragungen an Schulen	42
3.2.3. Zweite Version der Construct Map und der Items	43
3.2.4. Dritte Version der Construct Map und der Items	44
3.2.5. Finale Version der Items	49
3.3. Planung der Intervention	51
3.3.1. Didaktische Analyse	51
3.3.2. Stundenplanung	52
3.4. Erstellung der Fragebögen	54
3.5. Durchführung an den Schulen	55
3.6. Auswertung der Fragebögen	57

4. Ergebnisse	58
4.1. Analyse der Skala <i>relatedness</i>	58
4.1.1. KMO- und Bartlett-Test	58
4.1.2. Faktorenanalyse der Skala <i>relatedness</i>	59
4.1.3. Faktorenanalyse der Gruppenarbeit-Items	62
4.2. Einfluss der Intervention	64
4.3. Inhaltliche Analyse der Ergebnisse	66
4.3.1. Inhaltliche Analyse der Skala <i>relatedness</i>	66
4.3.2. Inhaltliche Analyse der Gruppenarbeit-Items	71
4.4. Schwierigkeiten und Anmerkungen zu den Items	72
4.5. Qualität des Fragebogens	75
5. Diskussion	77
5.1. Diskussion der neu erstellten Skala	77
5.2. Diskussion des Einflusses der Intervention	79
5.3. Inhaltliche Diskussion der Ergebnisse	80
6. Zusammenfassung und Ausblick	83
7. Literaturverzeichnis	85
8. Abbildungsverzeichnis	88
9. Tabellenverzeichnis	88
10. Anhänge A	89
10.1. Abstract in Deutsch	89
10.2. Abstract in Englisch	89
11. Anhänge B	91
11.1. Befragungen der Schülerinnen und Schüler	91
11.2. Arbeitsblatt	95
11.3. Fragebogen	98
11.3.1. Version 1	98
11.3.2. Version 2	100
11.4. Übersicht der Itemzuordnung im Fragebogen	102
11.5. Rotierte Komponentenmatrix der 30 Items	103

1) Einleitung

Im Alltag hört man oft Phrasen wie „Du bist aber motiviert.“ oder „Woher nimmst du nur die Motivation?“. Es scheint, dass manche Menschen motivierter sind als andere und das auch in Bezug auf gewisse Tätigkeiten, sei es in der Schule, im Studium, in der Arbeit oder sogar in der Freizeit. Nun stellt sich die Frage, warum manche Menschen eine große Motivation besitzen und andere hingegen zu überhaup nichts motiviert sind. Insbesondere für den schulischen Kontext wäre es für Lehrkräfte vorteilhaft zu wissen, wie sich die Motivation von Schülerinnen und Schülern steigern lässt und was diese negativ beeinflusst. Doch um diese Fragen zu beantworten, muss man sich mit dem Begriff „Motivation“ näher auseinandersetzen.

Deci und Ryan beschreiben Motivation folgendermaßen: „*To be motivated means to be moved to do something.*“ (Ryan & Deci, 2000, S.1). Aus diesem Satz lässt sich entnehmen, dass Motivation etwas mit einem „inneren Wollen“ zu tun hat. Dieses innere Wollen stellt die ausgezeichnetste Form der Motivation, die intrinsische Motivation, dar. Dabei führen Personen ihrer selbst willen eine Tätigkeit aus, ohne den Einfluss jeglicher äußerer Faktoren (Schlag, 2013). Das Handeln weist ein hohes Maß an Selbstbestimmung auf, oft wird bereits die Tätigkeit selbst als erfüllend wahrgenommen (Ryan & Deci, 2000).

Eine weitere Form der Motivation stellt die extrinsische dar. Personen, die extrinsisch motiviert sind, handeln aufgrund von äußeren Anreizen wie Strafandrohung, Belohnung, um Bezugspersonen eine Freude zu bereiten oder aufgrund von innerem Druck (Deci & Ryan, 1993). Um die Motivation von Schülerinnen und Schülern zu steigern, sollte die Fokussierung auf der Steigerung der intrinsischen Motivation liegen, was diese beeinflusst, wird in verschiedenen Motivationstheorien diskutiert. Heutzutage gibt es eine Vielzahl an Theorien, welche zu erklären versuchen, warum uns gewisse Ziele oder Aktivitäten stärker motivieren als andere und warum der Grad der Motivation von Person zu Person so verschieden ist.

Zur Evaluierung der Motivation von Schülerinnen und Schülern wurden bereits Fragebögen und Instrumente entwickelt. Die meisten dieser Instrumente sind jedoch sehr spezifisch oder weisen gewisse Mängel auf (Korner, 2015). Eines der bekanntesten Werkzeuge zur Ermittlung der Motivation ist das Intrinsic Motivation Inventory (IMI), welches von Deci und Ryan (2003) entwickelt wurde. Dieser Fragebogen ist in englischer Sprache verfasst und dient zur Evaluierung der intrinsischen Motivation von Personen. Im Rahmen ihrer „Cross-Age Peer Tutoring in Physik“ Studie versuchte Korner (2015) das IMI ins Deutsche zu übersetzen.

Schnell zeigten sich Schwierigkeiten auf, einerseits mit der Verständlichkeit der Items und andererseits mit dem Konstrukt selbst. Die Lernenden hatten Probleme mit der Beantwortung negativ gepolter Items, sowie mit der doppelten Verneinung (Korner, 2015). Da das IMI eigentlich für Teamsport entwickelt wurde, konnte es nicht einfach so in den schulischen Kontext übertragen werden. Aufgrund ihrer Erkenntnisse entschied sich Korner die Skalen des IMI neu zu entwickeln. Dazu verwendete sie die Methoden von Wilson (2005) und erstellte die Skala *effort/importance* komplett neu. Im Zuge ihrer Diplomarbeit entwickelten Klingenböck (2016) die Skala *interest/enjoyment* neu und Schmidt (2017) die Skalen *perceived competence*, *perceived choice* und *value/usefulness*. Es stellte sich heraus, dass diese Vorgangsweise gute Ergebnisse liefert und so ein Instrument zur Ermittlung der Motivation von Schülerinnen und Schülern entwickelt werden kann.

Die letzte Skala, *relatedness*, welche sich mit der sozialen Eingebundenheit der Lernenden befasst, wurde im Rahmen dieser Diplomarbeit entwickelt. Dazu wurden Items mithilfe des Four Building Blocks Approach von Wilson (2005) erstellt, welcher in Kapitel 2.6. genauer beschrieben wird. Die Skala wurde in Klassen der siebten und achten Schulstufe getestet. Dazu wurde in der Hälfte der Klassen eine Intervention zum Thema Magnetismus durchgeführt mit anschließendem Ausfüllen des Fragebogens, in der anderen Hälfte wurde nur der Fragebogen im regulären Unterrichtsverlauf ausgefüllt.

Ziel dieser Arbeit ist es, einen Beitrag zur Erstellung eines Instruments zur Evaluierung der Motivation von Schülerinnen und Schüler zu leisten, indem die letzte verbleibende Skala des IMI neu entwickelt wird. Weiters wird untersucht, ob die Antworten der Schülerinnen und Schüler unabhängig von einer zuvor durchgeführten Intervention sind.

Der fertige Fragebogen, bestehend aus den sieben Skalen, soll letztendlich das Konstrukt Motivation gut abbilden können, womit sich für Fachdidaktikerinnen und Fachdidaktiker weitere Forschungsmöglichkeiten ergeben. Der Fragebogen kann außerdem von Lehrkräften aller Fächer verwendet werden, um sich einen Überblick über die aktuelle Motivation der Lernenden zu verschaffen.

2) Theoretischer Hintergrund

In diesem Kapitel wird ein Überblick über die konstruktivistische Lerntheorie, die verschiedenen Motivationstheorien, Messinstrumente der Motivation sowie testtheoretischen Grundlagen und empirischen Methoden zur Überprüfung eines Testentwurfs gegeben.

2.1. Konstruktivistische Lerntheorien

Erste Gedanken zum Konstruktivismus gab es bereits vor Christus. Der griechische Philosoph Xenophanes sagte, dass *„kein Mensch je etwas Genaues erfassen werde, denn, was man sehe, sei immer nur Anschein“* (von Glasersfeld, 1997, p. 9). Dieser Aussage lässt sich entnehmen, dass neue Erkenntnisse auf unseren Erfahrungen beziehungsweise Wahrnehmungen beruhen und nicht unbedingt auf der Realität selbst (von Glasersfeld, 1997).

Interessant ist, wie neues Wissen überhaupt zustande kommt. Es stellt sich heraus, dass der Wissenserwerb ein sehr komplexer Vorgang ist. Hopf et al. schildern diese Komplexität auf neurophysiologischer Ebene folgendermaßen: Das Arbeitsgedächtnis verarbeitet akustische und visuelle Informationen *„mit Hilfe aktivierter kognitiver Strukturen aus dem Langzeitgedächtnis“* (Hopf, Wiesner, & Schecker, 2011, p. 31). Durch die Verknüpfung der neuen Informationen mit den kognitiven Strukturen kann eine Person den Informationen eine Bedeutung geben. Es wird vermutet, dass das Arbeitsgedächtnis ungefähr sieben Elemente, die für eine Person neu sind, gleichzeitig verarbeiten kann. Können die neuen Informationen mit denen im Langzeitgedächtnis verknüpft und gespeichert werden, hat die Person etwas Neues dazugelernt. Sind jedoch keine passenden Strukturen im Langzeitgedächtnis vorhanden um die Information zu verarbeiten, so kann die Person mit der neuen Information sozusagen nichts anfangen, sie nicht einordnen und ihr keine sinnvolle Bedeutung geben (Hopf et al., 2011). Aufgrund der Komplexität wurden Studien zum Lernen durchgeführt und verschiedene Lerntheorien entwickelt. Eine dieser ist die konstruktivistische Lerntheorie, welche im Folgenden beschrieben wird.

2.1.1. Konstruktivistische Grundlagen

Die Hauptaussage konstruktivistischer Lerntheorien ist, dass Wissen vom Lernenden auf Basis seiner Erfahrungen und Vorstellungen selbst konstruiert und nicht schlicht empfangen wird (Duit, 2010). Die Lernenden stehen im Zentrum der Lernhandlung, das Lernen selbst wird als *Konzeptwechsel* beziehungsweise *Konzeptentwicklung* aufgefasst (Korner, 2015). Wenn Schülerinnen und Schüler in den Unterricht kommen, haben sie bereits aus der Alltagserfahrung Vorstellungen zu Begriffen, Phänomenen und Prinzipien entwickelt.

Oft stimmen die Vorstellungen der Lernenden jedoch nicht mit den wissenschaftlichen Vorstellungen überein, wodurch Lernschwierigkeiten und Verständnisprobleme entstehen (Duit, 2010). Lehrpersonen machen immer wieder die Erfahrung, dass Lernende die klar kommunizierten Inhalte nicht verstehen (Korner, 2015). Da beim Lernen Wissen auf Basis der vorhandenen Vorstellungen aktiv konstruiert wird, reicht das Vermitteln von Informationen nicht aus, um den Lernprozess in Gang zu setzen. Knüpft der Unterricht nicht an das Wissen der Lernenden an, so ist es wahrscheinlich, dass diese gar nicht verstehen was die Lehrperson erklärt, oder was sie im Lehrbuch lesen (Duit, 2010).

Die Lehrperson kann somit nicht einfach ihr Wissen im Sinne des Nürnberger Trichters an die Schülerinnen und Schüler weitergeben (Korner, 2015). Dies ist jedoch auch heutzutage noch eine gängige Sichtweise. Das einfache Weiterreichen von Wissen ist deshalb nicht möglich, da die Sinnesdaten, die der Lernende empfängt, für ihn keine innewohnende Bedeutung haben. Sie erhalten erst eine Bedeutung, wenn der Lernende ihnen bewusst oder unbewusst eine Bedeutung verleiht (Duit, 2010). Wir Menschen können die Wirklichkeit an sich nicht erfassen, wir können uns lediglich ein Bild von ihr machen. Das Wahrgenommene ist nichts anderes als die Interpretation unseres Gehirns auf Basis des bereits Bekannten (von Glasersfeld, 1997). An dieser Stelle wird klar, dass Lernende den Unterrichtsinhalten oft eine andere Bedeutung verleihen als die Lehrperson. Die Lehrperson hat natürlich andere und viel komplexere Vorstellungen über ein Thema als die Lernenden. Haben die Lernenden jedoch noch nicht die Vorstellungen, die es braucht um einen gewissen Inhalt zu verstehen, so können sie dem Inhalt nicht die richtige Bedeutung verleihen. Dies ist nicht nur in der Schule beobachtbar, auch im Alltag treten immer wieder Situationen auf, in denen Personen aneinander einfach nicht verstehen.

Die konstruktivistischen Lerntheorien verdeutlichen, wie wichtig es ist, Vorerfahrungen der Lernenden zu berücksichtigen (Duit, 2010). Es braucht eine neue Herangehensweise, die das Konzept des Konstruktivismus berücksichtigt, um Wissen optimal zu übermitteln.

2.1.2. Conceptual change

Beim *Conceptual change* (Konzeptwechsel) sollen die vorunterrichtlichen Vorstellungen der Lernenden in wissenschaftlich anschlussfähige Vorstellungen überführt werden (Korner, 2015). Insbesondere in der Physik sollen die Lernenden von ihren Alltagsvorstellungen weg, hin zur physikalischen Sichtweise geführt werden. Die Frage, die sich nun stellt, ist, unter welchen Bedingungen kann der Konzeptwechsel stattfinden?

Studien haben gezeigt, dass es nicht gelingt die Lernendenvorstellungen gänzlich zu eliminieren, sie können nur Schritt für Schritt hin zur wissenschaftlichen Sichtweise geführt werden.

Es braucht somit mehrere Konzeptwechsel, damit die Lernenden der wissenschaftlich Sichtweise näher kommen (Duit, 2010).

In der Physik hat sich gezeigt, dass Schülerinnen und Schüler nach einem gut angeleiteten Konzeptwechsel die physikalische Sicht zwar verstehen, allerdings glauben sie trotzdem nicht, dass diese wahr ist. Beispielsweise können Lernende die physikalischen Grundlagen verstehen, warum ein beleuchteter Körper Licht aussendet, allerdings wollen sie diese Tatsache nicht wahrhaben. Aus solchen Situationen lässt sich schließen, dass beim Lernen immer affektive Aspekte eine Rolle spielen.

Um Lernendenvorstellungen zur wissenschaftlichen Vorstellung überzuführen, gibt es folgende drei Strategien: anknüpfen, umdeuten und konfrontieren. Anknüpfen und Umdeuten sind kontinuierliche Wege. Beim Anknüpfen versucht man an das Alltagsverständnis der Lernenden anzuknüpfen, wobei dieses so nah am physikalischen Verständnis wie möglich sein sollte (Duit, Treagust, & Widodo, 2008). Beim Umdeuten macht man den Lernenden klar, dass ihre Sichtweise nicht ganz falsch ist und dass sie durch eine Umdeutung ihrer Vorstellungen zu einer physikalisch sinnvollen Bedeutung kommen. Beispielsweise gibt es Lernende, die glauben, dass elektrischer Strom verbraucht wird, diese Vorstellung kann in die Umwandlung von Energie umgedeutet werden (Hopf et al., 2011).

Konfrontieren ist ein diskontinuierlicher Weg, bei dem es bei den Lernenden zu einem *kognitiven Konflikt* kommt. Die Vorstellung der Schülerinnen und Schüler wird der physikalischen Tatsache gegenübergestellt, zum Beispiel, indem ein Experiment anders ausgeht, als man glaubt. Es hat sich jedoch gezeigt, dass dieser Konflikt oft nicht ausreicht, um die Lernenden von ihren teils falschen Vorstellungen wegzuführen, oft verstehen sie den Konflikt schlichtweg nicht (Duit, 2010).

Damit ein erfolgreicher Konzeptwechsel durchgeführt werden kann, muss der Lernende mit seinen bisherigen Vorstellungen unzufrieden sein, man muss ihm ein alternatives Konzept anbieten, welches verständlich, plausibel und ertragreich ist. Die neue Vorstellung soll dem Lernenden helfen andere und neue Problemstellungen zu lösen. Erst dann kann eine Akkommodation der Vorstellungen erfolgen (Posner, Strike, Hewson, & Gertzog, 1982).

2.1.3. Conceptual Change im konstruktivistisch orientieren Unterricht

Die Qualität des Unterrichtens kann mithilfe des *Conceptual Change* Ansatzes verbessert werden, da bei dem Ansatz die Lernendenvorstellungen berücksichtigt werden. Dadurch können die Schülerinnen und Schüler schrittweise zu einem besseren Verständnis der Unterrichtsinhalte hingeführt werden.

In diesem Kapitel soll geklärt werden, wie Lernumgebungen aussehen müssen, um die eigenständige Konstruktion von Wissen zu ermöglichen (Widodo & Duit, 2004).

Widodo und Duit beschreiben fünf Kennzeichen konstruktivistischer Lernumgebungen (Widodo & Duit, 2004, p. 237):

- (1) *„Konstruktion des Wissens ermöglichen*
- (2) *Relevanz und Bedeutung der Lernerfahrungen*
- (3) *Soziale Interaktion*
- (4) *Unterstützung der Schüler beim eigenständigen Lernen*
- (5) *Wissenschaft, Wissenschaftler und wissenschaftliches Wissen“*

Die Kategorie *Konstruktion des Wissens ermöglichen* erfasst in welchem Maße und auf welche Weise an die bereits bestehenden Vorstellungen der Lernenden angeknüpft wird. Besonders wichtig ist es, den Lernenden ihren aktuellen Wissenstand bewusst zu machen, ihnen mitzuteilen, wofür die zu lernenden Unterrichtsinhalte verwendet werden können und eine Vorschau und einen Rückblick auf diese zu geben. Dadurch sollen die Unterrichtsinhalte miteinander vernetzt werden und nicht wie bisher einzelne Bruchstücke gelernt werden, bei denen die Lernenden keinen Zusammenhang erkennen. Um an das Vorwissen der Lernenden anknüpfen zu können, muss man zuerst in Erfahrung bringen, was sie bereits wissen und welche Vorstellungen sie haben. Dies sollte auf eine zeitökonomische Weise erfolgen. Weiters ist es empfehlenswert, sich mit den Denkweisen der Lernenden auseinanderzusetzen, dadurch kann man geeignete Aktivitäten, Unterrichtsmethoden, Experimente auswählen und angemessene Antworten geben. Besonders wichtig ist es den Schülerinnen und Schülern herausfordernde Aufgaben oder Problemstellungen zu geben, die zum Denken anregen und dieses fördern (Widodo & Duit, 2004).

In dem Punkt *Relevanz und Bedeutung der Lernerfahrungen* geht es darum, dass die Lernbedürfnisse der Lernenden berücksichtigt werden, um ihnen eine optimale Lernerfahrung zu ermöglichen. Dazu werden die neuen Inhalte in einen *„sozialen und materialen Kontext eingebettet“* (Widodo & Duit, 2004, p. 239). Finden die Lernenden eine persönliche Bedeutung für sich, so wird eher eine Bereitschaft da sein, die Inhalte zu lernen. Um eine persönliche Relevanz herzustellen sind Vorgänge, Phänomene, Beispiele oder Materialien aus dem Alltag sinnvoll. Es empfiehlt sich mit den Lernenden über Anwendungsmöglichkeiten des Gelernten zu diskutieren, damit das neue Wissen in einen Kontext eingebettet werden kann und somit länger zur Verfügung steht (Widodo & Duit, 2004).

Der Kategorie *soziale Interaktion* liegt der Fakt zugrunde, dass Wissen im sozialen Umfeld konstruiert wird. Die Lernenden sollten die Möglichkeit bekommen, mit anderen Lernenden oder der Lehrperson in verschiedenen Sozialformen zu interagieren. Dabei stellt die Kommunikation zwischen den Schülerinnen und Schülern und auch mit der Lehrperson eine wesentliche Rolle dar. Denn durch die Kommunikation miteinander wird das Verstehen anderer Standpunkte gefördert und damit die Reflexion über die eigenen Vorstellungen (Widodo & Duit, 2004).

Unterstützung der Schüler beim eigenständigen Lernen geht von der Annahme aus, dass mit dem Lernen eine Intention verfolgt wird und dass der Lernende für das Lernen selbst verantwortlich ist. Hier stellt sich die Frage, inwieweit man den Lernenden die Verantwortung für das Lernen geben kann und bis zu welchem Grade sie den Lernprozess mitbestimmen können. Eigenständiges Lernen kann beispielsweise durch Freiräume beim Lernen initiiert werden, aber auch dadurch, dass man die Schülerinnen und Schüler ermutigt sich mit ihren eigenen Ideen zu befassen, diese zu überdenken und zu analysieren. Die Lernenden sollen angeleitet werden über ihr Lernen zu reflektieren und es zu kontrollieren (Widodo & Duit, 2004).

Der letzten Kategorie, *Wissenschaft, Wissenschaftler und wissenschaftliches Denken*, auch *Nature of Science* genannt, liegt der Fakt zugrunde, dass Wissen vom Menschen konstruiert wird und auch falsch oder unvollständig sein kann. Die Schülerinnen und Schüler sollen erfahren wie Fachwissen konstruiert wird und dass dieses immer weiterentwickelt und manchmal auch revidiert wird. Erkenntnisse können sich immer wieder ändern. Außerdem wird Wissen von subjektiven Werten und Ansichten, den eigenen Interessen sowie von Erfahrungen beeinflusst. Es soll weiters eine Einsicht in die naturwissenschaftliche Denk- und Arbeitsweise gegeben werden, sowie die Rolle von Beobachtungen, Hypothesen, Theorien und Gesetzen verstanden werden. Die Lernenden sollen unterschiedliche wissenschaftliche Forschungsstrategien und die Grenzen von Theorien kennen lernen (Widodo & Duit, 2004).

Will man als Lehrperson einen erfolgreichen Konzeptwechsel durchführen, so sind die folgenden Schritte in der Strukturierung des Lernprozesses hilfreich (Duit, 2010, p. 4):

- „*Vertraut machen mit den Phänomenen*
- *Bewusstmachen der Schülervorstellungen*
- *Einführung in die physikalische Sichtweise*
- *Anwendung der neuen Sichtweise*
- *Rückblick auf den Lernprozess*“

Im Unterricht wird häufig auf die Schritte *Bewusstmachen der Schülervorstellungen* und *Rückblick auf den Lernprozess* verzichtet, welche jedoch von großer Bedeutung sind. Um die Unterrichtsinhalte optimal zu vermitteln, muss man sich mit den Lernendenvorstellungen befassen (Duit, 2010). Ausubel ist der Auffassung, dass der wichtigste Faktor, der den Lernprozess beeinflusst, das Vorwissen des Lernenden ist und dass die Lehrperson unter dessen Berücksichtigung unterrichten sollte (Duit et al., 2008). Generell sollte man sich an den Lernenden orientieren, man sollte sich in sie hineinversetzen und Sachverhalte aus deren Perspektive betrachten. Indem man Unterrichtsinhalte im Hinblick auf das Interesse, die Bedürfnisse und die kognitiven Fähigkeiten der Schülerinnen und Schüler plant, kann ein erfolgreicherer Lernprozess veranlasst werden (Korner, 2015).

2.2. Lernmotivation

In der fachdidaktischen Forschung gibt es eine Vielzahl an Motivationstheorien und Konstrukten. Der Grundgedanke der Motivationsforschung beziehungsweise der Lernmotivationsforschung ist, dass man wissen will, aus welchen Gründen und für welche Ziele jemand lernt. Insbesondere in Bezug auf schulisches Lernen ist die Kenntnis, wie das Lernverhalten der Schülerinnen und Schüler gesteigert und verbessert werden kann, von großer Bedeutung. Die Lernmotivationstheorien befassen sich mit dem *intentionalen, bewusst gesteuerten* und auf *bestimmte Ziele* gerichteten Lernen (Urhahne, 2008). Sie sollen das Erleben und Verhalten von Personen in Lernsituationen erklären und vorhersagen.

Eine Person, die zu etwas motiviert ist, hat gewisse Motive. Mit Motiven sind hier „*Beweggründe des Handelns*“ (Schlag, 2013, p. 11) gemeint. Sie sind der Grund für die Aufnahme einer Tätigkeit und sozusagen der „Motor“ des Handelns. Der Begriff Motivation steht für alle Motive, die in einer Situation wirken. Diese Motive können durchaus unterschiedliche sein, sogar widersprüchlich (Schlag, 2013).

Fragt man nun Schülerinnen und Schüler warum sie lernen, erhält man Antworten wie: aus Interesse am Fach, aufgrund von Spaß am Lernen, wegen der Bedeutsamkeit des Themas, des Wunsches besser zu werden oder besser als andere zu sein, um dem Leistungsdruck gerecht zu werden, die Erwartungen von Bezugspersonen zu erfüllen, gute Noten beziehungsweise gute Zeugnisse zu erhalten, Lob oder materielle Dinge zu erhalten und Strafe zu vermeiden (Urhahne, 2008).

In den Anfängen der Motivationsforschung wurde davon ausgegangen, dass Menschen aufgrund von Trieben Motivation für eine bestimmte Sache zeigen.

Heutzutage lassen sich folgende sieben Themenfelder von Motivationstheorien voneinander abgrenzen (Urhahne, 2008, p. 151):

- (1) „klassische Leistungsmotivationstheorie
- (2) erweiterte Erwartungs-Wert-Theorien
- (3) Attributionstheorien
- (4) Zielorientierungskonzepte
- (5) Theorien intrinsischer Motivation
- (6) Volitionsmodelle
- (7) Theorien sozialer Motivation“

Die *Leistungsmotivationstheorie* ist die meistuntersuchte Motivationstheorie. Kern der Theorie ist, dass Menschen sich selbst verbessern oder besser beziehungsweise schneller als andere sein wollen. Dabei wird ein Gütemaßstab angestrebt, an welchem der Erfolg oder Misserfolg des Ergebnisses einer Handlung gemessen wird (Urhahne, 2008).

Die Grundlage der Theorie geht auf das Risiko-Wahl-Modell von Atkinson zurück. Dieses besagt, dass leistungsmotiviertes Handeln aus dem Konflikt „zwischen der Hoffnung auf Erfolg und der Furcht vor Misserfolg“ (Urhahne, 2008, p. 152) entsteht. Somit wird eine Handlung ausgeführt um Erfolg herbeizuführen und Zufriedenheit durch eine Handlung zu erfahren (*Erfolgsmotiv*) oder um Misserfolg zu vermeiden und emotionalen Schmerz zu minimieren (*Misserfolgsmotiv*). Überwiegt die Tendenz Erfolg zu suchen, spricht man von einer erfolgsmotivierten Person, überwiegt die Tendenz Misserfolg zu vermeiden, spricht man von einer misserfolgsmotivierten Person. Das Leistungsmotiv ist von Person zu Person verschieden und wird als ein Persönlichkeitsmerkmal einer Person angenommen (Atkinson, 1957).

Ob eine Person nun aus einem Erfolgsmotiv oder einem Misserfolgsmotiv heraus handelt, hängt stark von den Gefühlen und Erfahrungen der Person ab. Affekterleben von Stolz bei Erfolg und Beschämung bei Misserfolg beeinflusst das Handlungsgeschehen wesentlich. Werden die eigenen Erwartungen nicht erreicht, so liegt dies oft an einer unrealistischen Zielsetzung (Urhahne, 2008). Die Leistungsmotivation setzt sich aus dem Motiv der Person, deren Erwartungen sowie dem Anreiz der Handlung zusammen (Atkinson, 1957).

Erwartungswert-Theorien besagen, dass Motivation ein Produkt aus Erwartung und Wert ist, dieses soll maximal werden. Wert bezeichnet hier, das, was man erreichen will und die Erwartung stellt die Wahrscheinlichkeit dar, es zu erreichen (Beckmann & Heckhausen, 2010).

Es wird zwischen vier Arten von Erwartungen unterschieden: *Ergebniserwartung*, *Selbstwirksamkeitserwartung*, *Erfolgserwartung* und *Ergebnis-Folge-Erwartung*.

Bei der *Ergebniserwartung* schätzt eine Person ein, ob mithilfe einer Handlung ein gewünschtes Ergebnis erreicht werden kann. Bei der *Selbstwirksamkeitserwartung* erfolgt eine subjektive Einschätzung der Person, dass sie eine zum Ziel führende Handlung auch ausführen kann. Studien bestätigen einen positiven Zusammenhang zwischen Selbstwirksamkeitserwartungen und Lernleistung (Urhahne, 2008). Bei der *Erfolgserwartung* schätzt eine Person die Wahrscheinlichkeit ein, ob ein bestimmtes Ergebnis realistisch gesehen erreicht werden kann. Die *Ergebnis-Folge-Erwartung* besagt, dass zwischen dem Ergebnis und den Folgen einer Handlung unterschieden werden muss. Beispielsweise kann das Ereignis einer guten Schularbeit verschiedene Folgen mit sich bringen, wie persönliche Freude, weniger Druck, Anerkennung oder Geschenke der Eltern.

Diese vier Erwartungen bilden ab, wie sicher sich eine Person über Handlungen, Ergebnisse und deren Folgen ist (Urhahne, 2008).

Den Erwartungswerttheorien liegt zugrunde, dass sich die Motivation nach der wahrgenommenen Wichtigkeit einer Aufgabe richtet. Es ist entscheidend, ob eine Handlung der Person Vergnügen bereitet und in Hinsicht auf das Erreichen zukünftiger Ziele nützlich ist. Der Mensch versucht außerdem die wahrgenommenen Kosten sowie verschiedene Aspekte wie empfundene Anstrengung, Angst zu versagen, verlorene Zeit und Energie so gering wie möglich zu halten (Urhahne, 2008). Personen entscheiden sich generell für Aufgaben, deren Anreizwert möglichst hoch ist (Beckmann & Heckhausen, 2010). Dabei bestimmen Erwartungen der Selbstwirksamkeit das Verhalten einer Person, beispielsweise wie viel Mühe für die Bewältigung einer Aufgabe aufgewendet wird und wie groß das Durchhaltevermögen bei Hindernissen jeglicher Art ist (Bandura, 1977).

Attributionale Theorien befassen sich mit der Wirkung von Attributionen auf das Erleben und Verhalten. Attributionen sind Ursachenzuschreibungen, die das Erleben und Verhalten beeinflussen. Es wird versucht das Handeln anderer Personen zu verstehen, wobei unter anderem nach den Gründen des Handelns gefragt wird. Die Kenntnis der Gründe ermöglicht es das zukünftige Verhalten einer Person vorherzusagen und auch zu beeinflussen (Stiensmeier-Pelster & Heckhausen, 2010). Mithilfe der Attributionen können Erfolg oder Misserfolg in Leistungssituationen erklärt werden. Die Erklärung erfolgt mithilfe folgender vier Eigenschaften: *eigene Fähigkeit*, *Anstrengung*, *Aufgabenschwierigkeit* und *Zufall*. *Fähigkeit* und *Anstrengung* sind internale Gründe, *Aufgabenschwierigkeit* und *Zufall* external und somit nicht von der Person beeinflussbar.

Führen Personen unausweichliche Misserfolge nicht auf mangelnde Fähigkeiten oder die eigene Intelligenz zurück, so bleibt die Erfolgserwartung konstant. Im Gegensatz dazu entsteht die sogenannte erlernte Hilflosigkeit durch den Glauben an die Unveränderlichkeit des eigenen Wissens und der eigenen Fähigkeiten. Für die Motivation und das Lernverhalten ist es entscheidend, wie Personen mit Misserfolg umgehen (Stiensmeier-Pelster & Heckhausen, 2010).

Laut Heckhausen kann Misserfolgsmotivation in Erfolgsmotivation überführt werden, indem man sich auf folgende drei Prozesse konzentriert (Brunstein & Heckhausen, 2010, p. 185):

- „Zielsetzung
- Ursachenerklärung und
- leistungsthematische Affekte“

Zielorientierungstheorien beschäftigen sich mit der Frage, warum Lernende mit gleichen Fähigkeiten unterschiedlich auf Misserfolge reagieren. Dazu differenziert man zwischen zwei Arten von Zielen, und zwar dem *Lernziel* und dem *Performanzziel* (Urhahne, 2008).

Personen, die ein Lernziel verfolgen, sind danach bestrebt die eigenen Fähigkeiten beziehungsweise Kompetenzen zu entwickeln und zu verbessern, sie glauben an die Veränderlichkeit des eigenen Wissens und der eigenen Fähigkeiten. Misserfolg und Fehler werden als Chance gesehen sich zu verbessern und die Aufmerksamkeit wird auf das Lernen gerichtet. Personen mit einem Lernziel versuchen selbstgesetzte Ziele zu erreichen und Herausforderungen zu bewältigen. Im schulischen Kontext ist eine Klasse, die Lernziele verfolgt, durch folgende Eigenschaften gekennzeichnet: Erfolg wird durch Fortschritt und Verbesserung definiert, Anstrengung und Lernen werden hoch geschätzt, Herausforderungen werden als Ansporn angesehen und Fehler werden als Teil des Lernprozesses betrachtet. Lehrpersonen sind darauf fokussiert wie Schülerinnen und Schüler lernen, anstatt die Leistungen der Lernenden untereinander zu vergleichen. (Anderman, Austin, & Johnson, 2002).

Handelt eine Person aufgrund eines Performanzziels, so ist ihr die Selbstachtung wichtig, die eigenen Fähigkeiten sollen gegenüber anderen bewiesen werden beziehungsweise Defizite sollen verborgen werden. Personen mit Performanzzielen glauben an die Unveränderlichkeit des eigenen Wissens und der eigenen Fähigkeiten (Urhahne, 2008). In diesem Fall zeichnet sich Erfolg durch gute Noten aus, gut entwickelte Fähigkeiten stehen im Vordergrund, Fehler rufen Angst hervor. In einer Klassengemeinschaft sind die Leistungen der Schülerinnen und Schüler im Vergleich zu anderen ausschlaggebend (Anderman et al., 2002). Bei anhaltenden Misserfolgen kann das Verhalten nicht aufrechterhalten werden, was zu hilflosem Verhalten führt.

Es ist metaanalytisch bestätigt, dass Personen, die ein Lernziel verfolgen erfolgreicher sind als Personen, die aufgrund eines Performanzziels handeln. Lernende können jedoch mehrere Zielstellungen zugleich verfolgen (Urhahne, 2008).

Intrinsische Motivation liegt vor, wenn eine Person eine Tätigkeit ausführt, weil ihr diese Freude bereitet, sie als interessant, spannend oder befriedigend erlebt wird. Deswegen wird versucht die Tätigkeit möglichst lange und häufig auszuführen, dabei muss nicht unbedingt ein Ziel erreicht werden (Rheinberg, 2010). Diese Form der Motivation ist die erstrebenswerteste, da die Lernenden aus sich heraus die Initiative ergreifen und gerne lernen. Intrinsisch motivierte Personen suchen nach Herausforderung und verwenden viel Zeit und Mühe darauf, diese erfolgreich zu bewältigen (Urhahne, 2008).

Zur intrinsischen Motivationstheorie gibt es folgende drei Subtheorien:

- Selbstbestimmungstheorie
- Flowtheorie
- Interessenstheorie

Die Selbstbestimmungstheorie der Motivation sowie die intrinsische Motivation werden im Kapitel 2.3. genauer beschrieben.

Das Flow-Erleben bezeichnet das völlige Aufgehen in einer Tätigkeit, ein exzessives Engagement, wobei alle nicht für die Tätigkeit notwendigen Kognitionen ausgeblendet werden. Weiters ist das Zeiterleben beeinträchtigt, Personen widmen sich stundenlang der Tätigkeit, ohne mitzubekommen wie viel Zeit vergangen ist (Rheinberg, 2010). Dabei liegt die gesamte Aufmerksamkeit bei der Aufgabe beziehungsweise Tätigkeit, man fühlt sich eins mit der Aufgabe. Befindet man sich im Flow, so ist man zu Höchstleistungen fähig (Schlag, 2013). Beispiele für solche Tätigkeiten sind, Maler, die an einem Bild malen, Chirurgen, die stundenlang operieren, das Arbeiten am PC oder das Erlernen neuer aufregender Dinge. Laut Rheinberg (2010) kann theoretisch jede Handlung im Flow ausgeführt werden, es gibt nur wenige Faktoren, die den Flow behindern, dazu zählen häufige Störungen (Handy, Personen, ...), Zeitdruck und ein schlechtes soziales Klima. In Bezug auf die Schule lässt sich folgendes sagen: Aufgaben sollten eine klare Struktur haben, genaue und sofortige Rückmeldung ermöglichen und die Anforderungen sollten in einem Gleichgewicht zu den Fähigkeiten der Lernenden stehen. Um weiterhin Flow zu erleben, müssen die Anforderungen gesteigert werden (Urhahne, 2008).

In der Interessenstheorie geht es darum, dass das Interesse beeinflusst, welchen Inhalten sich die Lernenden zuwenden und wie gut sie sich etwas merken und auch behalten.

Interesse entsteht aus einer Wechselbeziehung zwischen der Person und dem Gegenstand, bei diesem kann es sich um ein Themengebiet, eine konkrete Sache oder auch um eine Tätigkeit handeln. Interesse unterscheidet sich von anderen motivationalen Zuständen dadurch, dass der Gegenstand des Interesses für die Person eine besondere Bedeutung hat, Handlungen werden als selbstbestimmt wahrgenommen und es werden positive Gefühle dabei empfunden (Urhahne, 2008).

Volitionstheorien befassen sich mit der Willenskraft eines Individuums. Unter Volition versteht man regulative Prozesse einer Person, die entscheiden, welche Art der Motivation bei welchen Handlungen auf welche Weise realisiert werden (Heckhausen & Heckhausen, 2010).

Eine hohe Motivation an sich garantiert nicht, dass eine angestrebte Handlung auch tatsächlich umgesetzt wird. Mithilfe des Rubikon-Modells (Achtziger & Gollwitzer, 2010) kann der Handlungsprozess beschrieben werden, indem er in vier chronologisch aufeinander folgende Phasen, die durch diskrete Übergänge voneinander getrennt sind, gegliedert wird. Das Rubikon-Modell versucht das Entstehen, Heranreifen und Vergehen von Motivation zu beschreiben (Achtziger & Gollwitzer, 2010):

Phase 1: prädezieionale Handlungsphase

In der ersten Phase des Handlungsprozesses wägt eine Person zwischen verschiedenen Zielen und Wünschen ab, diese werden auf ihre Bedeutsamkeit und Realisierbarkeit geprüft. Erst dann erfolgt die Festlegung eines konkreten Handlungsziels, das erreicht werden soll.

Phase 2: präaktionale Phase

Diese Phase findet sich zwischen Zielbildung und Handlung, sie ist durch das Planen charakterisiert. Kann die Zielintention nicht sofort in eine Handlung umgesetzt werden, wird ein Vorsatz gebildet. Die Realisierung des in der vorherigen Phase festgelegten Zieles soll durch Kontrolle der Motivation, Aufmerksamkeit, Gefühle und einer guten Strategie zur Bewältigung von Misserfolgen vorangetrieben werden (Urhahne, 2008).

Phase 3: aktionale Phase

Hier steht das Handeln im Vordergrund, diese Phase ist dadurch charakterisiert, wie intensiv und ausdauernd die Handlung betrieben wird. Diese Intensität hängt von der Volitionsstärke des Ziels ab. Die Realisierung des Ziels soll durch ausdauerndes Verhalten und durch Anstrengungssteigerung bei auftretenden Schwierigkeiten erreicht werden (Achtziger & Gollwitzer, 2010).

Phase 4: postaktionale Handlungsphase

In der letzten Phase erfolgt eine Bewertung des Handlungsergebnisses und der eigenen Leistung. Ist die Person mit dem Handlungsergebnis zufrieden, so wird dieses deaktiviert. Erreicht die Person das Ziel nicht, so kann sie das Anspruchsniveau herabsetzen oder das alte Ziel durch ein neues ersetzen (Achtziger & Gollwitzer, 2010).

Soziale Motivationstheorien besagen, dass Motivation durch die Anerkennung anderer Menschen wie Lehrpersonen, Eltern oder Klassenkolleginnen und -kollegen beziehungsweise Freunden zustande kommt sowie durch das Herstellen von Beziehungen. Sozial motivationale und intrapsychische Prozesse sind relevant, um die Motivation von Schülerinnen und Schülern zu verstehen. Lernende, die durch Bezugspersonen emotionale und kognitive Unterstützung erhalten, sind meist erfolgreicher in Bezug auf das Lernen und auf die spätere berufliche Laufbahn (Wentzel, 1999). Das Anspruchsniveau und die Erwartungen der Eltern spielen ebenfalls eine wichtige Rolle. Die soziale Motivation kann auf ein psychologisches Grundbedürfnis nach *sozialer Eingebundenheit* zurückgeführt werden. Dabei ist es einem Individuum wichtig, Anschluss zu finden und nicht zurückgewiesen zu werden. Lernende, die sich von der Lehrperson geachtet und gerecht behandelt fühlen, zeigen eine höhere Motivation und eine positivere Einstellung zur Schule.

Soziale Motivation ist dadurch gekennzeichnet, dass bessere Lernleistungen angestrebt werden, um beispielsweise von Lehrpersonen oder den Eltern geachtet zu werden oder ihnen eine Freude zu machen (Urhahne, 2008). Je stärker der Glaube einer Person an die eigene Wirksamkeit und Kompetenz, desto wahrscheinlicher ist es, dass die Person mit starkem Willen ihre Ziele verfolgt (Wentzel, 1999).

Personen können unterschiedliche motivationale Gründe haben, diese können einander ergänzen oder sogar gegensätzlich sein.

Zusammenfassend kann gesagt werden, dass am Beginn der Lernhandlung immer die Bedürfnisse, Motive und Ziele der Person stehen. Während der Lernhandlung spielt das Interesse, der Flow, das Maß an Selbstbestimmung sowie die Anstrengung eine wesentliche Rolle, damit die Handlung fortgesetzt und bis zum Ende durchgeführt wird. Am Ende der Lernhandlung stellen sich entweder Gefühle von Stolz und Freude bei Erfolg, oder Gefühle von Scham, Trauer und Wut bei Misserfolg ein. Meistens hat eine Person mehrere Gründe, um eine Lernhandlung zu beginnen. Deswegen können die verschiedenen Lernmotivationstheorien nicht getrennt voneinander betrachtet werden, zur Beschreibung des Lernprozesses müssen Übereinstimmungen und Unterschiede berücksichtigt werden (Urhahne, 2008).

2.3. Selbstbestimmungstheorie der Motivation

Die Selbstbestimmungstheorie der Motivation (Self-Determination Theory) ist „eine organismische und dialektische Theorie der menschlichen Motivation“ (Deci & Ryan, 1993, p. 5), welche von Deci und Ryan begründet wurde. Organismisch bedeutet, dass der Mensch nach Integration strebt, sich weiterentwickeln will und sich selbst reguliert. Diese Prozesse werden durch intrinsische, motivationale Faktoren angetrieben, wobei die Bedürfnisse des Menschen im Mittelpunkt stehen. Dialektisch ist die Theorie insofern, als dass diese Prozesse der Integration in einer Wechselbeziehung mit der sozialen Umgebung der Person stehen (Deci & Ryan, 1993).

Der zentrale Begriff der Theorie ist das *Selbst*. In der Selbstbestimmungstheorie versteht man darunter einerseits einen Prozess, Personen entwickeln sich aufgrund der Interaktion mit dem sozialen Umfeld weiter und andererseits kann das *Selbst* als Ergebnis dieser Entwicklung angesehen werden. Kern der Theorie bilden die *angeborenen psychologischen Bedürfnisse, grundlegende Fähigkeiten* und die *Interessen* der Person (Deci & Ryan, 1993).

Die Selbstbestimmungstheorie geht davon aus, dass Menschen mit ihrem Verhalten eine Intention verfolgen. Menschen gelten demnach als motiviert, wenn sie mit ihrem Verhalten etwas erreichen wollen, wie zum Beispiel ein bestimmtes Ziel. Bei diesem Ziel kann es sich um ein schnell erreichbares, oder um ein langfristiges Ziel handeln, das in ein paar Jahren, Monaten oder Tagen erreicht werden kann. Derartige intentionale Handlungen gehen von der Person aus und zielen auf ein positives Erlebnis, wie etwas Interessantes, Spannendes, Aufregendes oder auf ein Ergebnis, welches andauerndes Engagement erfordert, wie beispielsweise das Erreichen einer guten Note, ab. Verhaltensweisen, die keine Absicht verfolgen, werden als *amotiviert* bezeichnet. Dazu zählen Handlungen, die kein erkennbares Ziel aufweisen sowie unkontrollierte Handlungen, wie zum Beispiel ein Wutanfall.

Die intentionalen, motivationalen Handlungen haben unterschiedliche qualitative Ausprägungen. Diese Ausprägungen werden nach dem Grad ihrer Selbstbestimmung beziehungsweise nach dem Ausmaß ihrer Kontrolliertheit eingeteilt. Am einen Ende dieses Kontinuums steht das selbstbestimmte Verhalten, am anderen Ende das kontrollierte Verhalten. Damit wird das Maß der Selbstbestimmtheit einer motivierten Handlung festgelegt. Frei gewählte Handlungen, bei denen das Individuum seinen Zielen und Wünschen nachgeht, gelten als selbstbestimmt und autonom. Aufgezwungene Handlungen durch andere Menschen oder inneren Druck gelten als kontrolliert, sie weisen deshalb ein geringes Maß an Selbstbestimmung auf, wodurch die Motivation des Individuums prinzipiell nicht sehr hoch sein kann (Deci & Ryan, 1993).

Eine der Fragen, die sich stellt, ist, warum sind Menschen überhaupt motiviert, oder warum werden gewisse Tätigkeiten motivierender empfunden, als andere.

Die meisten Theorien befassen sich nicht mit diesen Fragen und gehen davon aus, dass der Mensch eine ausreichend große psychische Energie hat, um seine Ziele zu verfolgen. Nach der Selbstbestimmungstheorie gibt es drei Quellen motivationaler Handlungsenergie und zwar die *physiologischen Bedürfnisse* (Triebe), *Emotionen* und die *psychologischen Bedürfnisse*. Menschliches Verhalten braucht alle drei Energiequellen. Die psychologischen Bedürfnisse gelten dabei als besonders wichtig, da sie diejenigen Prozesse beeinflussen, mit deren Hilfe der Mensch seine Triebe und Emotionen regelt. Die drei angeborenen psychologischen Bedürfnisse sind die folgenden (Deci & Ryan, 1993):

- Bedürfnis nach Kompetenz und Wirksamkeit
- Bedürfnis nach Autonomie und Selbstbestimmung
- Bedürfnis nach sozialer Eingebundenheit und Zugehörigkeit

Laut Deci und Ryan ist die Erfüllung dieser drei psychologischen Bedürfnisse in allen Kulturen ein Prädiktor für psychisches Wohlbefinden (Deci & Ryan, 2008).

Generell lässt sich sagen: Der Mensch will sich mit anderen verbunden fühlen, in einer Gemeinschaft etwas bewirken und sich dabei autonom fühlen. Somit wird die Frage beantwortet, warum Menschen gewisse Ziele mehr motivieren, als andere: es werden angeborenen Bedürfnisse befriedigt (Deci & Ryan, 1993).

Intrinsische & extrinsische Motivation

Wie bereits erwähnt kann motiviertes Verhalten differenziert werden. Intrinsisch motivierte Handlungen weisen ein Maximum an Selbstbestimmung auf, eine Person entscheidet frei darüber, welche Handlungen sie ausführt. Eine Tätigkeit wird dabei ihrer selbst willen ausgeführt. Intrinsische Motivation beinhaltet Neugier, Spaß, Spontanität, Interesse an Gegenständen, Sachverhalten und der Umwelt (Deci & Ryan, 1993).

Extrinsisch motivierte Handlungen sind dadurch gekennzeichnet, dass sie aufgrund ihrer Konsequenzen durchgeführt werden, sei es das Erreichen positiver Konsequenzen oder das Vermeiden negativer Auswirkungen (Schlag, 2013). Dazu zählen beispielsweise eine Belohnung, Bestrafung oder äußerer Druck beziehungsweise Zwang. Generell lässt sich sagen, dass extrinsische Motivation nicht spontan auftritt, sie wird durch Personen oder innerpsychische Prozesse in Gang gesetzt (Deci & Ryan, 1993).

Intrinsische und extrinsische Motivation sind jedoch keine Gegensätze, denn auch extrinsisch motiviertes Verhalten kann selbstbestimmt sein, dies ist in Abbildung 1 dargestellt.

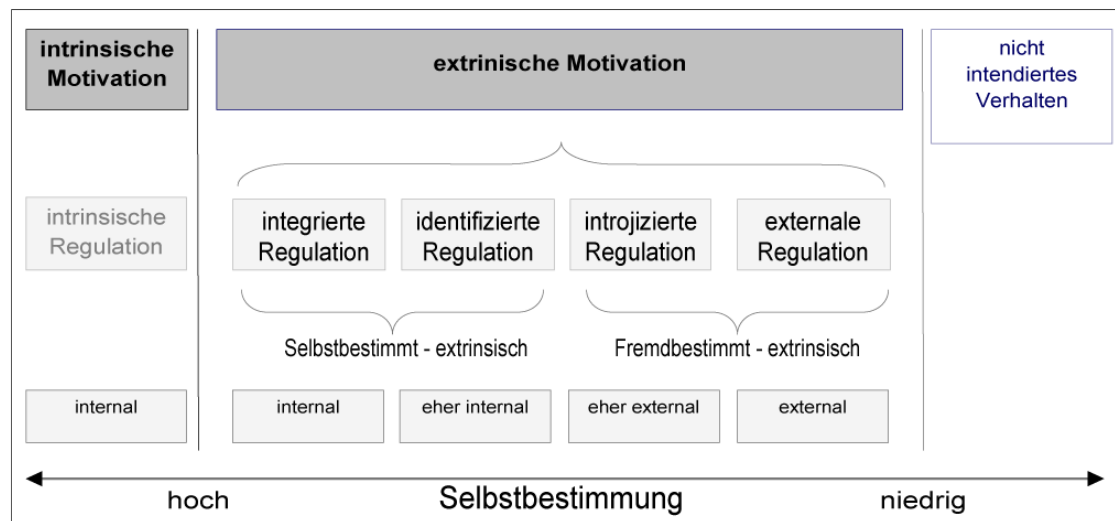


Abbildung 1: Selbstbestimmungskontinuum von handlungsbezogenen Werten (Steuer, Fasching, & Dresel)

Auf der linken Seite befindet sich die *intrinsische Motivation*, welche eine hohe Selbstbestimmung aufweist. Auf der rechten Seite befindet sich das *amotivierte Verhalten* mit dem niedrigsten Maß an Selbstbestimmung.

In der Mitte befindet sich die *extrinsische Motivation*, die in vier Gruppen der Verhaltensregulation unterteilt werden kann (Deci & Ryan, 1993).

Die erste Stufe heißt *externale Regulation*, bei der Handlungen ausgeführt werden, um eine Belohnung zu erhalten oder eine Bestrafung zu vermeiden. Handlungen auf dieser Stufe sind zwar intentional, aber von äußeren Faktoren abhängig, es wird weder Autonomie noch Freiwilligkeit wahrgenommen.

Bei der *introjizierten Regulation* ergreift eine Person eine Handlung aufgrund von innerem Druck, dazu gehören auch Handlungen, die für die Selbstachtung wichtig sind. Das Individuum führt eine Handlung aufgrund von gesellschaftlichen Normen oder eines schlechten Gewissens aus. Das Verhalten wird durch innere Kräfte beeinflusst, die Person will die Handlung nicht von sich aus ausführen. Ein Beispiel für diese Stufe wäre ein Lernender, der glaubt er muss die Matura machen, da er sonst minderwertiger wäre. Im Falle der introjizierten Regulation können Personen oft schlecht mit Misserfolgserlebnissen umgehen (Deci & Ryan, 1993).

Die nächste Stufe, die *identifizierte Regulation*, ist durch Verhaltensweisen, die für das Selbst als wichtig oder wertvoll empfunden werden, gekennzeichnet. Eine Handlung wird

ausgeführt, weil die Person sie für wichtig hält. Ein Schüler oder eine Schülerin lernt für die Matura, weil er oder sie studieren möchte. Lernende auf dieser Stufe haben Freude an der Schule und können Misserfolge gut bewältigen.

Die *integrierte Regulation* ist die höchste Stufe innerhalb der extrinsischen Motivation. Bei dieser Stufe erleben Personen ein hohes Maß an Selbstbestimmung, eine Person führt eine Handlung selbstständig und konsequent aus, um ein von sich aus gesetztes, als wichtig empfundenen Ziel, zu erreichen. Die integrierte Regulation stellt mit der intrinsischen Motivation die Basis selbstbestimmten Handelns dar (Deci & Ryan, 1993).

Insbesondere für Lehrpersonen ist es von Bedeutung zu wissen, was die intrinsische Motivation positiv und negativ beeinflusst, damit die Lernenden möglichst hoch motiviert sind. Studien haben gezeigt, dass die intrinsische Motivation abnimmt, wenn Personen extrinsischen Einflüssen wie Belohnung, Bestrafung, Bewertungen, Zeitdruck oder besonderen Auszeichnungen ausgesetzt sind (Deci & Ryan, 1993). Die intrinsische Motivation nimmt ab, da diese Faktoren als kontrollierend erlebt werden, wodurch die Selbstbestimmung sinkt. Unter gewissen Bedingungen kann extrinsische Belohnung jedoch zur Aufrechterhaltung der intrinsischen Motivation führen. Was sich positiv auf die intrinsische Motivation auswirkt, sind Rückmeldungen, die als selbstständigkeitsfördernd erlebt werden, sie halten die intrinsische Motivation nicht nur aufrecht, sondern verstärken diese. Ein Angebot an verschiedenen Wahlmöglichkeiten, respektvoller Umgang und ein optimales Anforderungsniveau wirken sich ebenfalls positiv aus. Einer der stärksten Faktoren, der die intrinsische Motivation fördert, ist positives, nicht kontrollierendes Feedback, dieses stärkt die *wahrgenommene Kompetenz*, welche zu den Grundbedürfnissen zählt (Deci & Ryan, 1993). Hier findet sich außerdem eine Parallele zu Hatties Studie „Visible learning“, in der Hattie feststellte, dass Feedback den Lernprozess sehr positiv beeinflussen kann (Steffens & Höfer, 2014).

Extrinsisch motiviertes Verhalten kann „*durch die Prozesse der Internalisation und Integration in selbstbestimmte Handlungen überführt werden*“ (Deci & Ryan, 1993, p. 9). Generell lässt sich sagen, ermöglichen soziale Umweltfaktoren die Befriedigung der drei psychologischen Bedürfnisse, so ist es wahrscheinlicher, dass intrinsische Motivation sowie die Integration extrinsischer Motivation auftritt.

Hochqualifiziertes Lernen kann somit nur durch ein vom Individuum ausgehendes Engagement erreicht werden kann. Durch äußere Faktoren zustande kommendes Lernen bleibt nur so lange aufrechterhalten, solange dieser Faktor vorhanden ist und als genügend stark empfunden wird (Korner, 2015).

Die Lehrperson sollte versuchen das Interesse der Lernenden anzuregen und konzeptuelles Lernen fördern. Weiters kann die Lehrperson das Autonomieerleben unterstützen, welches sehr wichtig ist für einen beständigen, länger anhaltenden Lernerfolg. Da die Eltern einen großen Einfluss auf das Lerngeschehen haben, wäre es ratsam diese stärker miteinzubeziehen.

Bemühen sich wichtige Bezugspersonen darum, dass die Befriedigung der psychologischen Bedürfnisse erfolgen kann, unterstützen diese weiters das Autonomiebestreben des Lernenden und ermöglichen die Erfahrung der eigenen Kompetenz, so erleben Lernende ein hohes Maß an Selbstbestimmung, was sich auf deren Motivation positiv auswirkt. Die Erfahrung, eigene Handlungen frei wählen zu können, ist dabei einer der wichtigsten Faktoren (Deci & Ryan, 1993).

2.4. Intrinsic Motivation Inventory (IMI)

Das Intrinsic Motivation Inventory ist ein multidimensionales Messinstrument zur Bestimmung der intrinsischen Motivation und der Selbstregulation von Personen bezüglich einer Aktivität. Es wurde von Deci und Ryan entwickelt und basiert auf deren Selbstbestimmungstheorie (Deci & Ryan, 2003). Auf Deutsch könnte man Intrinsic Motivation Inventory als "Fragenbogen zur intrinsischen Motivation" übersetzen. Es ist eines der wenigen guten Werkzeuge zur Bestimmung der Motivation, welches jedoch in englischer Sprache verfasst ist.

Das IMI besteht insgesamt aus 45 Items, die sich auf folgende sieben Skalen verteilen (Deci & Ryan, 2003, pp. 3,4):

- *interest/enjoyment* - Interesse/Vergnügen (7 Items)
- *perceived competence* - (selbst) wahrgenommene Kompetenz (6 Items)
- *effort/importance* - Einsatz/Wichtigkeit (5 Items)
- *value/usefulness* - Wert/Nützlichkeit (7 Items)
- *pressure/tension* - Druck/Anspannung (5 Items)
- *perceived choice* - wahrgenommene Wahlfreiheit (7 Items)
- *relatedness* - soziale Eingebundenheit (8 Items)

Die Skala *interest/enjoyment* ist die einzige Skala, welche die intrinsische Motivation selbst abbildet (Deci & Ryan, 2003).

Die Skalen *perceived competence* und *perceived choice* bilden positive Prädiktoren für intrinsische Motivation, die Skala *pressure/tension* einen negativen Prädiktor (Deci & Ryan, 2003).

Die Skala *value/usefulness* wird hauptsächlich für Studien der Internalisierung verwendet, wobei die Idee dahinter ist, dass Menschen Aktivitäten, welche sie als nützlich oder wertvoll für sich selbst empfinden, internalisieren und selbst regulieren (Deci & Ryan, 2003).

Die Skala *effort/importance* bildet ein eigenes Konstrukt ab, welches beschreibt, über wieviel Motivation eine Person verfügt. Die anderen Skalen hingegen geben die Ursachen und den Grad der Selbstbestimmung an (Korner, 2015).

Bei der Skala *relatedness* stehen die sozialen Interaktionen und Freundschaften im Vordergrund. Diese Skala versucht zu erfassen, inwieweit Personen soziale Eingebundenheit beziehungsweise Zugehörigkeit empfinden. Es wird vermutet, dass bei einem hohen Maß an sozialer Eingebundenheit eher intrinsische Motivation auftritt.

Laut Deci und Ryan sind die Skalen *interest/enjoyment*, *pressure/tension* und *effort/importance* faktorenanalytisch trennbar und stabil. Die Autoren stellten fest, dass die Anordnung der Items sowie die Auswahl der Subskalen eine vernachlässigbare Auswirkung auf die Ergebnisse haben. Für eigene Forschungszwecke kann bezüglich der Skalen variiert werden. Es können außerdem Items weggelassen werden, jedoch sollte jede Skala aus mindestens vier Items bestehen (Korner, 2015).

Die Auswertung der Items erfolgt über eine 7-stufige Likert-Skala. Bei der Likert-Skala bewertet die Testperson ein Item mit einer Abstufung von *trifft zu* bis *trifft nicht zu* oder von *sehr gut* bis *sehr schlecht*, Genauerer dazu in Kapitel 2.6. (Bühner, 2011).

Die Autoren schlagen vor, dass falls ein eigener Fragebogen zusammengestellt wird, der nicht alle Skalen beziehungsweise Items enthält, auf jeden Fall eine Faktorenanalyse durchgeführt werden sollte (Deci & Ryan, 2003).

2.5. Test- & Fragebogenkonstruktion

In diesem Kapitel werden Methoden zur Fragebogenkonstruktion und die Eigenschaften eines guten psychometrischen Tests, zu denen die Haupt- und Nebengütekriterien zählen, besprochen. Mit *psychometrischer Test* ist ein wissenschaftliches Verfahren zur Untersuchung eines oder mehrerer Persönlichkeitsmerkmale gemeint. Ziel dessen ist es, die absolute oder relative Ausprägung einer Eigenschaft oder Fähigkeit bei einer oder mehrerer Personen zu messen (Bühner, 2011).

Weiters werden Methoden zur Analyse der Items beschrieben und zwar die exploratorische und die konfirmatorische Faktorenanalyse.

2.5.1. Hauptgütekriterien

Zu den Hauptgütekriterien zählen Objektivität, Reliabilität, Validität und die Skalierbarkeit, sie werden an dieser Stelle genauer beschrieben.

Objektivität:

Ein Test soll bei der Durchführung und Auswertung objektiv sein, das heißt, wird derselbe Test von unterschiedlichen Personen durchgeführt, ausgewertet oder interpretiert, so muss die Testleistung der Person gleich bleiben und darf nicht variieren. Die gesamte Testdurchführung sowie Auswertung und Interpretation sollte standardisiert ablaufen (Bühner, 2011).

Es werden drei Arten von Objektivität unterschieden:

- Durchführungsobjektivität

Die Durchführung eines Tests muss immer nach gleicher Vorgehensweise erfolgen und sollte nicht variieren. Dazu müssen die Bedingungen, unter denen der Test, Fragebogen oder das Interview durchgeführt wird, klar definiert sein. Es sollte bereits im Vorhinein feststehen, welche Hilfestellungen gegeben werden und wie lange die Testung dauern soll. Insbesondere muss auf die Wortwahl bei der Instruktion geachtet werden, damit Rückfragen der Personen so minimal wie möglich gehalten werden können. Dadurch wird auch die Gefahr minimiert, dass den Personen möglicherweise Hilfestellungen in irgendeiner Form gegeben werden (Bühner, 2011).

- Auswertungsobjektivität

Die Auswertung eines Tests muss unabhängig von der auswertenden Person sein. Um dies zu erreichen sind genaue Auswertungsvorschriften nötig, beispielsweise in Form von Schablonen oder Auswertungsblättern bei Tests und Fragebögen.

Die Auswertungsobjektivität hängt von der Art des Fragens ab, bei offenen Fragestellungen ist es schwieriger die Auswertung vorzunehmen, als beispielsweise bei Likert-skalierten Items. Bei offenen Antwortformaten muss exakt festgelegt werden, ab wann ein Ergebnis zu werten ist und wie (Bühner, 2011).

- Interpretationsobjektivität

Die Beurteilung und Interpretation eines Tests sollte unabhängig von der beurteilenden Person sein. Jede auswertende Person sollte zum gleichen Ergebnis kommen und sagen können, ob ein Testergebnis unterdurchschnittlich, durchschnittlich oder überdurchschnittlich im Vergleich zu einer Normgruppe ist. Für die Interpretationsobjektivität sind vor allem ausreichend große Normstichproben und überprüfte Gütekriterien wichtig. Es empfiehlt sich standardisierte Interpretationsmöglichkeiten für einen Test vorzulegen (Bühner, 2011).

Reliabilität:

Reliabilität bezeichnet den Grad der Genauigkeit, mit dem ein Test eine gewisse Eigenschaft misst, unabhängig davon, was er zu messen beansprucht. Das heißt, man kann sagen, wie genau ein Test ein Merkmal misst, ohne zu wissen, was er misst. Die Reliabilität kann auf verschiedene Weisen geschätzt werden, die Basis dafür sind Korrelationskoeffizienten. Vom mathematischen Standpunkt her ist die Reliabilität als Varianzverhältnis definiert und zwar „als der Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte“ (Bühner, 2011, p. 144). Für Reliabilitätsabschätzungen werden aus der getesteten Stichprobe die entsprechenden Parameter, wie beispielsweise Varianzen, für die Grundgesamtheit geschätzt.

Zur Schätzung der Reliabilität gibt es folgende Methoden (Bühner, 2011):

- Konsistenzmethode

Die Reliabilität wird oft mit der Konsistenzmethode geschätzt, dies erfolgt mithilfe der Korrelation beziehungsweise Kovarianz von Items oder Testteilen. Man spricht bei derartigen Reliabilitätseinschätzungen von interner oder innerer Konsistenz. Die Schätzung kann durch eine Testhalbierung oder durch die Analyse der inneren Konsistenz erfolgen.

Für die Methode der *Testhalbierung* wird der Test zuerst einmal an einer Stichprobe durchgeführt, im nächsten Schritt wird er in zwei Testhälften aufgeteilt. Für die Berechnung wird die Korrelation der Rohwertpaare beider Testhälften ermittelt. Diese Korrelation wird nachher mithilfe einer Korrekturformel aufgewertet. Man kann die Items beispielsweise nach gerader oder ungerader Reihungsnummer, zufällig oder nach Testzeit aufteilen, um nur ein paar Methoden zu nennen (Bühner, 2011).

Für die Analyse der *inneren Konsistenz* wird der Test nur einmal an einer Stichprobe durchgeführt. Danach erfolgt eine Zerlegung des Tests in so viele Untertests, wie er Items besitzt. Man kann ihn auch in beliebig viele inhaltsähnliche Teile zerlegen, je nach Testformat und Konstrukt.

Die Berechnung der Reliabilität wird mithilfe der Itemvarianzen und Itemkovarianzen durchgeführt, beispielsweise durch das Cronbach- α . Konsistenzanalysen sind nur dann sinnvoll, wenn eindimensionale Konstrukte vorliegen. Bei Tests, die heterogene Aufgaben enthalten, ist die Konsistenzmethode nicht sinnvoll (Bühner, 2011).

Der **Cronbach- α** -Koeffizient stellt die gängige Methode zur Schätzung der inneren Konsistenz dar, er ist wie folgt definiert (Bühner, 2011, p. 166):

$$\alpha = \frac{k}{k-1} \cdot \left(1 - \frac{\sum_{i=1}^k \hat{\sigma}_i^2}{\hat{\sigma}_x^2} \right)$$

wobei k die Anzahl der Testitems bezeichnet, $\hat{\sigma}_i^2$ die Varianz des Testitems und $\hat{\sigma}_x^2$ die Gesamtvarianz des Tests.

Der Cronbach- α -Koeffizient ist ein kombinierter Index, er berücksichtigt nicht nur die Itemkorrelationenshöhe, sondern auch die Anzahl der Items (Bühner, 2011).

- Paralleltestmethode

Bei der Paralleltestmethode werden zwei Parallelformen des Tests bei derselben Stichprobe getestet. Hierzu muss es möglich sein, zwei Formen des Tests zu erstellen. Zur Berechnung wird die Korrelation der Rohwertpaare aus beiden Testformen ermittelt (Bühner, 2011).

Bei dieser Methode müssen Übungs- beziehungsweise Transfereffekte berücksichtigt werden. Wenn eine Person Test A durchgeführt hat, dann wird sie in Test B, der eine Parallelform von Test A darstellt, bessere Leistungen zeigen. Damit dies nicht passiert und kontrolliert werden kann, verwendet man ein „cross over design“, bei dem die Testabfolge variiert, zum Beispiel:

Gruppe 1: Test A – Test B, Gruppe 2: Test B – Test A (Bühner, 2011).

- Retestmethode

Bei der Retestmethode bearbeitet dieselbe Stichprobe den gleichen Test nach einem gewissen Zeitabstand wieder. Der Zeitabstand kann je nach Inhalt gewählt werden, zu beachten ist, dass er sich auf die Höhe des Korrelationskoeffizienten auswirken kann. Ist das Intervall zwischen den Testungen zu kurz, so erfolgt eine positiv verzerrte Schätzung der Reliabilität, Grund dafür sind die Transfer- oder Übungseffekte. Die Retest-Korrelation fällt geringer aus, je größer das Zeitintervall zwischen den beiden Testungen ist.

Geschätzt wird die Reliabilität durch die Korrelation der Rohwertpaare zwischen der ersten Testung und der Testwiederholung. Die Retestmethode ist für Situationen empfehlenswert, welche nicht von Lern- oder Übungseinflüssen abhängig sind. Es muss beachtet werden, dass die Durchführungsbedingungen zu beiden Testzeitpunkten gleich sind (Bühner, 2011).

Trennschärfe

Die Trennschärfe steht in einem direkten Zusammenhang mit der Reliabilität. Sie gibt an, wie gut ein Item eine bestimmte Eigenschaft misst und wird durch eine Korrelation quantifiziert. Derartige Trennschärfen nennt man Eigentrennschärfen, diese unterscheiden sich von den Fremdtrennschärfen. Hohe Eigentrennschärfen der Items begünstigen eine hohe Reliabilität des Testwerts. Rechnerisch erhält man die Trennschärfe, indem man „aus der Korrelation des Items mit dem gesamten Summenwert der übrigen Items das jeweilig zu korrelierende Item“ herausrechnet (Bühner, 2011, p. 172). Der verwendete Summenwert ist dadurch um den Beitrag des jeweiligen Items, für den man die Trennschärfe schätzt, korrigiert. Diese Korrektur wird Part-Whole-Korrektur genannt.

Die *Produkt-Moment-Korrelation* kann als Trennschärfe aufgefasst werden und wird auch Pearson-Korrelation genannt.

Für Likert-skalierte Items erfolgt die Berechnung der Trennschärfe folgendermaßen (Bühner, 2011, p. 172):

$$\hat{\rho}_{i(x-i)} = \frac{\hat{\rho}_{ix} \cdot \hat{\sigma}_x - \hat{\sigma}_i}{\sqrt{\hat{\sigma}_x^2 + \hat{\sigma}_i^2 - 2 \cdot \hat{\rho}_{ix} \cdot \hat{\sigma}_x \cdot \hat{\sigma}_i}}$$

wobei $\hat{\rho}_{i(x-i)}$ den Trennschärfekoeffizient der Aufgabe i mit der Skala x , bezeichnet. Aufgabe i wurde dabei nicht berücksichtigt. $\hat{\rho}_{ix}$ ist die Korrelation des Items i mit der Skala x , $\hat{\sigma}_i$ die Standardabweichung des Items i , $\hat{\sigma}_x$ die Standardabweichung des Tests bzw. der Skala x , $\hat{\sigma}_i^2$ die Varianz des Items i und $\hat{\sigma}_x^2$ die Varianz des Tests bzw. der Skala x .

Je stärker Personen einem Item zustimmen, desto höher fällt auch ihr Skalenwert auf der Skala ohne das entsprechende Item aus (Bühner, 2011).

Validität:

Man spricht von Validität, wenn ein Test das misst, was er zu messen beansprucht.

Dabei wird zwischen drei Arten der Validität unterschieden: der Inhaltsvalidität, der Konstruktvalidität und der Kriteriumsvalidität. Die Inhaltsvalidität entspricht der obigen Definition, es ist jedoch sehr schwierig diese zu bestimmen, da sie nicht empirisch überprüfbar ist. Deswegen bestimmt man die Validität eines Tests indirekt und zwar mithilfe der Konstruktvalidität und der Kriteriumsvalidität. Damit wird eigentlich nicht die Validität des Tests selbst ermittelt, sondern die Validität der abgeleiteten Aussagen (Bühner, 2011).

- **Inhaltsvalidität**

Für die Inhaltsvalidität muss ein Test und auch jedes Item das zu messende Merkmal hinreichend präzise erfassen. Mit präzise ist gemeint, dass das Konstrukt durch die Items abgebildet wird. Die Items müssen das Konstrukt präzise messen und dürfen dabei auf kein anderes Konstrukt abbilden. Eine hohe Inhaltsvalidität stellt einen der wichtigsten Faktoren bei der Testkonstruktion dar. Sie wird jedoch nicht numerisch erfasst, sondern aufgrund logischer und fachlicher Überlegungen bestimmt.

- **Konstruktvalidität**

Darunter versteht man, dass der Test jenes Merkmal misst, das er messen soll. Die Konstruktvalidität lässt sich in drei Validitäten aufspalten, die konvergente, diskriminante und die faktorielle Validität. Als konvergente Validität bezeichnet man hohe Korrelationen zwischen Tests, die Ähnliches messen. Diskriminante Validität liegt vor, wenn zwei Tests, die unterschiedliche Eigenschaften messen, inhaltlich aber ähnlich sind, genügend voneinander abgrenzbar sind. Für die faktorielle Validität müssen Items eines Tests inhaltlich einen Faktor abbilden (Bühner, 2011).

- **Kriteriumsvalidität**

Die Kriteriumsvalidität gibt den Zusammenhang der Testleistung zu einem oder mehreren Kriterien an, mit denen der Test aufgrund seines Messanspruchs korrelieren soll. Ermittelt wird die Kriteriumsvalidität basierend auf den Zusammenhängen zwischen Testkennwerten und Kriterien. Es kann zwischen folgenden vier Arten der Kriteriumsvalidität unterschieden werden: prognostische Validität, Übereinstimmungsvalidität, retrospektive Validität und inkrementelle Validität. Die ersten drei Validitäten befassen sich mit Korrelationen eines aktuellen Testergebnisses mit einem vergangenen oder zukünftigen Ergebnis.

Die inkrementelle Validität soll die Vorhersage eines Kriteriums über einen anderen Test hinaus ermöglichen (Bühner, 2011).

Skalierbarkeit:

Das Gütekriterium der Skalierbarkeit gibt an, ob Berechnungen mit einer gültigen Verrechnungsvorschrift durchgeführt wurden. Die Punktevergabe der Items und die daraus resultierenden Summenwerte müssen die Verhaltensrelation widerspiegeln. Die Verrechnung der Items zu Punkten ist von den Itemtrennschärfen abhängig. Die Gültigkeit der Rechenvorschrift sollte auf jeden Fall gegeben sein (Bühner, 2011).

2.5.2. Nebengütekriterien

Die Nebengütekriterien sind weitere Indikatoren eines guten psychometrischen Tests und wichtig für die Beurteilung der Güte. Sie werden im Folgenden kurz beschrieben (Bühner, 2011):

Normierung

Man bezeichnet einen Test als normiert, wenn die Möglichkeit besteht die Ergebnisse einer Testung mit anderen Stichproben zu vergleichen. Ist ein Test normiert, so kann man sagen, ob eine getestete Person ein unterdurchschnittliches, durchschnittliches oder überdurchschnittliches Ergebnis in Bezug auf andere Personen hat.

Vergleichbarkeit

Ein Test gilt als vergleichbar, wenn es ein oder mehrere Parallelförmigkeiten oder Tests mit demselben Gültigkeitsbereich gibt. Wird eine Person mit zwei ähnlichen Tests getestet, so sollten ähnliche Ergebnisse dabei herauskommen.

Ökonomie

Ein Test ist ökonomisch, wenn er eine kurze beziehungsweise angemessene Durchführungszeit aufweist, einfach zu handhaben ist, für eine Gruppentestung tauglich ist, wenig Material benötigt wird und schnell auswertbar ist.

Nützlichkeit

Ein Test ist nützlich, wenn er ein Merkmal misst, für dessen Untersuchung ein praktisches Bedürfnis besteht und noch keine zufriedenstellenden Tests vorhanden sind. Bevor ein Test erstellt wird, sollte überprüft werden, ob es bereits ähnliche Tests gibt.

Ist dies der Fall, so muss der neu entwickelte Test Vorteile gegenüber den anderen aufweisen, wie beispielsweise eine bessere Kriteriumsvalidität oder eine bessere Skalierbarkeit (Bühner, 2011).

Zumutbarkeit

Das Kriterium der Zumutbarkeit meint, dass ein Test die getestete Person in zeitlicher, psychischer und körperlicher Hinsicht schonen soll und nicht unnötig belastet.

Fairness

Das Gütekriterium der Fairness ist erfüllt, wenn die aus einem Test resultierenden Messwerte einzelne Gruppen der Testung nicht diskriminieren.

Nicht-Verfälschbarkeit

Ein Test ist nicht verfälschbar, wenn eine Person die Leistung im Test nicht willentlich oder unwillentlich in eine gewisse Richtung lenken kann, so dass sie daraus einen Vorteil zieht. Dieses Kriterium ist schwer zu kontrollieren, da Personen immer in der Lage sind das Testergebnis auf eine bestimmte Weise zu beeinflussen. Man sollte wenn möglich zumindest bekanntgeben, wie stark ein Test nach unten als auch nach oben verfälscht werden kann (Bühner, 2011).

2.5.3. Exploratorische Faktorenanalyse

Die exploratorische Faktorenanalyse dient dazu, Faktoren zu finden, wenn nicht genau bekannt ist, wie viele Eigenschaften die Items messen beziehungsweise welche Items Indikatoren bestimmter Eigenschaften darstellen. Mithilfe einer exploratorischen Faktorenanalyse sollen Zusammenhänge der Items untereinander durch eine möglichst geringere Anzahl dahinter liegender Faktoren erklärt werden. Diese Faktoren werden als Basis für die Modellbildung beziehungsweise zur Erklärung von Theorien herangezogen. Im Rahmen einer exploratorischen Faktorenanalyse wird die Korrelationsmatrix der Items analysiert. Die beiden gängigen Methoden dazu sind die Hauptachsenanalyse und die Maximum-Likelihood-Faktorenanalyse (Bühner, 2011).

In Bezug auf die Faktorenanalyse ist der Begriff Ladung wichtig, definiert ist sie als ein semipartielles standardisiertes Regressionsgewicht. *„Es gibt an, um wie viele Standardabweichungen sich im Durchschnitt die Ausprägung auf dem Item ändert, wenn sich die Ausprägung auf dem Faktor im Durchschnitt um die Standardabweichung ändert“* (Bühner, 2011, p. 300).

2.5.4. Konfirmatorische Faktorenanalyse

Bei der konfirmatorischen Faktorenanalyse können theoretisch oder empirisch gut fundierte Modelle auf ihre empirische Passung mit den Daten getestet werden oder mit alternativen Modellen verglichen werden (Bühner, 2011). Sie erlaubt es Kovarianzen oder Korrelationen zwischen beobachteten Variablen beziehungsweise Items auf latente Variablen zurückzuführen. Im Gegensatz zur exploratorischen Faktorenanalyse werden bei der konfirmatorischen Faktorenanalyse Kovarianzmatrizen analysiert sowie die Passung der Kovarianzmatrix mit einem reduzierten Modell untersucht. Bei der exploratorischen Faktorenanalyse wird eine Itemstruktur mithilfe der Itemkorrelationen gesucht.

Die empirischen Kovarianzen der Items sollen mithilfe linearer Gleichungen, auf Basis eines vom Anwender entwickelten Modells, erklärt werden. Die Items fungieren als manifeste oder beobachtbare Variablen (Bühner, 2011).

2.6. Messinstrumente der Motivation

Im englischsprachigen Raum gibt es eine Vielzahl an Messinstrumenten, welche die Motivation messen. Im deutschsprachigen Raum hingegen gibt es nur wenige und die bestehenden Instrumente sind meist sehr spezifisch. Korner (2015) wollte im Rahmen ihrer „Cross-Age Peer Tutoring in Physik“ (CAPT) Studie unter anderem die Lernmotivation von Schülerinnen und Schülern testen. Die bis dahin bestehenden Instrumente zur Evaluierung der Motivation wurden als nicht passend empfunden. Korner wollte ein Messinstrument konstruieren, das einen genaueren Einblick in die Motivation der Lernenden zulässt. Als theoretische Basis wurde die Selbstbestimmungstheorie von Deci und Ryan gewählt, da diese sowohl die intrinsische als auch die extrinsische Motivation berücksichtigt. Korner wollte nämlich unter anderem untersuchen, ob die Schülerinnen und Schüler von sich aus Lernen wollen und welchen Einfluss die Tutoren auf die Motivation der Lernenden haben (Korner, 2015). Die Selbstbestimmungstheorie ist somit am geeignetsten, alle anderen Theorien befassen sich nicht mit den Ursachen der Motivation.

Wird ein Instrument konstruiert, sollte das Alter der Lernenden, ihre Muttersprache, sowie ihre Lesekompetenz berücksichtigt werden. Das Instrument soll einerseits diese Rahmenbedingungen berücksichtigen, psychometrischen Anforderungen genügen und über eine verständliche und klare Sprache verfügen (Korner, 2015).

2.6.1. Bestehende Messinstrumente

Eines der Instrumente zur Ermittlung der Motivation stellt der *Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen* (FAM) (Rheinberg, Vollmeyer, & Burns, 2001) dar, er bezieht sich auf das Modell der Leistungsmotivation (Risiko-Wahl-Modell). Die aktuelle Motivation setzt sich aus Situationsfaktoren und der Motivstruktur der Person zusammen (Rheinberg et al., 2001). FAM besteht aus vier Subskalen: *Misserfolgsbefürchtung*, *Erfolgswahrscheinlichkeit* aus Sicht der Testperson, *Interesse* an der Situation und empfundener *Herausforderung*. Beantwortet werden die Items mit einer 7-stufigen Likert-Skala. Diese vier Faktoren können in einer faktorenanalytischen Untersuchung zufriedenstellend voneinander getrennt werden (Korner, 2015).

Ein weiterer Fragebogen zur Ermittlung der Motivation wurde von Müller et al. (2007) entwickelt, dem wiederum die Selbstbestimmungstheorie zugrunde liegt. Bei diesem Fragebogen wird auf die vier Formen der extrinsischen Verhaltensregulation eingegangen, die intrinsische Motivation erhält keine eigene Skala. Jede Skala im Fragebogen besteht aus vier oder fünf Items, die mit einer 6-stufigen Likert-Skala beantwortet werden.

Die Autoren entwickelten eine Formel zur Berechnung des sogenannten Selbstbestimmungsindex (SDI), in dem die Mittelwerte der vier Skalen folgendermaßen zusammengefasst werden (Müller, Hanfstingl, & Andreitz, 2007, p. 7):

$$\text{SDI} = 2 \cdot (\text{intrinsische Regulation}) + \text{identifizierte Regulation} \\ - \text{introjierte Regulation} - 2 \cdot (\text{externe Regulation})$$

Der SDI liegt im Intervall $[-15, +15]$. Ein positives Vorzeichen bedeutet selbstbestimmte Regulation, ein negatives Vorzeichen externe Regulation. Auch dieser Fragebogen weist ein gutes Maß an Reliabilität und Validität auf (Korner, 2015).

Der FAM konnte für die Zwecke von Korner nicht eingesetzt werden, da er einerseits für Personen ab einem Alter von 17 Jahren getestet wurde und da es sich andererseits um einen spezifischen Fragebogen handelt, der die *aktuelle Motivation in Lern- und Leistungssituationen* testet (Rheinberg et al., 2001). Der Fragebogen von Müller wurde nicht verwendet, da er sich auf extrinsisch reguliertes Verhalten beschränkt (Korner, 2015).

Das von Berger und Hänze (2004) entwickelte Instrument besteht aus den drei Skalen *soziale Eingebundenheit*, *Kompetenzerleben* und *Autonomieerleben*, welche die psychologischen Bedürfnisse der Selbstbestimmungstheorie darstellen (Deci & Ryan, 1993). Die Items können schnell bearbeitet werden und für verschiedene Unterrichtsmethoden sowie Unterrichtssituationen verwendet werden. Ein Grund für Korner dieses Instrument nicht zu verwenden waren die teilweise mäßig ausgeprägten Ladungen sowie die hohen Querladungen. Ein Item der Skala *Autonomieerleben* hat eine mäßige Ladung von 0,59 und eine hohe Querladung von 0,59 auf die Skala *soziale Eingebundenheit*. Weiters ist der Fragebogen für Schülerinnen und Schüler der Sekundarstufe 2 gedacht und wurde in der 12. Schulstufe getestet, Korner plante ihrer CAPT-Studie jedoch für die Sekundarstufe 1 (Korner, 2015).

Die *Kurzskala intrinsischer Motivation* (KIM) (Wilde, Bätz, Kovaleva, & Urhahne, 2009) besteht aus zwölf Items. Getestet wurde nachdem die Schülerinnen und Schüler einen außerschulischen Lernort besucht hatten. Die Autoren entschieden sich dafür, da sie der Meinung sind, dass der außerschulische Lernort intrinsische Motivation hervorruft und die Schülerinnen und Schüler dadurch Spaß am Lernen haben (Wilde et al., 2009). Der Fragebogen ist eine verkürzte und leicht veränderte Version des IMI. Dabei bildet die Kurzskala die Faktoren *interest/enjoyment*, *perceived competence*, *perceived choice* und *pressure/tension* mit je drei Items ab. Die vier Skalen ließen sich faktorenanalytisch gut voneinander trennen (Wilde et al., 2009).

Es stellte sich jedoch heraus, dass die Übersetzungen aus dem Englischen teilweise problematisch sind und dass manche Items auch einer anderen Skala zugeordnet werden könnten. KIM konnte für die Zwecke von Korner nicht verwendet werden, da die wichtige Skala *effort/importance* darin nicht enthalten ist (Korner, 2015).

Da keines der deutschsprachigen Instrumente für die Studie von Korner passend war, wurde beschlossen das IMI ins Deutsche zu übersetzen und an die Stichprobe anzupassen.

2.6.2. Probleme mit dem Intrinsic Motivation Inventory

Als ersten Versuch entwickelte Korner aus dem IMI einen Fragebogen zur Motivation, der aus sechs Skalen bestand (*perceived competence, perceived choice, interest/enjoyment, effort/importance, value/usefulness, relatedness*) und insgesamt 40 Items hatte (Deci & Ryan, 2003). Zur Beantwortung der Items wurde eine 5-stufige Likert-Skala verwendet. Das ursprüngliche IMI verwendet eine 7-stufige Likert-Skala, von dieser wurde jedoch abgesehen, da Schülerinnen und Schüler der Sekundarstufe 1 möglicherweise Schwierigkeiten mit der Differenzierung bekommen könnten (Korner, 2015).

Im Laufe der Testung stellte sich heraus, dass es bei manchen Items doch erhebliche Verständnisprobleme gab. Die Items wurden sorgfältig übersetzt und überarbeitet, durch Expertinnen und Experten sowie durch Lehrkräfte validiert, dennoch hatten Lernende mit gewissen Wörtern Probleme.

Als Beispiel wurde das Originalitem „*I felt really distant to this person.*“ (Deci & Ryan, 2003, p. 4) mit „*Ich fühlte mich der Person richtig fern.*“ (Korner, 2015, p. 105) übersetzt. Einige Schülerinnen und Schüler einer sechsten Schulstufe wussten nicht genau was mit dem Wort „fern“ gemeint ist. Nach der Pilotierung ließ Korner Items weg und versuchte unverständliche Items umzuschreiben. Negativ gepolte Items bereiteten den Lernenden besonders große Verständnisschwierigkeiten, deshalb wurden einige weggelassen.

Ein weiteres Problem stellte die Formulierung und Bedeutung der Items dar. Dieses Problem könnte dadurch entstehen, dass die Lernenden in ihrer emotionalen und sprachlichen Entwicklung noch nicht so weit sind, um die Unterschiede zu verstehen (Korner, 2015). Beispielsweise konnten die Lernenden keinen Unterschied zwischen den folgenden beiden Items erkennen:

„*I felt like I had to do this.*“ (Deci & Ryan, 2003, p. 4) – „*Mir kam vor, dass ich das tun musste.*“ (Korner, 2015, p. 106)

„*I did this activity because I had to.*“ (Deci & Ryan, 2003, p. 4) – „*Ich machte diese Aktivität, weil ich musste.*“ (Korner, 2015, p. 106)

Bei den beiden Items ist der *Ort der Entscheidung (locus of causality)* ein anderer, im ersten Item liegt er in der Person selbst, im zweiten Item außerhalb der Person, etwa bei jemandem, der die Person unter Druck setzt (Korner, 2015). Bei zwei weiteren Items, die widersprüchlich sind, empfanden die Lernenden keinen Widerspruch, es war für sie das Gleiche.

Weitere Probleme stellten die Itempolarität und die doppelte Verneinung dar. Negativ gepolte Items werden dazu eingesetzt, um zu verhindern, dass Personen den Items immer nur zustimmen (Bühner, 2011). Jedoch dürften die Schülerinnen und Schüler womöglich nicht verstanden haben, dass wenn sie dem negativ gepolten Item nicht zustimmen, bei der Likert-Skala „stimmt gar nicht“ auswählen müssen. Korner fand in Exitinterviews heraus, dass viele Lernende die doppelte Verneinung nicht anwenden konnten. Die Lernenden wussten nicht, ob sie beispielsweise bei dem Item *„Ich würde mit dieser Person lieber nicht mehr zusammenarbeiten.“* (Korner, 2015, p. 111) „stimmt völlig“ oder „stimmt gar nicht“ ankreuzen sollen, wenn sie doch gerne wieder mit der Person zusammenarbeiten wollen. Aus diesen Gründen verzichtete Korner auf negativ gepolte Items, die durch Verwendung von doppelter Verneinung formuliert wurden.

2.6.3. Entwicklung eines Messinstruments mit den Methoden von Wilson

In den Arbeiten von Korner (2014), Klingenböck (2016) und Schmidt (2017) erwies sich die Herangehensweise von Wilson zur Erstellung der Skalen als sehr nützlich. In dieser Arbeit wird für die Entwicklung der Skala *relatedness* ebenfalls nach Wilson vorgegangen. Dazu wird die Herangehensweise an dieser Stelle beschrieben.

Der Four Building Blocks Approach

Mithilfe des Four Building Blocks Approach (Wilson, 2005) kann ein Messinstrument entwickelt werden, indem vier Schritte immer wieder durchlaufen werden. Instrument meint an dieser Stelle eine Technik, etwas Beobachtetes aus der Realität in Beziehung zu etwas zu setzen, das gemessen werden kann. Dabei werden Rückschlüsse (Inferenzen) von den Beobachtungsdaten auf das zu messende Merkmal gemacht (Wilson, 2005).

Will man ein Instrument entwickeln, so muss man sich zuerst überlegen welches Konstrukt abgebildet werden soll, was das zu entwickelnde Instrument messen soll und wie die Messung verwendet wird. Im nächsten Schritt wird die Construct Map erstellt, welche das Konstrukt darstellen soll. Auf Basis der Construct Map können Items erstellt werden, dabei gilt es sich zu überlegen, welches Antwortformat für die Items am geeignetsten ist.

Die erhaltenen Antworten von den Items werden dann wieder mit dem Konstrukt selbst verglichen, wobei die Sinnhaftigkeit überprüft wird. Diese Schritte stellen einen Kreislauf dar, welcher in Abbildung 2 zu sehen ist.

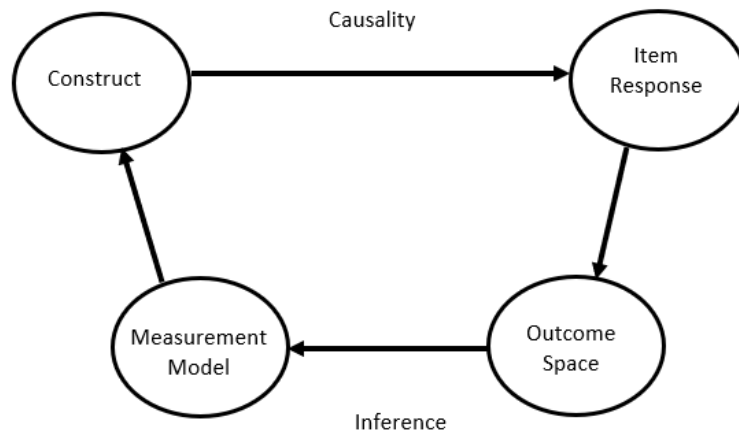


Abbildung 2: Four building blocks (Wilson, 2005, p. 31)

Dieser Kreislauf muss, je nachdem wie gut das Konstrukt abgebildet wurde, mehrmals durchlaufen werden. Treten Schwierigkeiten bei der Erstellung der Construct Map und Items auf, so sollte man sich mit dem Konstrukt näher beschäftigen, Theorien, auf denen das Konstrukt basiert, untersuchen, über bisherige Forschungsergebnisse recherchieren und sich intensiv mit der Literatur auseinandersetzen. Wilson empfiehlt sich mit anderen Personen vom Fach auszutauschen und sich verschiedene Meinungen einzuholen (Wilson, 2005).

Die vier Schritte des Kreislaufs werden im Folgenden genauer beschrieben:

Construct Map

Die Construct Map soll das zu untersuchende Konstrukt visuell darstellen und enthält dabei eine Definition seines Inhaltes. Ein Konstrukt ist eine Idee oder ein Konzept, welches erforscht werden soll. Laut Wilson wird angenommen, dass das Konstrukt, welches gemessen wird, zwei Extrema aufweist, es reicht von einer minimalen zu einer maximalen Merkmalsausprägung. Zwischen den beiden Extrema müssen passende Abstufungen beziehungsweise Ausprägungen gefunden werden. Die Construct Map wird im Laufe der Entwicklung des Instruments iterativ verbessert. Schritte, die zur Klärung führen, sind: das Konstrukt anderen erklären, die Items erstellen, Testen der Items mit wenig ProbandInnen und ein Vergleich der resultierenden Daten mit dem Konstrukt (Wilson, 2005).

Im Allgemeinen hat eine Construct Map zwei Spalten, in der ersten Spalte stehen die qualitativen Antworten der befragten Person, in der zweiten Spalte stehen die Item-Antwortmöglichkeiten. Am Beginn und am Ende der Construct Map befinden sich die Extrema des Konstrukts. Wilson empfiehlt Interviews als Hilfe zur Entwicklung und Verbesserung der Construct Map (Wilson, 2005).

Item Response

Im ersten Schritt der Itemkonstruktion überlegt man sich, in welchen realen Situationen sich das theoretische Konstrukt manifestiert. Dies stellt erstmals nur eine Vermutung dar. Laut Wilson ist es in der Praxis oft so, dass man zuerst die Items erstellt und im Nachhinein das Konstrukt genauer definiert. Die Itemerstellung dient somit zur Klärung und Spezifizierung des Konstrukts. Später müssen die Items jedoch als Realisierung des Konstrukts angesehen werden können.

Items können verschiedene Formen mit verschiedenen Antwortmöglichkeiten haben. Gängige Antwortmöglichkeiten wären beispielsweise das offene Antwortformat, bei dem die Personen frei einen Text schreiben und das geschlossene Antwortformat mit einer Likert-Skala (Wilson, 2005).

Die Likert-Skala ist eine Ratingskala, bei der Personen Aussagen bewerten beispielsweise mit „*trifft zu bis trifft nicht zu* oder von *sehr gut bis sehr schlecht*“ (Bühner, 2011, p. 110). Die Abstufung der Antwortkategorien kann je nach Bedarf gewählt werden, wobei Abstufungen von einer 3-stufigen Skala bis hin zu einer 7-stufigen Skala gängig sind.

Die erstellten Items müssen getestet und gegebenenfalls überarbeitet werden, beispielsweise in einer Pilot-Studie. Es ist besser, wenn man zu Beginn mehr Items hat, da sich während des Testens herausstellen kann, dass Items ungeeignet sind (Wilson, 2005).

Outcome Space und Measurement Model

Der Outcome Space stellt den Zwischenschritt zwischen den Itemantworten und dem Measurement Model dar. In diesem Schritt wird entschieden, welche Aspekte der Itemantworten als Basis für den Rückschluss auf das Konstrukt verwendet werden und wie diese Aspekte der Itemantworten kategorisiert und gewertet werden (Wilson, 2005). Bei offenen Formaten ist dies schwieriger als bei geschlossenen Formaten, wie zum Beispiel Likert-skalierten Items. Bei einer Likert-Skala müssen die Antworten nicht kategorisiert werden, da dies bereits von der Testperson gemacht wird. Der Outcome Space muss jedenfalls immer spezifisch für das Konstrukt und den Kontext, in dem es verwendet wird, sein. Die Punkte der Items müssen in Bezug zum Konstrukt gesetzt werden.

Im Falle der Likert-Skala werden die Antworten meistens mit 0,1,2,3,4 gewertet (Wilson, 2005). Für diese Arbeit wird eine 5-stufige Likert-Skala von eins bis fünf verwendet, wobei „1“ keine Zustimmung bedeutet und „5“ eine hohe Zustimmung.

Das Measurement Model überträgt die gewerteten Antworten der Items auf einen Punkt der Construct Map, es ist sozusagen eine Messvorschrift (Korner, 2015). Das Measurement Model soll die Antworten der Items wieder mit dem Konstrukt verknüpfen (Schmidt, 2017). Um dies zu erreichen gibt es zwei Möglichkeiten: Bei der ersten Möglichkeit bekommt eine Testperson eine bestimmte Anzahl von Items, denen dasselbe Konstrukt zugrunde liegt. Die Punkte der Items werden addiert oder es wird der Mittelwert gebildet. Wird ein hoher Gesamtscore erreicht, kann man auf eine hohe Merkmalausprägung schließen (Korner, 2015). Bei dieser Möglichkeit wird ein Zusammenhang zwischen Gesamtscore und Konstrukt hergestellt. Die zweite Möglichkeit fokussiert sich auf die Items, hier richtet sich der Gesamtscore danach, wie vielen Aussagen die Person zugestimmt hat beziehungsweise wie viele sie abgelehnt hat.

Für welche Möglichkeit man sich entscheidet hängt von der Wahl des Outcome Space ab (Korner, 2015).

3) Forschungsdesign

3.1. Forschungsfragen

Aufgrund von Problemen mit der Übersetzung und Übernahme der Items des IMI entwickelte Korner (2015) im Rahmen ihrer CAPT-Studie die Skala *effort/importance* neu. Es stellte sich heraus, dass für die Entwicklung eines guten Instruments zur Ermittlung der Motivation von Schülerinnen und Schülern alle Skalen neu entwickelt werden müssen. Im Rahmen ihrer Diplomarbeiten entwickelten Klingenböck (2016) die Skala *interest/enjoyment* und Schmidt (2017) die Skalen *perceived competence*, *perceived choice* und *value/usefulness* neu. In dieser Arbeit wird die letzte verbleibende Skala *relatedness* mit der gleichen Vorgehensweise entwickelt. Dazu werden die folgenden Forschungsfragen untersucht:

1. Lässt sich das Konstrukt „*relatedness*“ aus dem IMI auf die Schulsituation übertragen?
2. Kann eine trennscharfe Skala zum Teilkonstrukt *relatedness* entwickelt werden, die sich faktorenanalytisch bezüglich verschiedener Stichproben abgrenzen lässt?
3. Wie wirkt sich eine vor der Testung durchgeführte Intervention auf die Antworten der Schülerinnen und Schüler aus?

Die erste und die zweite Forschungsfrage befassen sich mit der neu zu entwickelnden Skala *relatedness* und deren zugrundeliegendem Konstrukt. Es soll untersucht werden, ob das Konstrukt *relatedness*, so wie es im IMI verwendet wurde, auch für den schulischen Kontext übernommen werden kann, oder ob es Änderungen beziehungsweise Überarbeitungen bedarf. Nachdem die Überlegungen zum Konstrukt gemacht wurden, können die Items der Skala *relatedness* erstellt werden. In der Analyse soll festgestellt werden, ob die neue Skala trennscharf ist und sich bezüglich verschiedener Stichproben abgrenzen lässt.

Bei der dritten Forschungsfrage soll ermittelt werden, wie stark beziehungsweise in welcher Form sich eine zuvor durchgeführte Intervention auf die Antworten der Schülerinnen und Schüler auswirkt. Da jüngere Lernende im Alter von zwölf bis vierzehn Jahren befragt werden, wird vermutet, dass sie sich bei der Beantwortung von Gruppenarbeit-Items leichter tun, wenn zuvor eine Intervention durchgeführt wurde. Durch die Intervention haben die Lernenden eine unmittelbare Situation, auf die sie sich bei der Beantwortung beziehen können. Allerdings könnte eine derartige Intervention das Antwortverhalten der Schülerinnen und Schüler und damit die Faktorstruktur der Skalen beeinflussen. Dies gilt es in der Analyse zu untersuchen.

3.2. Erstellung der Construct Map und der Items

Zur Erstellung der Construct Map und der Items müssen zuerst ein paar Überlegungen angestellt werden, beispielsweise was sind die beiden Extrema des Konstrukts, wie können Aussagen formuliert werden, die das Konstrukt abbilden, welche Abstufungen können gewählt werden, aber auch für welche Zielgruppe ist der Fragebogen gedacht. Dabei gilt es das Alter, die Bildung sowie die Sprachbeherrschung der Lernenden zu beachten (Bühner, 2011). Für diese Arbeit wird die gleiche Zielgruppe wie in den Studien zuvor verwendet und zwar Schülerinnen und Schüler der siebten und achten Schulstufe.

Um auf Evidenzen basierende Items zu erstellen braucht es die Construct Map, beides wird mithilfe des Four Building Blocks Approach von Wilson (2005) erstellt. Die Construct Map und die Items sollen im Laufe der Arbeit verbessert werden. Die erste Version der Construct Map und der Items bilde das Konstrukt erstmals nur grob ab.

3.2.1. Erste Version der Construct Map und der Items

Die Lernenden fühlen sich beim Arbeiten mit anderen sehr wohl und helfen einander.	
Lernende	Antwort zu den Items
Die Lernenden fühlen sich beim Arbeiten mit anderen sehr wohl.	Lernende arbeiten sehr gut miteinander zusammen und helfen einander.
Die Lernenden fühlen sich beim Arbeiten mit anderen wohl.	Lernende arbeiten gut miteinander zusammen und helfen einander meistens.
Das Arbeiten mit anderen ist für die Lernenden in Ordnung.	Lernende bearbeiten gemeinsam Aufgaben, wobei es ihnen gleichgültig ist, ob sie mit anderen arbeiten oder nicht.
Die Lernenden fühlen sich beim Arbeiten mit anderen nicht sehr wohl.	Lernende bearbeiten gemeinsam Aufgaben, wobei sie diese lieber alleine bearbeiten möchten.
Die Lernenden fühlen sich beim Arbeiten mit anderen nicht wohl.	Lernende können nicht miteinander arbeiten, sie stören einander und möchten alleine arbeiten.
Die Lernenden fühlen sich beim Arbeiten mit anderen unwohl und möchten die Aufgaben alleine bearbeiten.	

Tabelle 1: Version 1 der Construct Map zur Skala relatedness

Items - Version 1

- 1) Ich fühlte mich der Person nicht nah. (R)
- 2) Ich bezweifle, dass die Person und ich Freunde werden. (R)
- 3) Ich habe der Person vertraut.
- 4) Ich möchte mit dieser Person gerne öfters zusammenarbeiten.
- 5) Ich will mit dieser Person nicht mehr zusammenarbeiten. (R)
- 6) Wenn die Person und ich öfters etwas gemeinsam machen, könnten wir Freunde werden.
- 7) Ich fühle mich der Person nah.
- 8) Beim Bearbeiten der Aufgabe fühlte ich mich in der Gruppe wohl.
- 9) Ich konnte spüren, wie es der anderen Schülerin / dem anderen Schüler ging.
- 10) Ich wäre lieber mit meiner Freundin / meinem Freund in einer Gruppe gewesen. (R)
- 11) Durch die Gruppenarbeit habe ich mehr gelernt, als wie wenn ich die Aufgaben alleine gelöst hätte.

1) - 6) Items aus dem IMI (Deci & Ryan, 2003) ins Deutsche übersetzt

7), 9) Items aus Cross-Age Peer Tutoring in Physik (Korner, 2015, pp. 242, 245-246)

Die Items 8), 10), 11) wurden neu erstellt.

Die Bezeichnung „(R)“ steht für ein negativ gepoltes Item.

3.2.2. Befragungen an Schulen

Um die Construct Map sowie die Items zu verbessern und evidenzbasierter zu gestalten, wurde an verschiedenen Schulen in Wien eine Befragung von Schülerinnen und Schülern durchgeführt. Dies fand im Rahmen eines offenen Gespräches mit den Lernenden statt. Von einer schriftlichen Befragung wurde abgesehen, da bei einem Gespräch nachgefragt und eruiert werden kann, was die Lernenden im Falle von unklaren Aussagen meinen.

Im Zuge der Befragungen wurden die folgenden Fragen gestellt:

- Wann fühlt ihr euch in der Klasse eingebunden/miteinbezogen?
- Unter welchen Bedingungen fühlt ihr euch in der Klasse eingebunden/miteinbezogen?
- Wie hilft eine Gruppenarbeit, dass ihr euch mehr eingebunden fühlt?
- Unter welchen Bedingungen fühlt ihr euch in der Klasse wohl/unwohl?
- Unter welchen Bedingungen könnt ihr am besten mitarbeiten/ lernen?
- Woran merkt man, dass sich jemand in der Klasse wohl/unwohl fühlt?

Im Anhang B, Kapitel 9.1., befindet sich eine Übersicht über die Antworten der Schülerinnen und Schüler sowie weitere Daten.

3.2.3. Zweite Version der Construct Map und der Items

Auf Basis der Befragung der Schülerinnen und Schüler wurden die Construct Map und die Items angepasst und verbessert. In der Construct Map wurden keine großen Änderungen vorgenommen, die Items wurden jedoch überarbeitet.

Die Lernenden fühlen sich beim Arbeiten mit anderen sehr wohl und helfen einander.	
Lernende	Antwort zu den Items
Die Lernenden fühlen sich beim Arbeiten mit anderen sehr wohl.	Lernende arbeiten sehr gut miteinander zusammen und helfen einander.
Die Lernenden fühlen sich beim Arbeiten mit anderen wohl.	Lernende arbeiten gut miteinander zusammen und helfen einander meistens.
Das Arbeiten mit anderen ist für die Lernenden in Ordnung.	Lernende bearbeiten gemeinsam Aufgaben, wobei es ihnen gleichgültig ist, ob sie mit anderen arbeiten oder nicht.
Die Lernenden fühlen sich beim Arbeiten mit anderen nicht sehr wohl.	Lernende bearbeiten gemeinsam Aufgaben, wobei sie diese lieber alleine bearbeiten würden.
Die Lernenden fühlen sich beim Arbeiten mit anderen unwohl.	Lernende stören einander beim Bearbeiten der Aufgabe und möchten alleine arbeiten.
Das Arbeiten miteinander ist unmöglich, die Lernenden würden die Aufgaben lieber alleine bearbeiten.	

Tabelle 2: Version 2 der Construct Map zur Skala relatedness

Items - Version 2

Die Items 1. – 4. wurden aus der Arbeit von Korner (2015, pp. 245-246) und Schmidt (2017, p. 63) übernommen, Items 5. und 6. wurden aus der vorherigen Version übernommen. Die Items 7. – 12. wurden Items auf Basis der Lernendenbefragungen neu erstellt.

Die kursiv gesetzten Items wurden aus der ersten Version der Items übernommen.

1. Ich konnte ihr / ihm wirklich vertrauen.
2. Ich würde gerne Gelegenheit haben mit ihr / ihm öfter zusammenzuarbeiten.
3. Es ist möglich, dass das andere Kind und ich Freunde werden, wenn wir viel zusammenarbeiten.
4. *Ich konnte spüren, wie es der anderen Schülerin / dem anderen Schüler ging.*
5. *Beim Bearbeiten der Aufgabe fühlte ich mich wohl in der Gruppe.*
6. *Durch die Gruppenarbeit habe ich mehr gelernt, wie wenn ich die Aufgaben alleine bearbeitet hätte.*
7. Mit meinem Freund / meiner Freundin kann ich besser zusammenarbeiten.
8. Wir kamen in der Gruppe gut miteinander aus.
9. Ich hatte in der Gruppe weniger Angst etwas Falsches zu sagen.
10. Ich fühle mich nicht wohl, wenn eine / einer in der Gruppe alles macht. (R)
11. In der Gruppenarbeit haben wir viel miteinander gesprochen und einander geholfen.
12. Die anderen in der Gruppe waren gut drauf und lachten öfters.

3.2.4. Dritte Version der Construct Map und der Items

Auf Basis einer Diskussion unter Fachdidaktikerinnen und Fachdidaktikern im Forschungsseminar der fachdidaktischen Arbeitsgruppe Physik wurden die Construct Map und die Items komplett neu überarbeitet. In der Diskussion wurde angesprochen, dass die bisherige Construct Map stark auf Gruppenarbeit eingeht. Während der Diskussion wurde klar, dass das Konstrukt *relatedness* komplizierter und vielfältiger ist, als bisher gedacht. Da das IMI für Teamsport entwickelt wurde, kann das Konstrukt nicht einfach auf den Schulkontext übertragen werden (Korner, 2015).

In den Befragungen der Lernenden kristallisierten sich die Bereiche Klassenklima, Eigenschaften der Lehrperson sowie Eigenschaften der Schülerinnen und Schüler heraus, welche für das Konstrukt *soziale Eingebundenheit* von Bedeutung sind. Der Fokus wurde von der Gruppenarbeit weg, hin auf das Klassenklima und die sozialen Interaktionen in der Klasse gelenkt. Es wird davon ausgegangen, dass bei einem guten Klassenklima optimal zusammengearbeitet werden kann, wodurch auch die Motivation der einzelnen Lernenden steigen kann. Werden die Lernenden von ihren Mitschülerinnen beziehungsweise Mitschülern nicht respektvoll behandelt oder wird beispielsweise über ihre Aussagen ausgelacht, so führt dies vermutlich zu einer Verringerung der Motivation.

Weiters spielt die Beziehung zwischen den Lernenden untereinander, wie auch die Beziehung zwischen der Lehrperson und den Lernenden, eine wichtige Rolle in Hinblick auf die Motivation. Die Lernenden müssen einander respektvoll behandeln und ausreden lassen, die Lehrperson muss die Lernenden respektvoll behandeln und umgekehrt.

Generell wurde festgestellt, dass der Begriff *relatedness* geschärft werden muss, da sich die Dynamik in einer Schulklasse komplexer darstellt.

Auf Basis des Feedbacks der Fachdidaktikerinnen und Fachdidaktiker wurde beschlossen, zwei Construct Maps zu erstellen, eine, die Schülerinnen und Schüler betreffend (Version A) und eine, die Lehrperson betreffend (Version B). Diese sollen das Konstrukt *relatedness* beziehungsweise die soziale Eingebundenheit in einem allgemeineren Zusammenhang abbilden.

Die Lernenden fühlen sich aufgrund des guten Klassenklimas in der Schulumgebung sehr wohl.	
Lernende	Antwort zu den Items
Die Lernenden fühlen sich in der Schulumgebung sehr wohl.	Lernende arbeiten sehr gut miteinander zusammen, helfen einander und pflegen einen respektvollen Umgang miteinander.
Die Lernenden fühlen sich in der Schulumgebung wohl.	Lernende arbeiten gut miteinander zusammen und behandeln einander respektvoll.
Die Lernenden fühlen sich in der Schulumgebung mäßig wohl.	Lernende arbeiten häufig miteinander zusammen und pflegen meistens einen respektvollen Umgang miteinander.
Die Lernenden fühlen sich in der Schulumgebung nicht sehr wohl.	Lernende behandeln einander nicht immer respektvoll und helfen einander kaum.
Die Lernenden fühlen sich in der Schulumgebung überhaupt nicht wohl.	Lernende arbeiten nicht zusammen, helfen einander kaum und behandeln einander respektlos.
Die Lernenden fühlen sich aufgrund des schlechten Klassenklimas in der Schulumgebung unwohl, sie kommen mit den anderen Lernenden nicht gut aus.	

Tabelle 3: Version 3A der Construct Map zur Skala *relatedness*

Die Lernenden fühlen sich aufgrund des guten Lernklimas, das die Lehrperson schafft, in der Schulumgebung sehr wohl.	
Lernende	Antwort zu den Items
Die Lernenden fühlen sich in der Schulumgebung sehr wohl.	Die Lehrperson ist sehr engagiert, nimmt die Lernenden ernst und behandelt diese respektvoll.
Die Lernenden fühlen sich in der Schulumgebung wohl.	Die Lehrperson ist engagiert, nimmt die Lernenden ernst und behandelt diese respektvoll.
Die Lernenden fühlen sich in der Schulumgebung mäßig wohl.	Die Lehrperson behandelt die Lernenden meistens respektvoll, manche Fragen der Lernenden werden nicht adäquat beantwortet.
Die Lernenden fühlen sich in der Schulumgebung nicht so wohl.	Die Lehrperson handelt manchmal, als wäre sie den Lernenden überlegen und behandelt diese nicht immer respektvoll.
Die Lernenden fühlen sich in der Schulumgebung nicht wohl.	Die Lehrperson handelt, als wäre sie den Lernenden überlegen und pflegt keinen respektvollen Umgang mit ihnen.
Die Lernenden fühlen sich in der Schulumgebung unwohl, da die Lehrperson sie nicht respektvoll behandelt und ihre Fragen und Aussagen nicht ernst nimmt.	

Tabelle 4: Version 3B der Construct Map zur Skala relatedness

Items - Version 3

In der Expertenrunde wurde außerdem ein Itempanel (Wilson, 2005) durchgeführt, wobei die Fachdidaktikerinnen und Fachdidaktiker auf Basis der Befragungen der Lernenden und der zweiten Version der Construct Map selbst Items erstellten. Diese Vorgangsweise wird von Wilson zur Verbesserung der Items vorgeschlagen. Die so erstellten Items konnten in vier Gruppen unterteilt werden: Klassenklima, Gruppenarbeit, Items die Lehrperson betreffend und Items die Schülerinnen und Schüler betreffend. Das Itempanel wurde durchgeführt, um die bereits entwickelten Items mit den Vorschlägen der Fachdidaktikerinnen und Fachdidaktikern zu vergleichen und so eine Expertenvalidierung zu erhalten.

Im Rahmen der Diskussion wurde außerdem das Forschungsdesign an sich besprochen. In Frage gestellt wurde, ob es überhaupt sinnvoll ist eine Intervention durchzuführen, da diese die Schülerinnen und Schüler bei der Beantwortung der Items beeinflussen könnte und zeitaufwändig ist. Falls entschieden wird eine Intervention durchzuführen, sollte diese nicht allzu lange dauern, da durchaus mit einer Testzeit von ca. 30 Minuten gerechnet werden muss, wenn alle Skalen zusammen getestet werden sollen.

Für den Fall, dass keine Intervention durchgeführt wird, könnte man beispielsweise folgende Zusatzfragen am Ende des Fragebogens stellen: „Wie oft hast du schon in Gruppen gearbeitet?“ oder „Beschreibe eine Situation, an die du beim Ausfüllen gedacht hast!“.

Für das Forschungsdesign ergaben sich die folgenden drei Möglichkeiten:

Methode 1):

Zuerst werden nur die Items der Skala *relatedness* alleine getestet, um die besten Items auszuwählen. Nach der ersten Testung werden die Schülerinnen und Schüler zu der Verständlichkeit der Items befragt, sie sollen ihre Meinung beziehungsweise Verbesserungsvorschläge zu den Items äußern. Nach der Überarbeitung testet man alle Items der sieben Skalen zusammen, ohne zuvor eine Intervention durchgeführt zu haben. Dazu müssen sich die Lernenden jedoch auf eine zuvor durchgeführte Tätigkeit beziehen können, da diese den anderen Skalen zugrunde liegt. Dies müsste vorher mit den Lehrkräften besprochen werden.

Methode 2):

In der Pilotierung werden nur die Items der Skala *relatedness* alleine getestet, um die besten Items auszuwählen. Im nächsten Schritt wird eine verkürzte Version der Intervention in allen Klassen durchgeführt, damit sich vor allem die jüngeren Schülerinnen und Schüler bei ihren Antworten auf eine Situation beziehen können. Danach erfolgt das Ausfüllen des Fragebogens, welcher nur aus zwei bis drei Skalen besteht, wobei die Skala *relatedness* immer dabei ist.

Methode 3):

Es werden wiederum zuerst die Items der Skala *relatedness* alleine getestet, um die besten Items auszuwählen. Im nächsten Schritt wird in der Hälfte der Klassen eine verkürzte Version der Intervention durchgeführt, in der anderen Hälfte wird nur der Fragebogen ausgeteilt. Auch hier soll der Fragebogen nur aus zwei bis drei Skalen bestehen, wobei die Skala *relatedness* immer dabei ist. Bei dieser Möglichkeit kann zusätzlich untersucht werden, wie und ob sich eine Intervention auf die Beantwortung des Fragebogens auswirkt.

Die folgenden neuen Items wurden mithilfe der Vorschläge aus der Diskussion der Expertenrunde erstellt. Die Items wurden auf zwei Gruppen aufgeteilt: Items, die Lehrperson betreffend und Items, die Schülerinnen und Schüler betreffend.

Items die Lehrperson betreffend:

relL1 Meine Fragen werden im Unterricht ernst genommen und die Lehrperson versucht sie immer zu beantworten.

relL2 Wenn ich aufzeige, sieht mich die Lehrperson an, auch wenn sie mich nicht dran nehmen kann.

relL3 Die Lehrperson hört mir zu, wenn ich etwas im Unterricht sage.

relL4 Die Lehrperson gibt uns nicht genügend Zeit zum Nachdenken/Überlegen. (R)

relL5 Die Lehrperson lässt uns in Kleingruppen arbeiten.

relL6 Die Lehrperson unterstützt mich, meine Sicht zu formulieren.

relL7 Wenn ich etwas nicht verstehe, traue ich mich bei der Lehrperson nachzufragen.

relL8 Ich bin mit der Beurteilung/Notengebung durch die Lehrperson zufrieden.

relL9 Es kommen immer dieselben Schüler dran. (R)

relL10 Jeder, der etwas sagen will, kommt zu Wort.

relL11 Manchmal lerne ich mehr von meinen Mitschülern, als von der Lehrperson. (R)

relL12 Ich werde von der Lehrperson ernst genommen und respektvoll behandelt.

relL13 Die Lehrperson beachtet mich nicht, wenn ich etwas zum Unterricht sagen will. (R)

relL14 Es gibt Regeln, wie man sich zu Wort meldet und diese werden eingehalten.

relL15 Die Lehrperson gibt mir genügend Zeit zum Nachdenken/Überlegen.

Items die Schülerinnen und Schüler betreffend:

relS1 Wenn ich im Unterricht etwas sage, kann ich ausreden, ohne dass mich jemand unterbricht.

relS2 Wir hören einander zu.

relS3 Wir helfen einander in der Klasse.

relS4 Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.

relS5 Ich werde von meinen Mitschülern ausgelacht, wenn ich eine Frage stelle oder meine Meinung äußere. (R)

relS6 Meine Mitschüler behandeln mich mit Respekt.

relS7 Strukturierter Unterricht ist für mich wichtig.

relS8 Meine Mitschüler sind gemein zueinander. (R)

relS9 Bei einer Gruppenarbeit teilen wir die Arbeit auf alle auf.

- relS10* Ich kann mich nur schwer konzentrieren, wenn ich in einer Gruppe arbeite. (R)
- relS11* Ich arbeite lieber alleine. (R)
- relS12* In meiner Klasse lachen wir oft miteinander.
- relS13* Ich traue mich im Unterricht Fragen zu stellen.
- relS14* In der Gruppe teilen wir die Arbeit auf alle auf.
- relS15* In einer Gruppenarbeit stören die anderen nur und tratschen. (R)
- relS16* Zusammen mit anderen kann ich bessere Ergebnisse erzielen.

3.2.5. Finale Version der Items

Die finale Version der Items ergab sich nach der Pilotierung, der kursiv gesetzte Text wurde umgeschrieben beziehungsweise ergänzt. Item *relS14* wurde gestrichen, da es zu Item *relS9* sehr ähnlich ist. Mit dieser Version der Items wurde die Testung durchgeführt. Eine genauere Diskussion der Items erfolgt in Kapitel 4.4.

Items die Lehrperson betreffend:

- relL1* Meine Fragen werden im Unterricht ernst genommen und die Lehrperson versucht sie immer zu beantworten.
- relL2* Die Lehrperson merkt, dass ich aufzeige, auch wenn sie mich nicht dran nehmen kann.
- relL3* Die Lehrperson hört mir zu, wenn ich etwas im Unterricht sage.
- relL4* Die Lehrperson gibt uns nicht genügend Zeit zum Nachdenken/Überlegen. (R)
- relL5* Die Lehrperson lässt uns in Kleingruppen arbeiten.
- relL6* Die Lehrperson unterstützt mich, meine Sicht/*Meinung* zu formulieren.
- relL7* Wenn ich etwas nicht verstehe, traue ich mich bei der Lehrperson nachzufragen.
- relL8* Ich bin mit der Beurteilung (Notengebung) durch die Lehrperson zufrieden.
- relL9* Es kommen immer dieselben Schüler dran. (R)
- relL10* Jeder, der etwas sagen will, kommt zu Wort.
- relL11* Manchmal lerne ich mehr von meinen Mitschülern, als von der Lehrperson. (R)
- relL12* Ich werde von der Lehrperson ernst genommen und respektvoll behandelt.
- relL13* Die Lehrperson beachtet mich nicht, wenn ich etwas zum Unterricht sagen will. (R)
- relL14* Es gibt Regeln, wie man sich zu Wort meldet und diese werden eingehalten.
- relL15* Die Lehrperson gibt mir genügend Zeit zum Nachdenken/Überlegen.

Items die Schülerinnen und Schüler betreffend:

relS1 Wenn ich im Unterricht etwas sage, kann ich ausreden, ohne dass mich jemand unterbricht.

relS2 Wir hören einander zu.

relS3 Wir helfen einander in der Klasse.

relS4 Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.

relS5 Ich werde von meinen Mitschülern ausgelacht, wenn ich eine Frage stelle oder meine Meinung äußere. (R)

relS6 Meine Mitschüler behandeln mich mit Respekt.

relS7 Strukturierter Unterricht ist für mich wichtig.

relS8 Meine Mitschüler sind gemein zueinander. (R)

relS9 Bei einer Gruppenarbeit teilen wir die Arbeit auf alle auf.

relS10 Ich kann mich nur schwer konzentrieren, wenn ich in einer Gruppe arbeite. (R)

relS11 Ich arbeite lieber alleine. (R)

relS12 In meiner Klasse lachen wir oft miteinander.

relS13 Ich traue mich im Unterricht Fragen zu stellen.

~~*relS14* In der Gruppe teilen wir die Arbeit auf alle auf.~~

relS15 In einer Gruppenarbeit stören die anderen nur und tratschen. (R)

relS16 Zusammen mit anderen kann ich bessere Ergebnisse erzielen.

3.3. Planung der Intervention

Der Fokus bei der Planung der Intervention liegt auf dem gemeinsamen Arbeiten der Schülerinnen und Schüler. Es werden Methoden gewählt, bei denen die Lernenden miteinander interagieren. Das Thema der Intervention sind Elementarmagnete. Bei der Planung wurde darauf geachtet, dass die Intervention maximal 40 Minuten dauert.

3.3.1. Didaktische Analyse¹

Im Nachfolgenden wird Schülerinnen und Schüler mit „SuS“ abgekürzt.

Unterrichtsplanung zum Thema „Magnetismus“, 3./4. Klasse Unterstufe	
Schlüsselbegriffe der Unterrichtseinheit	magnetische Pole, Elementarmagnete
Grobziele - Was sollen die SuS am Ende der Unterrichtseinheit wissen/können?	Die SuS wissen dass sich ungleichnamige Pole anziehen und gleichnamige Pole abstoßen. ... dass es sich bei magnetischen Kräften um Wechselwirkungen handelt. ... dass es keine magnetischen Monopole gibt. ... dass die Magnetisierung von Materialien durch die Ausrichtung von Elementarmagneten bzw. Weiss-Bezirken beschrieben werden kann. Die SuS können die Bedeutung von Magnetismus für unsere Gesellschaft in einigen Bereichen erkennen und wiedergeben. ... magnetische Kräfte als Wechselwirkung beschreiben.
Warum sollen SuS das lernen? (Bedeutung für die Gesellschaft, Zukunft, Wissenschaft, für die Lernenden selbst, Interesse)	Gegenwart/Gesellschaft: Medizin (MRT), Mülltrennung, Magnetfeld der Erde (Polarlichter, Zugvögel, Kompass) Zukunft: effizientere Energieversorgung Wissenschaft: CERN, Supraleitung SchülerInnen: Kühlschrankschrankmagnete, Schließmechanismen (Handyhülle, Taschen, ...)

¹ Format aus der Lehrveranstaltung SE Begleitseminar zur Schulpraxis bei Prof. Mag. Dr. Ilse Bartosch

Lehrplan (BMB, 2000, p. 4)	4. Klasse <i>„Elektrizität bestimmt unser Leben:</i> • <i>Einsicht in den Zusammenhang zwischen elektrischer und magnetischer Energie gewinnen; Permanentmagnet und Elektromagnet; elektromagnetische Induktion“</i>
Elementare Grundideen (Key ideas) (Hewitt, 2000)	• Gleichnamige Pole stoßen sich ab, ungleichnamige Pole ziehen sich an. • Es gibt keine magnetischen Monopole. • Ein Elementarmagnet ist die kleinste Einheit eines (magnetisierbaren) Materials, in der Magnetismus auftritt.
Lernendenvorstellungen (Raduta, 2005) und (Hewitt, 2000)	• Die magnetischen Pole haben die gleiche Rolle wie die Ladungen in der Elektrizität. • Magnetismus wird allein auf eine anziehende Kraftwirkung zwischen Magnet und magnetisierbaren Stoff reduziert. Verschiedene Pole oder Richtungsabhängigkeiten werden nicht berücksichtigt.

Tabelle 5: Didaktische Analyse

3.3.2. Stundenplanung²

Zeit	Phase	Aktivität Lehrperson	Aktivität SuS
5 min	-	Begrüßung, Erklärung des Ablaufs	
3 min	Engage	- zeigt Magnete her - fragt SuS worum es sich handelt - magnetisiert Drahtstück - Fragen: Was passiert, wenn ich den Draht in der Mitte auseinanderschneide? Ist er dann noch magnetisch? Haben die beiden Teilstücke noch zwei Pole?	- beantworten die erste Frage - SuS sollen im Experiment und in einer Gruppendiskussion Antworten auf diese Fragen finden

² Format aus der Lehrveranstaltung SE Begleitseminar zur Schulpraxis bei Prof. Mag. Dr. Ilse Bartosch

7 min	Explore	- teilt die SuS in 4-5 Gruppen ein - teilt jeder Gruppe ein magnetisiertes Drahtstück, einen Seitenschneider und einen Kompass aus - beantwortet Fragen der SuS/ hilft den SuS	- überprüfen mithilfe eines Kompasses welcher Pol der Nordpol bzw. der Südpol ist - schneiden den Draht in der Mitte durch und überprüfen die zuvor gestellten Fragen - versuchen in der Gruppe eine Erklärung für ihre Beobachtung zu finden
20 min	Explain	- teilt das Arbeitsblatt aus - geht in der Klasse herum und beantwortet Fragen der SuS	- bearbeiten gemeinsam in der Gruppe das Arbeitsblatt - stellen Fragen
5 min	Elaborate	- leitet folgendes Spiel an (Bresler, Heepmann, & Hörter, 2012): Alle SuS stellen sich im Klassenraum auf. Die SuS strecken die Arme aus und jede bzw. jeder symbolisiert einen Elementarmagneten. Die linke Hand steht für den Nordpol, die rechte Hand für den Südpol. Zuerst stehen die SuS in einer beliebigen Richtung. L: Ihr stellt jetzt ein nicht magnetisches Material dar. Wie müsstet ihr euch hinstellen, dass ihr ein magnetisches Material darstellt? Eventuell noch die SuS so wie sie stehen in der Mitte teilen, als Analogon zum Drahtstück.	- nehmen am Spiel teil
15 min	-	Klärung eventueller Fragen	Ausfüllen des Fragebogens

Tabelle 6: Stundenplanung

Das Spiel der Elaborate-Phase dient als Puffer, falls die Bearbeitung des Arbeitsblattes schneller als gedacht erfolgt, wenn die Zeit zu knapp ist, wird es nicht durchgeführt. Das Arbeitsblatt befindet sich in Anhang B, 11.2.

3.4. Erstellung der Fragebögen

Der Fragebogen der Pilotierung umfasste alle 31 neu erstellten Items der Skala *relatedness* (dritte Version der Items). Nachdem die Schülerinnen und Schüler den Fragebogen ausgefüllt hatten, wurden sie in einem offenen Gespräch befragt, ob es Unklarheiten gab oder ob ihnen gegebenenfalls Fehler aufgefallen sind. Während des Ausfüllens gab es bereits Fragen, auf Basis derer wurde ein Item gestrichen und es wurden kleine Änderungen vorgenommen, welche der finalen Version der Items entsprechen (siehe Kapitel 3.2.5.). Eine genauere Diskussion dieser Schwierigkeiten erfolgt in Kapitel 4.4.

Mit den 30 Items wurden nach der Pilotierung zwei verschiedene Fragebogenversionen erstellt, welche getestet werden sollen. Version 1 ist der Fragebogen, der im regulären Unterrichtsverlauf getestet wird, Version 2 wird nach der zuvor durchgeführten Intervention ausgefüllt. Die beiden Versionen unterscheiden sich lediglich durch den Einleitungstext, die Reihung der Items ist gleich.

Version 1: Skala *relatedness* ohne Intervention (30 Items)

Version 2: Skala *relatedness* mit Intervention (30 Items)

Da sich alle anderen Skalen, bis auf die Skala *relatedness*, stark auf eine zuvor durchgeführte Tätigkeit beziehen, wurde entschieden die Skala alleine zu testen. Wäre eine weitere Skala hinzugenommen worden, so hätte in allen Klassen eine Intervention oder von den Lehrkräften eine Gruppenarbeit durchgeführt werden müssen.

Die Beantwortung der Items erfolgt mit einer 5-stufigen Likert-Skala, anders als in der Literatur vorgeschlagen mit einer 7-stufigen Likert-Skala. Grund dafür ist das Alter der Schülerinnen und Schüler. In der siebten und achten Schulstufe dürfte die 7-stufige Likert-Skala eine zu feine Unterteilung darstellen, zwischen der die Lernenden wahrscheinlich nur schwer differenzieren können (Schmidt, 2017).

Für die Fragebögen wurden die Items durchmischt, die Zuordnung welche Nummer des Fragebogens welchem Item entspricht, ist in Tabelle 22, Anhang B zu sehen. Im Anhang B befinden sich weiters Version 1 und Version 2 des Fragebogens.

3.5. Durchführung an den Schulen

Die Durchführung der Testung erfolgte nach Methode 3 (siehe Kapitel 3.2.4.), jedoch wurde beschlossen keine weitere Skala mitzutesten.

Die Testung gliederte sich in zwei Teile. Die eine Gruppe von Schülerinnen und Schülern bekam im regulären Unterrichtsverlauf Fragebogenversion 1, die andere Gruppe füllte Fragebogenversion 2 nach der zuvor durchgeführten Intervention aus.

Die Testung erfolgte an fünf verschiedenen Schulen in Wien, die aufgrund der Anonymität in weiterer Folge mit Buchstaben bezeichnet werden. Dabei handelt es sich um vier Gymnasien und um eine Neue Mittelschule. Tabelle 7 gibt eine Übersicht, welche Fragebögen an welcher Schule getestet wurden beziehungsweise welche Version wo verwendet wurde. Insgesamt wurden die Fragebögen in zwölf Klassen, davon sechs Klassen der 7. Schulstufe und sechs Klassen der 8. Schulstufe, getestet. Inklusive Pilotierung wurden die Fragebögen von 284 Schülerinnen und Schülern ausgefüllt.

	Pilotierung	Version 1	Version 2
Schule A	FB#1 - #21	FB#216 - #238	FB#196 - #215
Schule B	FB#22 - #47	FB#75 - #103 FB#239 - #262	-
Schule C	-	-	FB#48 - #74
Schule D	-	FB#128 - #150 FB#151 - #172	FB#104 - #127 FB#173 - #195
Schule E	-	-	FB#263 - #284

Tabelle 7: Übersicht der Schulen und Fragebogenversionen

Die Bezeichnung „FB#“ steht für die Nummer des Fragebogens, beispielsweise „FB#1“ bedeutet Fragebogen Nummer 1. Die Gesamtanzahl an ausgefüllten Fragebögen der Pilotierung und der beiden Versionen ist die folgende:

Pilotierung: N = 47

Version 1: N = 121

Version 2: N = 116

In Schule A wurden drei Klassen getestet, eine 7. Schulstufe (FB#196 - #215) und zwei 8. Schulstufen (FB#1 - #21, FB#216 - #238), in Schule B drei 7. Schulstufen (FB#22 - #47, FB#75 - #103, FB#239 - #262), in Schule C eine 8. Schulstufe (FB#48 - #74), in Schule D zwei 7. Schulstufen (FB#128 - #150, FB#173 - #195) und zwei 8. Schulstufen (FB#104 - #127, FB#151 - #172) und in Schule E eine 8. Schulstufe (FB#263 - #284).

Die Testung sowie die Interventionen wurden ausschließlich von der Autorin dieser Arbeit durchgeführt. Es wurde versucht, den gleichen Ablauf einzuhalten sowie die gleiche Wortwahl zu benutzen. Die Stundenplanung für die Klassen mit der Intervention konnte, wie in Kapitel 3.3.2. beschrieben, eingehalten werden, das Spiel wurde jedoch in keiner Klasse gebraucht. Da die Stundenplanung für die siebte und für die achte Schulstufe gemacht wurde, bestand die Möglichkeit, dass Schülerinnen und Schüler der achten Schulstufe die Aufgaben und das Arbeitsblatt eventuell als zu leicht empfinden. Dies bewahrheitete sich jedoch nicht, da die Thematik bei einigen Lernenden anscheinend nicht mehr so präsent war oder die Konzepte von der Lehrperson noch nicht besprochen wurden. Festgestellt konnte dies anhand der Fragen der Lernenden und anhand der Bearbeitungsdauer des Arbeitsblattes werden. Somit waren der Aufwand und die Schwierigkeit für die siebte und die achte Schulstufe im Durchschnitt etwa gleich. An der Stundenplanung musste im Laufe der Testung nichts verändert werden, da der Zeitplan stets eingehalten werden konnte. Vereinzelt wurde es jedoch zeitlich etwas knapp, da manche Schülerinnen und Schüler sich intensiv mit dem Arbeitsblatt beschäftigt hatten und dieses unbedingt fertig machen wollten. Im Durchschnitt nahm das Ausfüllen des Fragebogens circa sieben Minuten in Anspruch, was kürzer als angenommen war.

3.6. Auswertung der Fragebögen

Die Auswertung der Fragebögen erfolgte mithilfe des Statistik- und Analyseprogramms SPSS von IBM. Der erste Schritt bestand darin, die Fragebögen durchnummerieren und die Antworten der Schülerinnen und Schüler in die Tabelle von SPSS einzutragen. Die Werte gehen von eins bis fünf, wobei 5 eine hohe Zustimmung bedeutet und 1 gar keine Zustimmung. Für fehlende oder unklare Antworten wurde die Zahl 99 verwendet. Die negativ gepolten Items, im Text davor mit (R) bezeichnet, wurden umkodiert.

Umkodiert wurde wie folgt:

5 → 1

4 → 2

3 → 3

2 → 4

1 → 5

Kreuzte eine Schülerin oder ein Schüler beispielsweise bei Item *relL4* „Die Lehrperson gibt uns nicht genügend Zeit zum Nachdenken/Überlegen.“ *stimmt gar nicht* an, was mit 1 kodiert wurde, so wurde die 1 in eine 5 umkodiert.

Danach konnte mit der Auswertung begonnen werden, welche in Kapitel 4 beschrieben und diskutiert wird.

Zur Auswertung konnten alle Fragebögen herangezogen werden. Manche Schülerinnen und Schüler schrieben Anmerkungen auf den Fragebogen oder direkt zu den Items, diese werden in Kapitel 4.4. genauer diskutiert.

4) Ergebnisse

In diesem Kapitel werden die Ergebnisse der Studie genauer beschrieben, es folgen inhaltliche Analysen der Ergebnisse der neu erstellten Skala *relatedness* sowie der Bezug zu den anderen Skalen. Unter anderem werden Schwierigkeiten mit den Items sowie Anmerkungen der Schülerinnen und Schüler diskutiert und die Qualität des erstellten Fragebogens wird besprochen.

Für die Auswertung wurden die Daten der Pilotierung nicht verwendet.

4.1. Analyse der Skala *relatedness*

4.1.1. KMO- und Bartlett-Test

Bevor eine Faktorenanalyse gemacht wurde, erfolgte die Bestimmung des Kaiser-Meyer-Olkin-Koeffizients (KMO-Koeffizient), weiters wurde der Bartlett-Test auf Sphärizität durchgeführt. Anhand der Werte dieser beiden Tests kann entschieden werden, ob eine Faktorenanalyse durchgeführt werden kann und ob diese sinnvoll ist.

Der KMO-Koeffizient lässt Aussagen zu, ob die Items für eine Faktorenanalyse geeignet sind, dazu wird die Höhe der Korrelationen in der Korrelationsmatrix analysiert (Schmidt, 2017). Laut Bühner kann ein KMO-Koeffizient zwischen 0,60 und 0,69 als mäßig klassifiziert werden, zwischen 0,70 und 0,79 als mittel, zwischen 0,80 und 0,89 als gut und größer 0,90 als sehr gut.

Der Bartlett-Test ist ein Signifikanztest. Er wird signifikant, wenn die Items korreliert sind, sie eignen sich dann für eine Faktorenanalyse. Signifikant ist die Prüfgröße für $p < 0,05$. In diesem Fall weichen die Korrelationen signifikant von null ab. Der Test hängt von der Stichprobengröße ab, größere Stichproben sind immer empfehlenswert (Bühner, 2011).

Der KMO-Koeffizient und der Bartlett-Test der erhobenen Daten sind in Tabelle 8 zu sehen.

KMO- und Bartlett-Test		
Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin.		,827
Bartlett-Test auf Sphärizität	Ungefähres Chi-Quadrat	2074,774
	df	435
	Signifikanz nach Bartlett	,000

Tabelle 8: KMO-Koeffizient und Bartlett-Test der *relatedness* Items

Für die 237 zur Analyse verwendeten Fragebögen (FB#48 - FB#284) ergab sich ein KMO-Koeffizient von 0,827, welcher laut Bühner als gut klassifiziert werden kann.

Der Bartlett-Test erweist sich mit einer Signifikanz von $p < 0,001$ als höchstsignifikant, was auf eine hohe Korrelation der Items schließen lässt. Auf Basis dieser Ergebnisse kann mit den Daten eine Faktorenanalyse durchgeführt werden.

4.1.2. Faktorenanalyse der Skala *relatedness*

Die 30 Items der Skala *relatedness* wurden einer exploratorischen Faktorenanalyse unterzogen, wobei mithilfe der rotierten Komponentenmatrix Items mit geringer Ladung iterativ ausgeschlossen wurden. In Tabelle 23, Anhang B, sind die Ladungen aller Items auf die Skala *relatedness* dargestellt, in Tabelle 9 die ausgewählten Items. Die Items wurden aufgrund ihrer Faktorladung und unter Berücksichtigung des Inhaltes ausgewählt. Im Prozess der Elimination wurde iterativ immer ein Item weggelassen, danach wurde wieder eine Faktorenanalyse durchgeführt und die neuen Faktorladungen wurden begutachtet. Ergaben sich schlechtere Ladungen als zuvor, so wurde das Item wieder hinzugenommen, verbesserten sich die Ladungen, wurde das Item weggelassen.

Rotierte Komponentenmatrix^a		
	Komponente	
	1	2
relL1	,795	
relS2	,313	,679
relL3	,802	
relS3		,728
relS4		,821
relS5		,660
relL6	,672	
relS6		,776
relL10	,724	
relL12	,773	
Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung. ^a a. Die Rotation ist in 3 Iterationen konvergiert.		

Tabelle 9: Rotierte Komponentenmatrix der neu erstellten Skala *relatedness*

Als Resultat dieses Prozesses wurden die Items *relL1*, *relS2*, *relL3*, *relS3*, *relS4*, *relS5*, *relL6*, *relS6*, *relL10* und *relL12* ausgewählt. Die Ladungen der Items sind Tabelle 9 zu entnehmen. Die Items laden mit Werten über 0,6 auf eine Skala, Querladungen kleiner als 0,25 wurden der Übersicht halber unterdrückt. Alle ausgewählten Items weisen eine Querladung kleiner 0,25 auf, außer Item *relS2*, welches eine Querladung von 0,313 aufweist.

Weiters wurde mittels Kaiser-Kriterium (Tabelle 10) überprüft, in wie viele Faktoren die Skala demnach zerfällt. Zwei der Eigenwerte sind über 1, wobei nach dem Eigenwertkriterium eine 2-faktorielle Lösung angemessen ist (Bühner, 2011). Zwei Faktoren erklären zusammen 58,068% der Gesamtvarianz der Items, was für diese Arbeit als gut und zufriedenstellend angesehen werden kann. Bei diesen beiden Faktoren handelt es sich, wie in der Theorie (Kapitel 3.2.4) beschrieben, um die soziale Interaktion der Lernenden untereinander und um die Interaktion von Lernenden und Lehrperson.

Beim Screeplot (Abbildung 3) ist auf der x-Achse die Anzahl der Faktoren aufgetragen und auf der y-Achse der dazugehörige Eigenwert. Bei dieser Extraktionsmethode wird nach einem bedeutsamen Eigenwertabfall gesucht, wobei nur die Werte vor dem „Knick“ betrachtet werden (Bühner, 2011). Der Screeplot für die Items weist einen Knick bei Faktor 3 auf, womit auch dieser Test bestätigt, dass zwei Faktoren für diese Items gerechtfertigt sind.

Erklärte Gesamtvarianz									
Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen		
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	3,843	38,426	38,426	3,843	38,426	38,426	3,008	30,084	30,084
2	1,964	19,642	58,068	1,964	19,642	58,068	2,798	27,984	58,068
3	,811	8,111	66,179						
4	,719	7,195	73,374						
5	,557	5,565	78,939						
6	,495	4,951	83,890						
7	,472	4,721	88,612						
8	,418	4,184	92,796						
9	,383	3,826	96,622						
10	,338	3,378	100,000						
Extraktionsmethode: Hauptkomponentenanalyse.									

Tabelle 10: Erklärte Gesamtvarianz der Skala relatedness

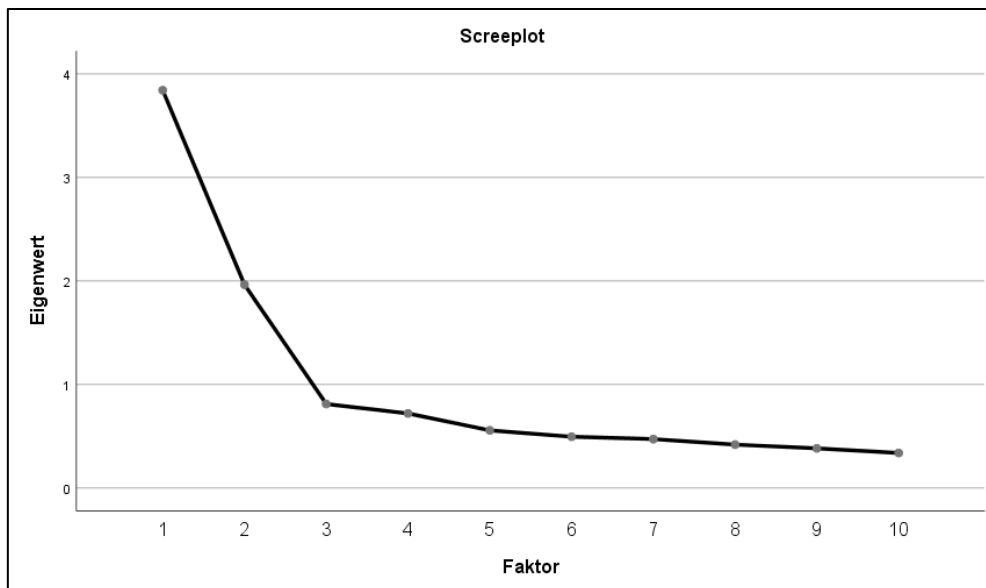


Abbildung 3: Screeplot der Faktoren der Skala relatedness

Tabelle 11 stellt die Item-Skala-Statistik dar, welche aus der Trennschärfenanalyse erhalten wurde. Aus der Tabelle kann man die Trennschärfe (korrigierte Item-Skala-Korrelation) der einzelnen Items entnehmen, die Werte befinden sich im Bereich von 0,320 bis 0,591. Die Trennschärfe von vier Items kann als mittel (0,30 - 0,50) eingestuft werden, sechs Items weisen eine hohe Trennschärfe ($> 0,50$) auf (Bühner, 2011, p. 81). Die Trennschärfe sollte generell größer als 0,3 sein, da das Item sonst möglicherweise ein anderes Konstrukt misst. Ist die Trennschärfe zu hoch (nahe bei eins) wären die Items redundant, man würde keine zusätzlichen Informationen aus den anderen Items bekommen (Bühner, 2011). Die Trennschärfen der ausgewählten Items können somit als gut gewertet werden.

	Item-Skala-Statistiken			
	Skalenmittelwert , wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item- Skala- Korrelation	Cronbachs Alpha, wenn Item weggelassen
relL1	32,38	35,882	,522	,796
relS2	34,07	34,142	,591	,788
relL3	32,20	36,925	,506	,799
relS3	33,56	34,549	,499	,798
relS4	33,15	33,287	,585	,788
relS5	32,85	36,973	,320	,818
relL6	32,64	36,854	,401	,808
relS6	33,31	34,730	,521	,796
relL10	33,27	33,800	,568	,790
relL12	32,38	36,138	,466	,801

Tabelle 11: Item-Skala-Statistiken für die Skala relatedness

Der Wert des Cronbach-Alphas für die ausgewählten zehn Items der neuen Skala *relatedness* beträgt 0,815. Da das Cronbach-Alpha ein Maß für die Reliabilität ist, kann an dieser Stelle bereits gesagt werden, dass die Skala mit guter Genauigkeit die soziale Eingebundenheit misst.

4.1.3. Faktorenanalyse der Gruppenarbeit-Items

Da die Gruppenarbeit-Items in der obigen Analyse aufgrund der geringen Ladung ausgeschlossen wurden, wurde eine exploratorische Faktorenanalyse nur mit diesen Items durchgeführt, um zu überprüfen, ob diese einen eigenen Faktor bilden. Der KMO-Koeffizient für die Items fällt mit 0,658 mäßig aus, die Korrelation der Items ist mit $p < 0,001$ jedoch höchstsignifikant. Aufgrund dieser Werte kann eine Faktorenanalyse durchgeführt werden.

KMO- und Bartlett-Test		
Maß der Stichprobeneignung nach Kaiser-Meyer-Olkin.		,658
Bartlett-Test auf Sphärizität	Ungefähres Chi-Quadrat	140,030
	df	10
	Signifikanz nach Bartlett	,000

Tabelle 12: KMO-Koeffizient und Bartlett-Test für die Gruppenarbeit-Items

Auf Basis des Kaiser-Kriteriums kann aus Tabelle 13, der erklärten Gesamtvarianz der Gruppenarbeit-Items, entnommen werden, dass ein Faktor 40,509% der Gesamtvarianz der Items erklärt.

Erklärte Gesamtvarianz				Summen von quadrierten Faktorladungen für Extraktion		
Komponente	Anfängliche Eigenwerte					
	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2,025	40,509	40,509	2,025	40,509	40,509
2	,958	19,152	59,662			
3	,915	18,291	77,953			
4	,592	11,841	89,794			
5	,510	10,206	100,000			
Extraktionsmethode: Hauptkomponentenanalyse.						

Tabelle 13: Erklärte Gesamtvarianz der Gruppenarbeit-Items

Für die Gruppenarbeit-Items wurde bei der Faktorenanalyse einmal ein Faktor extrahiert und einmal zwei Faktoren, die Ergebnisse sind in Tabelle 14 und in Tabelle 15 zu sehen. Da bei dieser Analyse nur ein Eigenwert größer als eins ist und die Extrahierung zweier Faktoren

keine besonders guten Ergebnisse liefert, können die Gruppenarbeit-Items zu einem Faktor zusammengefasst werden. Die Ladungen auf die Skala sind mit 0,396 bis 0,748 in Ordnung. Item *relS11* weist eine negative Ladung auf, trotz nochmals überprüfter, richtiger Polarität.

Komponentenmatrix^a	
Komponente	
1	
relS9	,396
relS10	,748
relS11	-,745
relL15	,632
relS15	,595
Extraktionsmethode: Hauptkomponentenanalyse. a. 1 Komponenten extrahiert	

Tabelle 14: Komponentenmatrix der Gruppenarbeit-Items, ein Faktor extrahiert

Rotierte Komponentenmatrix^a		
Komponente		
	1	2
relS9		,908
relS10	,682	,311
relS11	-,836	
relL15	,440	,519
relS15	,657	
Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung. ^a		

Tabelle 15: Rotierte Komponentenmatrix der Gruppenarbeit-Items, zwei Faktoren extrahiert

Die Gruppenarbeit-Items bilden alleine einen Faktor, sie scheinen jedoch zur Skala *relatedness* nicht zu passen. Im Zuge der Faktorenanalyse aller 30 Items wurde außerdem versucht drei Faktoren zu extrahieren, was jedoch zu keinem besonders guten Ergebnis führte. Daraus lässt sich schließen, dass die optimale Anzahl an Faktoren für die Skala *relatedness* zwei Faktoren sind und die Gruppenarbeit-Items für diese nicht so relevant sind, wie Items mit anderem Inhalt.

4.2. Einfluss der Intervention

In allen folgenden Analysen wurde ein Signifikanzniveau von 0,05 gewählt. Das heißt, ab einer Signifikanz kleiner oder gleich 0,05 wird die Nullhypothese H_0 abgelehnt und die Alternativhypothese H_1 angenommen. Werte kleiner oder gleich 0,001 werden als höchstsignifikant gewertet (Rasch, 2014).

Um den Einfluss der Intervention zu untersuchen, wurde ein t-Test mit den ausgewählten zehn Items durchgeführt. Mithilfe des t-Tests soll festgestellt werden, ob sich die Daten der beiden Fragebogenversionen systematisch voneinander unterscheiden, beziehungsweise ob sich eine zuvor durchgeführte Intervention auf die Beantwortung des Fragebogens auswirkt (Rasch, 2014). Tabelle 16 zeigt die Gruppenstatistik der ausgewählten Items, FB_Version 1 steht für Fragebogen Version 1, der Fragebogen, bei dem keine Intervention durchgeführt wurde. Bei FB_Version 2 wurde die Intervention vor der Beantwortung des Fragebogens durchgeführt.

Gruppenstatistiken					
	FB_Version	N	Mittelwert	Std.- Abweichung	Standardfehler des Mittelwertes
relL1	1	121	4,28	1,010	,092
	2	116	4,22	,921	,086
relS2	1	121	2,55	1,118	,102
	2	114	2,63	1,050	,098
relL3	1	121	4,50	,828	,075
	2	116	4,41	,823	,076
relS3	1	121	2,99	1,235	,112
	2	115	3,22	1,082	,101
relS4	1	120	3,34	1,170	,107
	2	115	3,67	1,226	,114
relS5	1	121	3,71	1,143	,104
	2	114	3,91	1,141	,107
relL6	1	121	4,21	,894	,081
	2	114	3,81	1,080	,101
relS6	1	121	3,21	1,125	,102
	2	115	3,53	1,071	,100
relL10	1	120	3,34	1,156	,106
	2	115	3,42	1,139	,106
relL12	1	120	4,35	,976	,089
	2	112	4,18	1,015	,096

Tabelle 16: Gruppenstatistiken der beiden Fragebogenversionen

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
relL1	Varianzen sind gleich	,112	,738	,521	235	,603	,065	,126	-,182	,313
	Varianzen sind nicht gleich			,522	234,424	,602	,065	,125	-,182	,313
relS2	Varianzen sind gleich	,364	,547	-,550	233	,583	-,078	,142	-,357	,201
	Varianzen sind nicht gleich			-,551	232,998	,582	-,078	,141	-,356	,201
relL3	Varianzen sind gleich	,088	,767	,845	235	,399	,091	,107	-,121	,302
	Varianzen sind nicht gleich			,845	234,681	,399	,091	,107	-,121	,302
relS3	Varianzen sind gleich	1,518	,219	-1,490	234	,138	-,226	,151	-,524	,073
	Varianzen sind nicht gleich			-1,495	232,499	,136	-,226	,151	-,523	,072
relS4	Varianzen sind gleich	,085	,771	-2,097	233	,037	-,328	,156	-,636	-,020
	Varianzen sind nicht gleich			-2,095	231,153	,037	-,328	,156	-,636	-,020
relS5	Varianzen sind gleich	,026	,873	-1,352	233	,178	-,202	,149	-,495	,092
	Varianzen sind nicht gleich			-1,352	232,224	,178	-,202	,149	-,495	,092
relL6	Varianzen sind gleich	4,426	,036	3,098	233	,002	,400	,129	,145	,654
	Varianzen sind nicht gleich			3,080	219,733	,002	,400	,130	,144	,655
relS6	Varianzen sind gleich	,157	,692	-2,263	234	,025	-,324	,143	-,606	-,042
	Varianzen sind nicht gleich			-2,266	233,999	,024	-,324	,143	-,605	-,042
relL10	Varianzen sind gleich	,013	,908	-,506	233	,614	-,076	,150	-,371	,219
	Varianzen sind nicht gleich			-,506	232,814	,614	-,076	,150	-,371	,219
relL12	Varianzen sind gleich	,006	,939	1,311	230	,191	,171	,131	-,086	,429
	Varianzen sind nicht gleich			1,309	227,301	,192	,171	,131	-,087	,429

Tabelle 17: t-Test der ausgewählten Items in Bezug auf die beiden Fragebogenversionen

Der Levene-Test für Varianzgleichheit zeigt, dass bis auf Item *relL6* ($p < 0,05$) Varianzhomogenität vorliegt. Somit ist die Durchführung eines t-Tests zur Analyse der Mittelwertunterschiede der beiden Fragebogenversionen gerechtfertigt.

Die zu testende Nullhypothese H_0 lautet: Es liegen keine Unterschiede in den Mittelwerten der beiden Fragebogenversionen vor.

Für die Items *relS4*, *relL6* und *relS6* unterscheiden sich die Mittelwerte der beiden Gruppen (mit vorheriger Intervention – ohne vorherige Intervention) signifikant ($p < 0,05$). Die Intervention hat somit die Beantwortung dieser drei Items beeinflusst. Für alle restlichen Items sind die Mittelwertunterschiede der beiden Gruppen nicht signifikant. Daher wird für den Großteil der Items die Nullhypothese, dass sich die Intervention nicht auf die Beantwortung der Items auswirkt, beibehalten.

Eine genauere Diskussion und Analyse der Ergebnisse erfolgt in Kapitel 5.2.

4.3. Inhaltliche Analyse der Ergebnisse

In diesem Kapitel erfolgt eine genauere Analyse und Diskussion der Ergebnisse. Weiters wird die Motivation an den verschiedenen Schulen sowie der beiden getesteten Schulstufen genauer untersucht.

4.3.1. Inhaltliche Analyse der Skala *relatedness*

Der KMO-Koeffizient der Skala fiel mit 0,827 gut aus, die Items sind mit einer Signifikanz von $p < 0,001$ hochkorreliert. Mithilfe der Ladungen einer exploratorischen Faktorenanalyse wurden die Items mit niedriger Ladung ausgeschlossen, wobei auf den Inhalt der Items geachtet wurde.

Bei den ausgewählten Items traten außerdem keine Verständnisschwierigkeiten auf, was für die Auswertung und deren Genauigkeit von Vorteil ist.

Die zehn ausgewählten Items sind hier im Überblick zu sehen:

relL1 Meine Fragen werden im Unterricht ernst genommen und die Lehrperson versucht sie immer zu beantworten.

relL3 Die Lehrperson hört mir zu, wenn ich etwas im Unterricht sage.

relL6 Die Lehrperson unterstützt mich, meine Sicht/Meinung zu formulieren.

relL10 Jeder, der etwas sagen will, kommt zu Wort.

relL12 Ich werde von der Lehrperson ernst genommen und respektvoll behandelt.

relS2 Wir hören einander zu.

relS3 Wir helfen einander in der Klasse.

relS4 Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.

relS5 Ich werde von meinen Mitschülern ausgelacht, wenn ich eine Frage stelle oder meine Meinung äußere. (R)

relS6 Meine Mitschüler behandeln mich mit Respekt.

Mit Ladungen von 0,660 bis 0,821 laden die Items gut auf die Skala *relatedness*. Die Trennschärfe der Items liegt in einem Bereich von 0,320 bis 0,591, welche laut Bühner (2011) als mittel bis hoch eingestuft werden können. Aus der Tabelle der Trennschärfeanalyse (Tabelle 11) kann weiters entnommen werden, dass das Weglassen von Item *relS5* zu einer kleinen Verbesserung des Cronbach-Alphas führen würde, dieses Item hat auch die geringste Trennschärfe.

Da das Item jedoch inhaltsmäßig sehr gut zu den anderen passt und die Verbesserung des Cronbach-Alphas nur sehr klein wäre, wurde es beibehalten.

Aus der Hauptkomponentenanalyse lässt sich entnehmen, dass wie in der Theorie erwartet, die Skala *relatedness* von zwei Faktoren gebildet wird.

Der Mittelwert für die Skala *relatedness* mit den zehn ausgewählten Items, beträgt 3,63 mit einer Standardabweichung von 0,59. In weiterer Folge kann nun untersucht werden, wie hoch die Motivation der Schülerinnen und Schüler an den verschiedenen Schulen war. In Abbildung 4 sind die Mittelwerte der Skala *relatedness* für die verschiedenen Schulen sowie der Gesamtmittelwert zu sehen. Abbildung 5 zeigt Boxplots der Motivation an den verschiedenen Schulen. Die Daten für die Boxplots der Motivation wurden aus den Werten der Skala *relatedness* (der zehn ausgewählten Items) für jede Schule mittels SPSS berechnet.

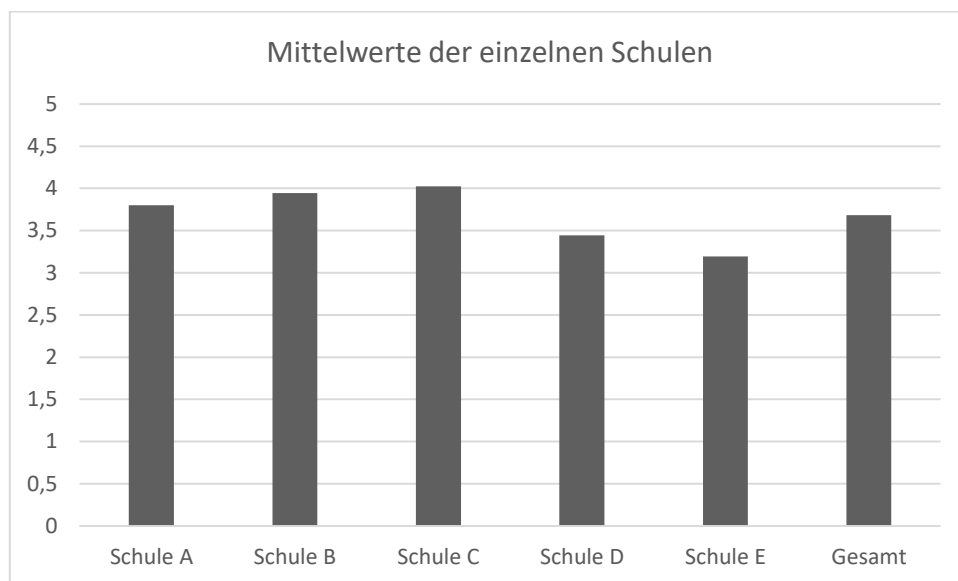


Abbildung 4: Mittelwerte der Skala *relatedness* für die verschiedenen Schulen

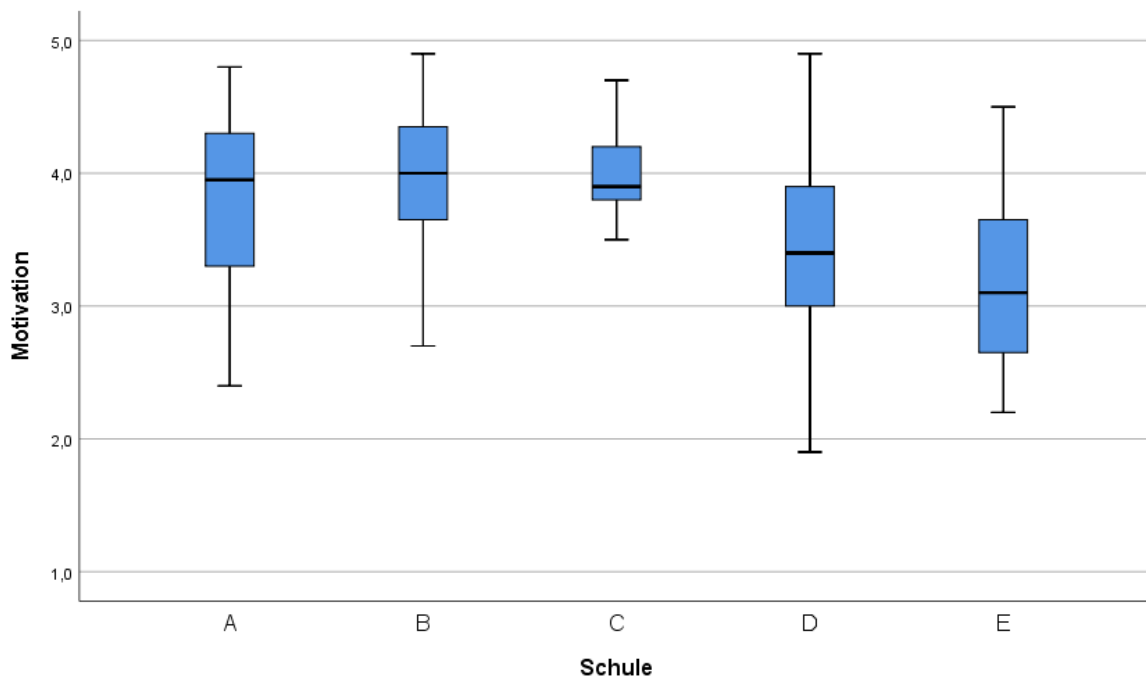


Abbildung 5: Boxplots der Motivation in den verschiedenen Schulen

Um die Unterschiede zwischen den Schulen zu untersuchen wird die folgende Nullhypothese H_0 aufgestellt: Es liegen keine Unterschiede in den Mittelwerten der verschiedenen Schulen vor.

Da es sich bei den verschiedenen Schulen um mehr als zwei disjunkte Teilgesamtheiten (Stichproben) handelt, muss zuerst überprüft werden, ob eine ANOVA (Analysis of Variance) durchgeführt werden kann, oder ob ein Kruskal-Wallis-Test gemacht werden muss.

Die Voraussetzungen für eine ANOVA sind die folgenden: Das zu untersuchende Merkmal, in diesem Fall die Motivation, muss in den Teilgesamtheiten normalverteilt mit einer homogenen Streuung sein.

Die Approximation der Motivationsverteilung der einzelnen Schulen durch eine Normalverteilung ergab Varianzen von 0,11 bis 0,51. Aufgrund dieser Varianzen liegt das Kriterium der homogenen Streuung nicht vor. Die Motivation an den verschiedenen Schulen ist eher schlecht normalverteilt. Da die Voraussetzungen für eine ANOVA unzureichend erfüllt sind, wurde ein Kruskal-Wallis-Test durchgeführt (Eckstein, 2016).

Die asymptotische Signifikanz des Kruskal-Wallis-Tests lieferte mit einem p-Wert kleiner 0,001 ein höchstsignifikantes Ergebnis. Somit wird die Nullhypothese abgelehnt, es gibt höchstsignifikante Unterschiede der Motivation in den verschiedenen Schulen.

Um diese Unterschiede genauer zu untersuchen, wurden die Schulen mittels Mann-Whitney-U-Test paarweise miteinander verglichen.

Dabei wurde festgestellt, dass Schule D und Schule E eine Gruppe weniger motivierter Schülerinnen und Schüler bilden und die Schulen A, B und C eine Gruppe stärker motivierter Lernender.

Weiters kann untersucht werden, ob sich die Motivation in den beiden Schulstufen unterscheidet. In Abbildung 6 sind die Mittelwerte der beiden Schulstufen für jedes Item dargestellt, in Abbildung 7 der Mittelwert der Skala *relatedness* für beide Schulstufen.

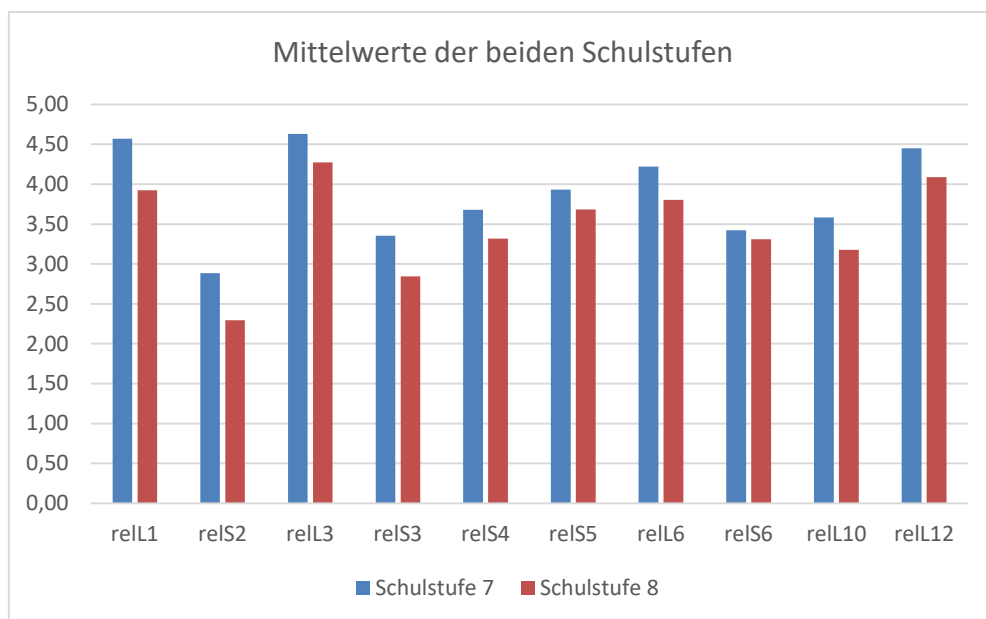


Abbildung 6: Mittelwerte der Items für die 7. und 8. Schulstufe

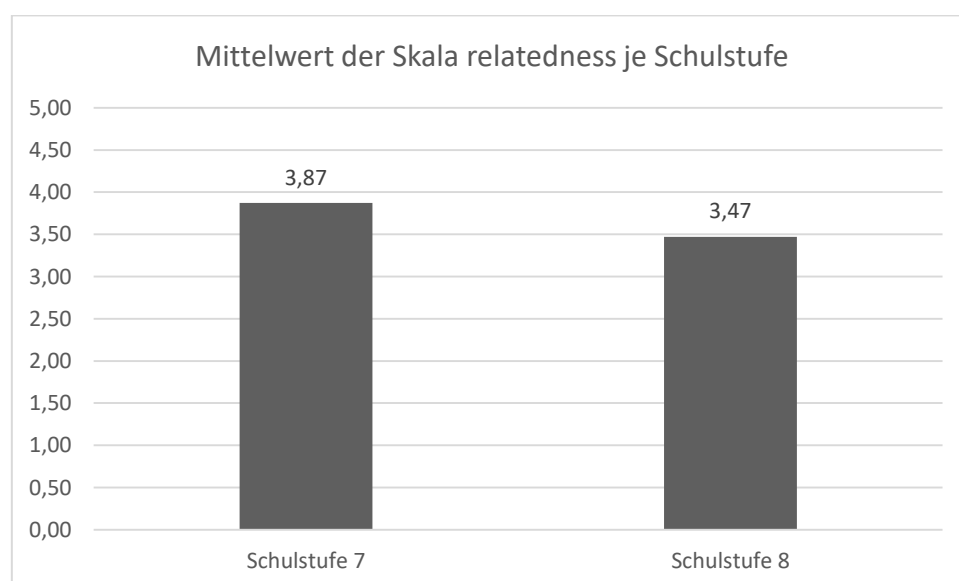


Abbildung 7: Mittelwert der Skala *relatedness* für die 7. und 8. Schulstufe

Test bei unabhängigen Stichproben										
		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit					95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
reIL1	Varianzen sind gleich	24,266	,000	5,467	235	,000	,648	,118	,414	,881
	Varianzen sind nicht gleich			5,456	189,880	,000	,648	,119	,414	,882
reIS2	Varianzen sind gleich	,195	,659	4,322	233	,000	,589	,136	,321	,858
	Varianzen sind nicht gleich			4,330	230,082	,000	,589	,136	,321	,857
reIL3	Varianzen sind gleich	13,917	,000	3,425	235	,001	,359	,105	,153	,566
	Varianzen sind nicht gleich			3,420	214,416	,001	,359	,105	,152	,566
reIS3	Varianzen sind gleich	,268	,605	3,413	234	,001	,507	,148	,214	,799
	Varianzen sind nicht gleich			3,413	233,914	,001	,507	,148	,214	,799
reIS4	Varianzen sind gleich	2,273	,133	2,319	233	,021	,362	,156	,054	,669
	Varianzen sind nicht gleich			2,316	229,340	,021	,362	,156	,054	,669
reIS5	Varianzen sind gleich	3,284	,071	1,671	233	,096	,248	,149	-,045	,541
	Varianzen sind nicht gleich			1,670	230,473	,096	,248	,149	-,045	,542
reIL6	Varianzen sind gleich	7,197	,008	3,239	233	,001	,417	,129	,163	,670
	Varianzen sind nicht gleich			3,236	222,590	,001	,417	,129	,163	,671
reIS6	Varianzen sind gleich	,064	,800	,779	234	,437	,112	,144	-,172	,397
	Varianzen sind nicht gleich			,779	233,503	,437	,112	,144	-,172	,397
reIL10	Varianzen sind gleich	2,167	,142	2,734	233	,007	,403	,147	,113	,694
	Varianzen sind nicht gleich			2,736	229,455	,007	,403	,147	,113	,694
reIL12	Varianzen sind gleich	9,906	,002	2,808	230	,005	,362	,129	,108	,616
	Varianzen sind nicht gleich			2,808	207,831	,005	,362	,129	,108	,616

Tabelle 18: t-Test der ausgewählten Items in Bezug auf die beiden Schulstufen

Mittels Levene-Test (Tabelle 18) konnte Varianzhomogenität für die beiden Gruppen, siebte und achte Schulstufe, festgestellt werden. Der durchgeführte t-Test lieferte die Ergebnisse, dass sich die Items *reIL1*, *reIS2*, *reIL3*, *reIS3* und *reIL6* höchstsignifikant voneinander unterscheiden, die Items *reIS4*, *reIL10* und *reIL12* unterscheiden sich signifikant. Für die Items *reIS5* und *reIS6* wird die Nullhypothese, dass sich die Gruppen nicht signifikant unterscheiden, beibehalten. Da für 80% der Items die Nullhypothese abgelehnt werden muss, wird diese für die gesamte Skala abgelehnt.

Das heißt, es gibt signifikante Unterschiede in der Motivation zwischen Schülerinnen und Schüler der siebten und achten Schulstufe.

4.3.2. Inhaltliche Analyse der Gruppenarbeit-Items

Die Gruppenarbeit-Items zeigten bei der Faktorenanalyse keine guten Ergebnisse hinsichtlich der Ladungen auf die Skala *relatedness*, weshalb diese ausgeschlossen wurden. In einem Versuch, die Items zu einer eigenen Skala zusammenzufügen, wurden mittelmäßige Ergebnisse erhalten. Der KMO-Koeffizient von 0,658 fiel mäßig aus, die Korrelation der Items mit $p < 0,001$ ist jedoch höchstsignifikant.

Ein Faktor erklärt 40,509% der Gesamtvarianz der Items, was in Ordnung ist. Zwei zugrunde liegende Faktoren würden 59,662% der Gesamtvarianz erklären, wobei jedoch nicht klar wäre, um welche zwei Faktoren es sich handeln sollte. Die Annahme eines Faktors wird weiters dadurch unterstützt, dass entsprechend dem Kaiser-Kriterium nur ein Eigenwert größer eins ist.

Bei den Ergebnissen der Faktorenanalyse fiel auf, dass Item *relS11* „Ich arbeite lieber alleine.“ eine negative Ladung aufweist. Bei der Analyse mit zwei Faktoren wurde ebenfalls ein negativer Wert dieses Items erhalten. Somit wurde zunächst überprüft, ob die Werte des Items in SPSS korrekt umkodiert wurden, da es sich um ein negativ gepoltes Item handelt. Die Werte wurden korrekt umkodiert, woraus sich schließen lässt, dass das Item nicht zu den restlichen passt.

Aufgrund dieser Ergebnisse kann gesagt werden, dass Gruppenarbeit-Items nicht geeignet erscheinen, um das Konstrukt *relatedness* zu repräsentieren, sie können somit weggelassen werden.

4.4. Schwierigkeiten und Anmerkungen zu den Items

Bei der Pilotierung bekamen die Schülerinnen und Schüler den Fragebogen zum Ausfüllen und wurden im Vorhinein angewiesen darauf zu achten, ob sie alles verstehen, ob es bei Aussagen Unklarheiten gibt und ob ihnen Fehler im Text auffallen. Nachdem die Lernenden den Fragebogen ausgefüllt hatten, wurden sie in Kleingruppen nacheinander aus der Klasse geholt und zu den obig genannten Punkten befragt.

Es gab Schülerinnen und Schüler, die sagten, dass der Fragebogen etwas lang sei. Der Großteil war der Meinung, dass die Länge durchaus in Ordnung sei. Zu Unklarheiten kam es bei Item *re/L2*: „Wenn ich aufzeige, sieht mich die Lehrperson an, auch wenn sie mich nicht dran nehmen kann.“ Einige Lernende wussten nicht, was mit dem Item gemeint ist, deswegen wurde es umgeändert auf „Die Lehrperson merkt, dass ich aufzeige, auch wenn sie mich nicht dran nehmen kann.“

Bei Item *re/L6* „Die Lehrperson unterstützt mich, meine Sicht zu formulieren.“ wussten ebenfalls einige Schülerinnen und Schüler nicht was gemeint ist. In der Diskussion der Kleingruppen ergab sich, dass durch die Hinzunahme des Wortes „Meinung“ klarer wird, worauf das Item abzielt. Es wurde geändert auf „Die Lehrperson unterstützt mich, meine Sicht/Meinung zu formulieren.“. Es gab Lernende, die bei Item *re/S7* „Strukturierter Unterricht ist für mich wichtig.“ nicht wussten, was mit „strukturierter Unterricht“ gemeint ist. Das Item wurde umformuliert zu „Gut geplanter Unterricht ist für mich wichtig.“.

In Tabelle 19 sind die Anmerkungen, welche die Lernenden im Zuge der Pilotierung auf den Fragebogen geschrieben haben, zu sehen.

Item-bezeichnung	Item	Anmerkung
<i>re/S5</i>	Ich werde von meinen Mitschülern ausgelacht, wenn ich eine Frage stelle oder meine Meinung äußere.	bei falschen Fragen

Tabelle 19: Anmerkungen auf den Fragebögen bei der Pilotierung

Für die eigentliche Testung wurden die oben genannten Änderungen durchgeführt und es wurde ein Item, *re/S14* „In der Gruppe teilen wir die Arbeit auf alle auf.“, gestrichen, da es bereits ein ähnliches gab. Während der Testung gab es kaum Fragen, vereinzelt wollten Schülerinnen und Schüler sicher gehen, dass sie ein Item richtig verstehen und fragten nach.

Zu den folgenden Items wurde nachgefragt:

relL2 „Die Lehrperson merkt, dass ich aufzeige, auch wenn sie mich nicht dran nehmen kann.“

relS7 „Gut geplanter Unterricht ist für mich wichtig.“

relL14 „Es gibt Regeln, wie man sich zu Wort meldet und diese werden eingehalten.“

In Tabelle 20 sind die Anmerkungen, welche die Lernenden im Zuge der Testung auf den Fragebogen schrieben, angeführt.

Item-bezeichnung	Item	Anmerkung
<i>relS1</i>	Wenn ich im Unterricht etwas sage, kann ich ausreden, ohne dass mich jemand unterbricht.	die Schüler
<i>relL3</i>	Die Lehrperson hört mir zu, wenn ich etwas im Unterricht sage.	er versucht
<i>relL9</i>	Es kommen immer dieselben Schüler dran.	Die immer aufzeigen.
<i>relS9</i>	Bei einer Gruppenarbeit teilen wir die Arbeit auf alle auf.	machen wir nicht
		wir machen keine
<i>relL10</i>	Jeder, der etwas sagen will, kommt zu Wort.	außer es ist am Ender der Stunde und es läutet schon
<i>relS12</i>	In meiner Klasse lachen wir oft miteinander.	miteinander wurde ausgebessert zu übereinander

Tabelle 20: Anmerkungen auf den Fragebögen bei der Testung

Item *relL2* „Die Lehrperson merkt, dass ich aufzeige, auch wenn sie mich nicht dran nehmen kann.“ verursachte trotz Umformulierung vereinzelt Verständnisprobleme, weshalb es für die fertige Skala nicht verwendet wurde. Schülerinnen und Schüler verstanden vermutlich nicht, dass dieses Item darauf abzielt, dass die Lehrperson die Schülerin beziehungsweise den Schüler wahrnimmt und das beispielsweise durch Blickkontakt signalisiert.

Item *relS7* „Gut geplanter Unterricht ist für mich wichtig.“ wurde von Lernenden oft nicht mit der Unterrichtsplanung in Verbindung gebracht, wodurch sich Fragen ergaben. Es gab Lernende die sich unter dem Ausdruck „gut geplanter Unterricht“ nichts vorstellen konnten. Eine kurze Erklärung beseitigte jedoch schnell die Verständnisprobleme.

Bei Item *re/L14* „Es gibt Regeln, wie man sich zu Wort meldet und diese werden eingehalten.“ waren Lernende verunsichert und fragten beispielsweise nach, ob damit das „Aufzeigen“ gemeint ist. Möglicherweise bereitete der Ausdruck „zu Wort melden“ Probleme.

Eine weitere Problematik ergab sich mit den Gruppenarbeit-Items. In einer Klasse, in der keine Intervention durchgeführt wurde, machte die Lehrperson nur sehr selten Gruppenarbeiten, wodurch die Lernenden sich nicht sicher waren, was sie ankreuzen sollen. Zwei der Lernenden haben am Fragebogen „machen wir nicht“ und „wir machen keine“ dazugeschrieben. Da die Gruppenarbeit-Items aufgrund ihrer Ladung weggelassen wurden, lässt sich der Fragebogen beziehungsweise die Skala *relatedness* in vielfältigeren Bereichen einsetzen.

Enthält ein Fragebogen Items, die auf gemeinsames Arbeiten abzielen, so ist es ratsam, eine Intervention davor durchzuführen oder die Lehrkräfte im Vorhinein zu fragen, ob sie Gruppenarbeit mit den Schülerinnen und Schülern machen.

Wie schon Korner (2015) feststellte, fiel auf, dass manche Schülerinnen und Schüler Schwierigkeiten mit der doppelten Verneinung hatten. Drei Lernende fragten, was sie bei Item *re/L4* „Die Lehrperson gibt uns nicht genügend Zeit zum Nachdenken/Überlegen.“ ankreuzen sollen, wenn die Lehrperson ihnen schon genug Zeit gibt. Weiters gab es immer wieder Lernende, die gewisse Items laut vorlasen und zu lachen anfangen, Item *re/S2* „Wir hören einander in der Klasse zu.“ und *re/S3* „Wir helfen einander in der Klasse.“ zählten dazu.

Die ausgewählten Items (siehe 4.3.1.) sind einfach formuliert und können somit von der siebten bis zur zwölften Schulstufe eingesetzt werden. Der Fragebogen eignet sich womöglich sogar für jüngere Schülerinnen und Schüler, doch um diese Annahme zu bestätigen, müssen weitere Untersuchungen durchgeführt werden. Die Skala *relatedness* ist im Gegensatz zu allen anderen Skalen von einer zuvor durchgeführten Intervention beziehungsweise Tätigkeit unabhängig. Dieses Ergebnis konnte aus dem t-Test für die beiden Fragebogenversionen entnommen werden. Von den anderen Skalen beziehen sich die Skala *value/usefulness* und *effort/importance* sehr stark auf eine zuvor durchgeführte Tätigkeit, aber auch einige Items der anderen Skalen. In künftigen Arbeiten kann versucht werden etwas von der Aufgabenbezogenheit wegzukommen, falls dies möglich ist.

Da die Skala *relatedness* mit zehn Items recht kurz ist, kann sie von Lehrkräften eingesetzt werden, um sich ein schnelles Feedback von den Schülerinnen und Schülern einzuholen, wie diese den Unterricht und das Klassenklima wahrnehmen.

4.5. Qualität des Fragebogens

In diesem Kapitel wird die Qualität der neu erstellten Skala *relatedness* anhand der in Kapitel 2.5. aufgelisteten Gütekriterien beurteilt. Mit *Fragebogen* ist im Folgenden die neu erstellte Skala *relatedness* bestehend aus den zehn Items gemeint.

Hauptgütekriterien

Da die entwickelte Skala von einer zuvor durchgeführten Intervention unabhängig ist, kann der Fragebogen einfach im regulären Unterrichtsverlauf ausgeteilt werden und bedarf nur ein paar kurzer Instruktionen, die immer gleich formuliert werden können. Damit ist die Durchführungsobjektivität gegeben, der Fragebogen kann nach gleicher Vorgehensweise und unter klaren Bedingungen durchgeführt werden. Die Items sind klar und einfach formuliert, wodurch sich während einer Testung kaum Fragen ergeben sollten.

Der Fragebogen verwendet Likert-skalierte Items, die eine von der auswertenden Person unabhängige Auswertung ermöglichen, wodurch die Auswertungsobjektivität gegeben ist. Für die Interpretationsobjektivität ist eine ausreichend große Stichprobe nötig, welche vorliegt. Die standardisierte Interpretationsmöglichkeit des Tests beziehungsweise des Fragebogens ist gegeben, da aufgrund der Likert-Skalierung jede auswertende Person zum gleichen Ergebnis kommt und beurteilen kann, ob der erhaltene Wert durchschnittlich, unter- bzw. überdurchschnittlich ist (Bühner, 2011). Insgesamt kann der Fragebogen als *objektiv* angesehen werden.

Die *Reliabilität* des Fragebogens wurde mittels Konsistenzmethode geschätzt, wobei die innere Konsistenz analysiert wurde. Für das Cronbach- α ergab sich ein Wert von 0,815, welcher als gut gewertet werden kann.

Die *Validität* ist aufgrund der Erstellung der Items mithilfe der Methoden von Wilson (2005) gegeben. Indem der Kreislauf des Four Building Blocks Approachs mehrmals durchlaufen wurde, ist sichergestellt, dass die Items auf das Konstrukt abbilden, wodurch die Inhaltsvalidität hoch ist (Schmidt, 2017).

Die resultierenden Testwerte spiegeln die Merkmalsrelation adäquat wider, womit das Kriterium der *Skalierbarkeit* gegeben ist (Schmidt, 2017). Die wenig streuenden Trennschärfen von 0,320 bis 0,591 sind weiters ein Hinweis darauf, dass das Gütekriterium der Skalierbarkeit erfüllt ist (Bühner, 2011).

Nebengütekriterien

Durch die Likert-Skalierung der Items können die Ergebnisse der einzelnen Personen verglichen werden, wodurch der Fragebogen *normiert* ist.

Direkt *vergleichbar* ist der Fragebogen nicht, da es keinen ähnlichen gibt. Man kann die Skala *relatedness* mit den bereits von Korner (2015), Klingenböck (2016) und Schmidt (2017) erstellten Skalen vergleichen, hinsichtlich beispielsweise der Faktorladungen, der Gesamtvarianz oder anderen Variablen.

Der Fragebogen ist *ökonomisch*, mit zehn Items ist er kurz und weist eine angemessene Durchführungszeit auf, er ist für eine Gruppentestung tauglich, wenig materialintensiv und aufgrund der Likert-Skalierung schnell und einfach auswertbar. Selbst die Vollversion ist mit einer Testdauer von durchschnittlich sieben Minuten nicht sehr zeitintensiv. Aufgrund dessen sind die Kriterien der *Zumutbarkeit* und der *Fairness* erfüllt, der Fragebogen unterscheidet nicht zwischen Gruppen von Personen.

Die *Nützlichkeit* des Fragebogens ist gegeben, da es im deutschsprachigen Raum nur wenig zufriedenstellende Fragebögen zur Evaluierung der Motivation von Schülerinnen und Schülern gibt. Weiters können mit der Skala *relatedness* Faktoren wie Klassenklima und die Beziehung zwischen Lehrkraft und Lernenden schnell und einfach evaluiert werden (Bühner, 2011).

Insgesamt sind die Haupt- und Nebengütekriterien eines guten psychometrischen Tests zufriedenstellend erfüllt.

5) Diskussion

In diesem Kapitel erfolgt eine genauere Diskussion der Ergebnisse und der ausgewählten Items. Weiters wird der Konnex von der neu erstellten Skala zu den bereits vorhandenen hergestellt und auf die Forschungsfragen eingegangen.

5.1. Diskussion der neu erstellten Skala

In der ersten Version der Items lag der Fokus auf einzelnen Personen, mit denen ein Lernender interagiert und darauf, wie sich die Dynamik in einer Gruppenarbeitsphase gestaltet. Nach den Befragungen der Lernenden wurde schnell klar, dass die Skala *relatedness* das Klassenklima, den Umgang der Lernenden miteinander sowie den Umgang der Lehrperson mit den Lernenden berücksichtigen muss. Das Konstrukt wurde nach den Befragungen überdacht und in der Construct Map überarbeitet. Der Fokus wurde nun weg von einzelnen Personen, hin zum Klassenklima als Ganzes und der sozialen Interaktion in der Klasse gelegt. Weiters wurde die starke Aufgabenbezogenheit beziehungsweise der Bezug auf eine zuvor durchgeführte Tätigkeit fast komplett aufgegeben. Von den 30 zu testenden Items waren fünf Gruppenarbeit-Items, der Großteil der Items bezog sich somit auf die sozialen Interaktionen. In der Diskussion der Expertenrunde wurde vermutet, dass Gruppenarbeiten beziehungsweise die Interaktionen während einer Gruppenarbeit, zur sozialen Eingebundenheit beitragen. Diese Vermutung bewahrheitete sich jedoch nicht, es kann aber auch sein, dass die befragten Lernenden selten in Gruppen arbeiteten, wodurch sie bei der Beantwortung der Items möglicherweise etwas unsicher waren.

Die Skala sollte laut Theorie in zwei Faktoren zerfallen, nämlich der sozialen Interaktion der Lernenden untereinander und der Interaktion zwischen Lehrperson und Lernenden. Diese beiden Faktoren sollen die soziale Komponente der Motivation von Schülerinnen und Schülern abbilden. Die exploratorische Faktorenanalyse bestätigte, dass die Skala in zwei Faktoren zerfällt.

Betrachtet man den Inhalt der Items, so fällt auf, dass all den ausgewählten Items der respektvolle Umgang miteinander zugrunde liegt. Das heißt, pflegen sowohl die Lernenden untereinander einen respektvollen Umgang, als auch die Lehrperson mit den Lernenden und umgekehrt, so ist es wahrscheinlicher, dass Schülerinnen und Schüler eine hohe Motivation aufweisen. Der respektvolle Umgang manifestiert sich hauptsächlich durch gegenseitiges Zuhören und Helfen.

Der Mittelwert der neu erstellten Skala *relatedness* beträgt 3,63 mit einer Standardabweichung von 0,59. Betrachtet man nun den Mittelwert der siebten und der achten Schulstufe, so fällt auf, dass Lernende der siebten Schulstufe mit einem Mittelwert von 3,87 motivierter sind als Lernende der achten Schulstufe, mit einem Mittelwert von 3,47. Abbildung 6 kann weiters entnommen werden, dass Lernende der siebten Schulstufe durchgehend motivierter waren als Lernende der achten Schulstufe. Die Gründe dafür können sehr verschieden sein, Einflussfaktoren können unter anderem die Tagesverfassung, der sozioökonomische Hintergrund der Lernenden, der Standort der Schule, die Schulform sowie die Zeit der Testung sein (Schmidt, 2017). Eine wichtige Rolle dürfte jedoch das Alter der Lernenden spielen. Jüngere Lernende sind womöglich im Durchschnitt interessierter und offener für Neues, was jedoch nur eine Vermutung darstellt.

Um die Unterschiede der Motivation an den verschiedenen Schulen besser analysieren zu können, werden diese kurz beschrieben. Der sozioökonomische Hintergrund der Lernenden konnte einerseits durch die Lage der Schule und andererseits durch Erzählungen der Lehrpersonen grob eingeschätzt werden. Der sozioökonomischen Hintergrund wurde aufgrund der beiden Indikatoren Bildung und Beruf der Eltern festgemacht (Bauer, 2016).

Bei Schule A handelt es sich um ein privates Gymnasium mit einem sehr geringen Anteil an nicht-deutschsprachigen Schülerinnen und Schülern. Schule B ist ein großes Gymnasium mit durchschnittlichem sozioökonomischen Hintergrund. Bei Schule C handelt es sich um ein Gymnasium mit sehr gutem sozioökonomischen Hintergrund. Schule D ist die einzige Neue Mittelschule. Bei Schule E handelt es sich ebenfalls um ein Gymnasium. Die Schulen B, D und E weisen einen ähnlichen sozioökonomischen Hintergrund auf, ebenso wie die Schulen A und C. In den Schulen A, C und D wurde besonders viel Wert auf den respektvollen Umgang miteinander gelegt.

Betrachtet man die Mittelwerte der Skala *relatedness* für die einzelnen Schulen, so lässt sich feststellen, dass die Lernenden der verschiedenen Schulen mit Werten von 3,19 bis 4,02 mittel bis hoch motiviert waren. Schule C, welche einen sehr guten sozioökonomischen Hintergrund hat, wies dabei mit 4,02 den höchsten Wert an Motivation auf. Interessanterweise wies Schule D, in der Lernende mit weniger gutem sozioökonomischen Hintergrund gingen, eine höhere Motivation auf, als Schule E mit etwas besserem sozioökonomischen Hintergrund. Der Kruskal-Wallis-Test bestätigte, dass es zwischen den Schulen signifikante Unterschiede in der Motivation gibt. Mittels eines Mann-Whitney-Tests konnten zwei Gruppen gebildet werden, Schule D und Schule E bilden im Durchschnitt eine Gruppe weniger motivierter Lernender und die Schulen A, B, C eine Gruppe stärker motivierter Lernender.

Schule D hat die größte Spannweite, daraus lässt sich schließen, dass es hier sehr motivierte und auch sehr un motivierte Schülerinnen und Schüler gab. Schule C hat eine geringe Spannweite, die Lernenden weisen allesamt eine hohe Motivation auf.

5.2. Diskussion des Einflusses der Intervention

Die Analyse des Einflusses der Intervention zeigt, dass es insgesamt keine signifikanten Unterschiede zwischen der Beantwortung der Items mit Intervention und ohne Intervention gibt. Der t-Test konnte für die Items *relS4*, *relL6* und *relS6* als signifikant ($p < 0,05$) eingestuft werden, die restlichen Items sind als nicht signifikant zu werten. Dieses Ergebnis ist zum Teil überraschend, da zumindest Item *relS4* „Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.“ und Item *relS6* „Meine Mitschüler behandeln mich mit Respekt.“ inhaltsmäßig von einer Intervention unabhängig sein könnten.

Die Mittelwertunterschiede der beiden Fragebogenversionen sind in Abbildung 8 für jedes Item dargestellt.

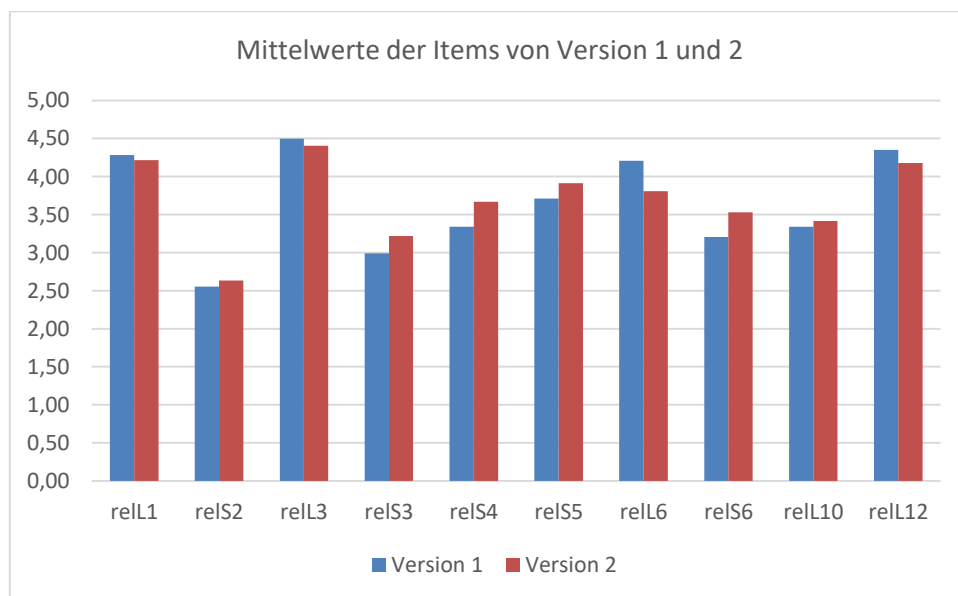


Abbildung 8: Mittelwerte der Items von Fragebogen Version 1 und 2

Für künftige Arbeiten an der Skala *relatedness* kann aufgrund dieser Ergebnisse gesagt werden, dass es keiner Intervention vor der Testung bedarf. Weiters kann überprüft werden, ob die anderen Skalen auch unabhängig von einer zuvor durchgeführten Intervention getestet werden können. Man könnte versuchen enger mit den Lehrpersonen zusammenzuarbeiten und den Fragebogen in den regulären Unterrichtsverlauf zu integrieren.

Dadurch besteht auch nicht die Möglichkeit, dass die Lernenden verwirrt sind, auf wen sie sich bei der Beantwortung des Fragebogens beziehen sollen.

Vor allem die Skalen *perceived competence*, *perceived choice* und *interest/enjoyment* würden sich zur Verallgemeinerung eignen und zur gemeinsamen Testung mit der Skala *relatedness*, ohne zuvor eine Intervention durchzuführen.

5.3. Inhaltliche Diskussion der Ergebnisse

In der ersten Forschungsfrage sollte untersucht werden, ob sich das Konstrukt *relatedness* aus dem IMI auf die Schulsituation übertragen lässt. Beim Erstellen der Construct Map und der Items wurde festgestellt, dass die übersetzten Items des IMI das Konstrukt im schulischen Kontext nicht gut abbilden. Grund dafür ist, dass sich das Konstrukt *relatedness* beziehungsweise soziale Eingebundenheit im Schulkontext wesentlich komplexer darstellt. Das Konstrukt wurde deswegen im Laufe der Arbeit komplett neu überarbeitet. In einer Gruppendiskussion der fachdidaktischen Arbeitsgruppe Physik wurden zwei wichtige Bereiche der sozialen Eingebundenheit, welche die Motivation von Schülerinnen und Schülern beeinflussen, ausgemacht. Diese Bereiche sind das Klassenklima und die sozialen Interaktionen in der Klasse, dazu gehören die Interaktionen der Lernenden untereinander und die Interaktion mit der Lehrperson. Aufgrund dessen wurde angenommen, dass die Skala *relatedness* in zwei Faktoren zerfällt und zwar in Items, die Lernenden betreffend und Items, die Lehrperson betreffend. Diese Annahme wurde durch die exploratorische Faktorenanalyse bestätigt. Im Gegensatz zur Skala *relatedness* liegt allen anderen Skalen nur ein Faktor zugrunde.

Als zweite Forschungsfrage sollte untersucht werden, ob eine trennscharfe Skala *relatedness* erstellt werden kann und ob sie sich faktorenanalytisch bezüglich verschiedener Stichproben abgrenzen lässt.

Die Analyse der neu erstellten Skala lieferte gute Trennschärfen im Bereich von 0,320 bis 0,591. Somit konnte eine trennscharfe Skala *relatedness* erstellt werden.

Um zu überprüfen, ob sich die Skala faktorenanalytisch bezüglich verschiedener Stichproben abgrenzen lässt, wurde eine Faktorenanalyse für ausgewählte Stichproben durchgeführt, die Ladungen der Stichproben auf die Skala müssen dann ähnlich sein (Korner, 2015).

Als Erstes wurden die Stichproben der beiden Fragebogenversionen untersucht. Die aus den beiden Faktorenanalysen erhaltenen Ladungen sind ähnlich. Es wurde weiters Schule A und Schule B verglichen, da diese einen vergleichbaren Stichprobenumfang aufweisen.

Auch hier konnte eine für beide Schulen durchgeführte Faktorenanalyse bestätigen, dass die Ladungen auf die Skala ähnlich sind. Es kann somit gesagt werden, dass sich die neu erstellte Skala *relatedness* faktorenanalytisch bezüglich verschiedener Stichproben abgrenzen lässt.

Im Laufe der Entwicklung der Skala wurden die Gruppenarbeit-Items gestrichen, da diese schlecht zu den restlichen Items passten. In Bezug auf die soziale Eingebundenheit bedeutet das, dass andere Faktoren wichtiger sind, als die Arbeitsaufteilung und Dynamik bei Gruppenarbeiten. Durch das Wegfallen der Gruppenarbeit-Items wurde die Skala unabhängig von einer zuvor durchgeführten Tätigkeit und kann somit theoretisch für jedes Unterrichtsfach und im regulären Unterrichtsverlauf spontan eingesetzt werden. In Folgearbeiten sollte dies jedoch noch überprüft werden.

Der t-Test zur Überprüfung der Mittelwertgleichheit von Version 1 und Version 2 des Fragebogens ergab nicht signifikante Unterschiede, womit analytisch bestätigt wurde, dass die Ergebnisse von einer zuvor durchgeführten Intervention beziehungsweise Tätigkeit unabhängig sind, was auch die dritte Forschungsfrage beantwortet.

Zusätzlich bietet die Vollversion der Skala *relatedness* mit den 30 Items, wie sie in Anhang B, 11.3.1, dargestellt ist, Lehrkräften die Möglichkeit sich ein schnelles Feedback der Lernenden einzuholen, um einen Überblick über das Klassenklima sowie die sozialen Interaktionen zu erhalten. Mit einer Testdauer von durchschnittlich sieben Minuten ist die Vollversion des Fragebogens zeitsparend und kann von Lehrpersonen mittels Excel schnell und einfach ausgewertet werden.

Zur Auswertung können die Daten spaltenweise für jeden Fragebogen in Excel eingegeben werden (siehe Abbildung 9). Die Kodierung der Antworten kann wie in dieser Arbeit erfolgen, wobei „1“ *stimmt gar nicht* bedeutet, „2“ *stimmt eher nicht*, bis hin zu „5“ *stimmt völlig*. #1 steht für den ersten Fragebogen. Je nach Bedarf können auch nur gewisse Items ausgewählt werden. Sind alle Daten eingegeben, muss nur mehr über jede Zeile gemittelt werden, um den Mittelwert des Items zu erhalten. Danach können die Ergebnisse interpretiert werden.

A	B	C	D	E
Itemnr.	#1	#2	...	Mittelwert
1	3	5		4
2	4	4		4
3	2	4		3
4	2	4		3
5	4	5		4,5
6	3	4		3,5
7	1	1		1
8	3	4		3,5
9	4	3		3,5
10	1	2		1,5
11	3	3		3
12	4	4		4
13	3	1		2
...	...	...		...
28	5	3		4
29	4	1		2,5
30	2	5		3,5

Abbildung 9: Möglichkeit zur Auswertung der Daten mittels Excel

Zusammenfassend kann gesagt werden, dass gute Ergebnisse für die Skala *relatedness* erhalten wurden, in weiteren Arbeiten kann diese noch überarbeitet und weiterentwickelt werden. Außerdem muss noch überprüft werden, ob sich zusammen mit den anderen Skalen eine Faktorenstruktur reproduzieren lässt und ob die Skala *relatedness* zusammen mit den anderen Skalen trennscharf und stabil ist (Schmidt, 2017).

Rückblickend auf den Weg zur Erstellung der Skala *relatedness* lässt sich sagen, dass die Diskussion in der fachdidaktischen Arbeitsgruppe Physik einen großen Beitrag zur Entwicklung der Skala leistete. Für künftige Arbeiten ist es empfehlenswert, sich schon früh in der Arbeitsphase mit Expertinnen und Experten auszutauschen und sich Feedback beziehungsweise Verbesserungsvorschläge einzuholen. Weiters empfiehlt es sich Befragungen an Schulen durchzuführen, um auf Evidenzen basierende Items zu erstellen, falls die jetzigen Items nochmals überarbeitet werden sollen. Es kann in Erwägung gezogen werden, auf eine zuvor durchgeführte Intervention zu verzichten und sich gegebenenfalls mit den Lehrkräften abzusprechen, damit der Zeitaufwand verringert werden kann. Falls möglich, sollte eine größere Stichprobe getestet werden. Als letzten Schritt müssen alle Skalen gemeinsam getestet werden, um letztendlich ein Instrument zur Ermittlung der Motivation von Schülerinnen und Schülern zu erhalten.

6) Zusammenfassung und Ausblick

Mit dieser Arbeit wurde ein weiterer Schritt hin zu einem deutschsprachigen Instrument zur Evaluierung der Motivation von Schülerinnen und Schülern gemacht. Die bestehenden Instrumente beziehungsweise Fragebögen sind teils sehr spezifisch oder können meist nur für bestimmte Zwecke eingesetzt werden.

Korner wollte im Rahmen ihrer Arbeit „Cross-age peer tutoring in Physik“ die Motivation der Schülerinnen und Schüler testen, wobei sich schnell Schwierigkeiten aufzeigten. Die bestehenden deutschsprachigen Instrumente waren für ihre Arbeit nicht geeignet, weswegen sich Korner entschied das Intrinsic Motivation Inventory (Deci & Ryan, 2003) auf Deutsch zu übersetzen. Das IMI ist ein englischer Fragebogen zur Evaluierung der Motivation, der aus folgenden sieben Skalen besteht:

interest/enjoyment, effort/importance, relatedness, value/usefulness, perceived choice, perceived competence und pressure/tension.

Probleme ergaben sich einerseits, da das IMI für Teamsport entwickelt wurde und nicht so einfach auf den schulischen Kontext übertragen werden konnte und andererseits, weil die Lernenden Probleme mit der Beantwortung negativ gepolter Items hatten. Manche Items bereiteten den Lernenden generell Verständnisprobleme. Aus diesen Gründen beschloss Korner die Skala *effort/importance* mithilfe des Four Building Blocks Approach von Wilson (2005) neu zu entwickeln. Klingenböck (2016) entwickelte daraufhin die Skala *interest/enjoyment* neu und Schmidt (2017) die Skalen *value/usefulness, perceived choice* und *perceived competence*. Die Vorgehensweise, mit der Methode von Wilson, erwies sich als besonders vielversprechend.

Die letzte verbleibende Skala, *relatedness*, wurde in dieser Arbeit neu entwickelt. Dazu wurde ebenfalls der Four Building Blocks Approach von Wilson (2005) verwendet. Der erstellte Fragebogen wurde in Klassen der siebten und achten Schulstufe an insgesamt 284 Lernenden getestet. Die Hälfte der Klassen füllte den Fragebogen im regulären Unterrichtsverlauf aus, in der anderen Hälfte wurde zuvor eine Intervention zum Thema Magnetismus durchgeführt.

Auf Basis einer exploratorischen Faktorenanalyse wurden zehn Items ausgewählt, die hoch auf die Skala *relatedness* laden. Die Annahme, dass die neu erstellte Skala in zwei Faktoren zerfällt, wurde durch die Analyse bestätigt. Diese beiden Faktoren sind das Klassenklima und die Beziehung zwischen Lehrperson und Lernenden.

Bei der Analyse stellte sich heraus, dass die Intervention keinen signifikanten Einfluss auf die Beantwortung der Items hatte.

In künftigen Arbeiten kann die Skala *relatedness* gemeinsam mit den anderen Skalen getestet werden. Weiters muss überprüft werden, ob diese sich stabil und trennscharf mit den anderen Skalen verhält. Es besteht die Möglichkeit, die bereits bestehenden Skalen zu verallgemeinern und vom starken Bezug einer zuvor durchgeführten Tätigkeit etwas wegzukommen, falls dies möglich ist. Inwiefern dies machbar ist, kann gegebenenfalls in einer Expertenrunde diskutiert werden. Alle bestehenden Skalen können nochmals überarbeitet und zusammen getestet werden, um letztendlich die gewünschte Faktorstruktur zu erhalten.

Weitere Untersuchungsmöglichkeiten stellt die Motivation von Schülerinnen und Schülern in verschiedenen Schulstufen dar. In der Analyse zeigte sich, dass die Motivation von Lernenden der siebten Schulstufe größer ist, als die Motivation von Lernenden der achten Schulstufe. Man könnte Studien in verschiedenen Schulstufen durchführen um herauszufinden, warum die Motivation in höheren Schulstufen abnimmt und ob das generell der Fall ist, oder nur in der getesteten Stichprobe beobachtet wurde.

Zusammenfassend lässt sich sagen, dass die Analyse der neu erstellten Skala *relatedness* gute Ergebnisse lieferte. Insgesamt wurde das Instrument zur Ermittlung der Motivation von Schülerinnen und Schülern verbessert.

Zusätzlich bietet die Vollversion des Fragebogens Lehrkräften die Möglichkeit, sich einen schnellen Überblick über das Klassenklima sowie über die sozialen Interaktionen in der Klasse zu verschaffen.

7) Literaturverzeichnis

- Achtziger, A., & Gollwitzer, P. (2010). Motivation und Volition im Handlungsverlauf. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation und Handeln*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Anderman, E., Austin, C., & Johnson, D. (2002). The development of goal orientation. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation*.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, Vol.64(6), 359.
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191-215.
- Barmeier, M., Boldt, J., & Ciprina, H. J. (2009). Prisma Physik 4.
- Bauer, J. (2016). Sozioökonomischer Status. *Zentralblatt für Arbeitsmedizin, Arbeitsschutz und Ergonomie*, Vol. 66(5), 318-320.
- Beckmann, J., & Heckhausen, H. (2010). Motivation durch Erwartung und Anreiz. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation und Handeln*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- BMB. (2000). Bundesministerium für Bildung, Lehrplan Physik, Unterstufe. https://www.bmb.gv.at/schulen/unterricht/lp/lp_ahs_unterstufe.html am 16.12.2017.
- Bresler, S., Heepmann, B., & Hörter, C. (2012). Natur und Technik - Physik.
- Brinkman, R. Mathematik und Physik für Schüler, Lehrer und Eltern. https://www.brinkmann-du.de/physik/sek1/ph09_06a.htm am 02.01.2018.
- Brunstein, J., & Heckhausen, H. (2010). Leistungsmotivation. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation und Handeln*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (3., aktualisierte und erw. Aufl. ed.). München [u.a.]: Pearson Studium.
- Deci, E., & Ryan, R. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. [The theory of self-determination of motivation and its relevance to pedagogics]. *Zeitschrift für Pädagogik*, 39(2), 223-238.
- Deci, E., & Ryan, R. (2003). Intrinsic Motivation Inventory.
- Deci, E., & Ryan, R. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian psychology/Psychologie canadienne*, 49(3), 182-185.

- Duit, R. (2010). PIKO-Briefe.
- Duit, R., Treagust, D., & Widodo, A. (2008). Teaching science for conceptual change: Theory and practice *International handbook of research on conceptual change* (pp. 629-646): Routledge.
- Eckstein, P. (2016). *Angewandte Statistik mit SPSS*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Fachportal-Physik. (2017). Landesbildungsservers Baden-Württemberg, Magnetismus und Elektrostatik - ein Vergleich
- Guse, C., Noble, E., Schneiderbauer, L., Kramer, P., & Liebel, V. (2013). Das Physik-Skript. <http://www.schule.promathika.de/index.php?n=PhysikSkript>. PhysikSkript am 09.01.2018.
- Heckhausen, J., & Heckhausen, H. (2010). *Motivation und Handeln*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Hewitt, P. G. (2000). Conceptual Physics. *Pearson Prentice Hall*.
- Hopf, M., Wiesner, H., & Schecker, H. (2011). *Physikdidaktik kompakt*. Freising: Aulis-Verlag.
- Klingenböck, A.-M. (2016). *Ermittlung der intrinsischen Motivation von SchülerInnen der Unterstufe in der Elektrizitätslehre*. Wien.
- Korner, M. (2015). *Cross-age Peer Tutoring in Physik*. Wien.
- Müller, F. H., Hanfstingl, B., & Andreitz, I. (2007). Skalen zur motivationalen Regulation beim Lernen von Schülerinnen und Schülern. *Adaptierte und ergänzte Version des Academic*, 242.
- 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.
- Raduta, C. (2005). General students' misconceptions related to Electricity and Magnetism.
- Rasch, B. (2014). *Quantitative Methoden* (4 ed.): Springer Berlin Heidelberg
- Rheinberg, F. (2010). Intrinsische Motivation und Flow-Erleben. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation und Handeln*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen. *QCM: A Questionnaire to assess current motivation in learning situations*, 57.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, 25(1), 54-67.

- Schlag, B. (2013). *Lern- und Leistungsmotivation*: Springer Fachmedien Wiesbaden.
- Schmidt, F. (2017). *Entwicklung weiterer Skalen eines Messinstruments zur Bestimmung der intrinsischen Motivation von SchülerInnen*. Wien.
- Steffens, U., & Höfer, D. (2014). Die Hattie-Studie: Forschungsbilanz & Handlungsperspektiven. <http://www.sqa.at/course/view.php?id=36> am 10.06.2018.
- Steuer, G., Fasching, M., & Dresel, M. Lern- und Leistungsmotivation I: Grundlagen und Komponenten. <https://wuecampus2.uni-wuerzburg.de/moodle/mod/book/view.php?id=322112&chapterid=5263> am 07.01.2018.
- Stiensmeier-Pelster, J., & Heckhausen, H. (2010). Kausalattribution von Verhalten und Leistung. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation und Handeln*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Urhahne, D. (2008). Sieben Arten der Lernmotivation. *Psychologische Rundschau*, 59(3), 150-166.
- von Glasersfeld, E. (1997). *Kleine Geschichte des Konstruktivismus* (Vol. 8): Österreichische Zeitschrift für Geschichtswissenschaften.
- Wentzel, K. R. (1999). Social-Motivational Processes and Interpersonal Relationships: Implications for Understanding Motivation at School. *Journal of Educational Psychology*, 91(1), 76-97.
- Widodo, A., & Duit, R. (2004). Konstruktivistische Sichtweisen vom Lehren und Lernen und die Praxis des Physikunterrichts. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 233-255.
- Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009). Überprüfung einer Kurzsкала intrinsischer Motivation (KIM). *Zeitschrift für Didaktik der Naturwissenschaften*, 15.
- Wilson, M. A. (2005). *Constructing measures*. Mahwah, N.J. [London]: Lawrence Erlbaum Associates Taylor & Francis.

8) Abbildungsverzeichnis

Abbildung 1: Selbstbestimmungskontinuum von handlungsbezogenen Werten (Steuer, Fasching, & Dresel)	21
Abbildung 2: Four building blocks (Wilson, 2005, p. 31)	37
Abbildung 3: Screeplot der Faktoren der Skala relatedness	61
Abbildung 4: Mittelwerte der Skala relatedness für die verschiedenen Schulen	67
Abbildung 5: Boxplots der Motivation in den verschiedenen Schulen	68
Abbildung 6: Mittelwerte der Items für die 7. und 8. Schulstufe	69
Abbildung 7: Mittelwert der Skala relatedness für die 7. und 8. Schulstufe	69
Abbildung 8: Mittelwerte der Items von Fragebogen Version 1 und 2	79
Abbildung 9: Möglichkeit zur Auswertung der Daten mittels Excel	82

9) Tabellenverzeichnis

Tabelle 1: Version 1 der Construct Map zur Skala relatedness	41
Tabelle 2: Version 2 der Construct Map zur Skala relatedness	43
Tabelle 3: Version 3A der Construct Map zur Skala relatedness	45
Tabelle 4: Version 3B der Construct Map zur Skala relatedness	46
Tabelle 5: Didaktische Analyse	51
Tabelle 6: Stundenplanung	52
Tabelle 7: Übersicht der Schulen und Fragebogenversionen	55
Tabelle 8: KMO-Koeffizient und Bartlett-Test der relatedness Items	58
Tabelle 9: Rotierte Komponentenmatrix der neu erstellten Skala relatedness	59
Tabelle 10: Erklärte Gesamtvarianz der Skala relatedness	60
Tabelle 11: Item-Skala-Statistiken für die Skala relatedness	61
Tabelle 12: KMO-Koeffizient und Bartlett-Test für die Gruppenarbeit-Items	62
Tabelle 13: Erklärte Gesamtvarianz der Gruppenarbeit-Items	62
Tabelle 14: Komponentenmatrix der Gruppenarbeit-Items, ein Faktor extrahiert	63
Tabelle 15: Rotierte Komponentenmatrix der Gruppenarbeit-Items, zwei Faktoren extrahiert	63
Tabelle 16: Gruppenstatistiken der beiden Fragebogenversionen	64
Tabelle 17: t-Test der ausgewählten Items in Bezug auf die beiden Fragebogenversionen	65
Tabelle 18: t-Test der ausgewählten Items in Bezug auf die beiden Schulstufen	70
Tabelle 19: Anmerkungen auf den Fragebögen bei der Pilotierung	72
Tabelle 20: Anmerkungen auf den Fragebögen bei der Testung	73
Tabelle 21: Zuordnung der Nummer im Fragebogen zur Itembezeichnung bei der Pilotierung	102
Tabelle 22: Zuordnung der Nummer im Fragebogen zur Itembezeichnung für Version 1 und 2	102
Tabelle 23: Rotierte Komponentenmatrix der Vollversion der Skala relatedness	103

10) Anhänge A

10.1. Abstract in Deutsch

Im deutschsprachigen Raum gibt es nur wenige Instrumente, welche die Motivation von Schülerinnen und Schülern messen können. Diejenigen, die es gibt, sind meist sehr spezifisch. Ein gutes Instrument zur Evaluierung der Motivation stellt das Intrinsic Motivation Inventory dar, welches von Deci und Ryan entwickelt wurde und aus sieben Skalen besteht. Es ist in englischer Sprache verfasst und somit schwer anwendbar im deutschsprachigen Unterricht. Der erste Versuch das IMI ins Deutsche zu übersetzen wurde von Korner (2015) im Rahmen ihrer CAPT-Studie unternommen, wobei sie mithilfe des Four Building Blocks Approach von Wilson die Skala *effort/importance* neu erstellte. Diese Vorgehensweise erwies sich als nützlich und lieferte gute Ergebnisse. Klingenböck (2016) entwickelte die Skala *interest/enjoyment* neu und Schmidt (2017) die Skalen *value/usefulness*, *perceived choice*, *perceived competence*. In dieser Arbeit wurde die letzte verbleibende Skala, *relatedness*, neu entwickelt, welche an Schülerinnen und Schülern der siebten und achten Schulstufe getestet wurde. Die Skala *relatedness* ist trennscharf und die Items laden hoch auf die Skala, generell wurden vielversprechende Ergebnisse erhalten. Weiters bietet die Vollversion des Fragebogens Lehrkräften die Möglichkeit sich ein schnelles Feedback über das Klassenklima und die sozialen Interaktionen in der Klasse einzuholen. In künftigen Arbeiten können die Skalen zusammen getestet und verbessert werden, um letztendlich ein Instrument zur Evaluierung der Motivation von Schülerinnen und Schülern zu erhalten.

10.2. Abstract in Englisch

In the German-speaking area there are only a few instruments that evaluate the motivation of students. This diploma thesis contributes to the development of such an instrument by developing a new *relatedness* scale of the Intrinsic Motivation Inventory. The scale *effort/importance* has been developed by Korner (2014), *interest/enjoyment* by Klingenböck (2016) and the scales *value/usefulness*, *perceived choice*, *perceived competence* by Schmidt (2017). Korner began her work on the IMI because there was no satisfying instrument in German to evaluate motivation. For her purposes she used the Four Building Blocks Approach from Wilson (2005) and successfully created the new scale *effort/importance*. The further scales have been developed with the same methods and proved to be stable and selective. In order to develop items for the *relatedness* scale Wilson's approach was used in this thesis as well. The items were then tested amongst seventh- and eighth-graders. With the results of an exploratory factor analysis ten items with the highest loadings onto the scale were chosen. The new developed scale has proven to be selective, in general good results were obtained. Future research can focus on combining the present scales in order to finally create a questionnaire which evaluates the motivation of students.

11) Anhänge B

Die Antworten der Schülerinnen und Schüler wurden inhaltsgetreu wiedergegeben, wobei die untenstehenden Aussagen nicht wortgetreu sind.

Im Arbeitsblatt wurden die Referenzen nachträglich hinzugefügt, die Lernenden erhielten es der Übersicht halber ohne Referenzen.

11.1. Befragungen der Schülerinnen und Schüler

Befragung Nr. 1

Schule in 1150 Wien

8. Schulstufe

Anzahl der Schülerinnen und Schüler: 23 (m: 15, w: 8)

Datum: 6.12.2017, 12 Uhr

Antworten der Schülerinnen und Schüler:

- Lehrperson und Lernende arbeiten gemeinsam
- Mitschülerinnen und Mitschüler sind leise
- die Lehrperson muss Humor verstehen
- die Lehrperson muss streng sein, aber Humor verstehen
- die Lernenden müssen brav, leise und respektvoll sein
- die Lehrperson muss eine Leidenschaft für das Fach haben und diese auch ausstrahlen
- die Energie der Lehrperson muss spürbar sein
- zu Gruppenarbeit:
 - mit Mitschülerinnen und Mitschülern zu diskutieren hilft bei der Erfassung neuer Inhalte
 - selbst überlegen, diskutieren
 - ein „Wettkampf“ beim Erklären untereinander hilft, dass man motivierter ist
 - durch gegenseitiges helfen fühlt man sich wohl, aber es braucht Disziplin dazu
 - austauschen untereinander ist wichtig
 - bei Gleichaltrigen können Erklärungen besser verstanden werden, die Lehrperson sollte aber schon auch erklären
 - es ist besser nicht neben Freunden zu sitzen, damit man nicht in Versuchung kommt zu tratschen

Befragung Nr. 2

Schule in 1150 Wien (+2 externe Lernende)

Schulstufe 8,9,10,12 (Lernende, die an der Physikolympiade teilnehmen)

Anzahl der Schülerinnen und Schüler: 9 (m: 7, w: 2)

Datum: 6.12.2017, 14:35 Uhr

Antworten der Schülerinnen und Schüler:

- Freiraum ist wichtig, dass man sich wohl fühlt
- Klassenklima muss gut sein
- jeder soll so akzeptiert werden, wie er ist (einerseits unter den Lernenden, aber auch die Lehrperson soll von den Lernenden akzeptiert werden und umgekehrt)
- die Lehrperson soll objektiv und gerecht sein und die Lernenden in den Unterricht miteinbeziehen
- Gruppenarbeit:
 - bei einer Gruppenarbeit ist es nicht gut, wenn einer alles macht
 - gut ist eine Gruppenarbeit, wenn die Anforderungen passen und wenn die Aufgaben gut verteilt sind
- es ist wichtig einander zu helfen, auch wenn das während des Unterrichts nicht immer möglich ist
- Lernende fühlen sich unwohl, ...
 - wenn die Lehrperson so tut, also wäre sie den Lernenden charakterlich und vom Wissen her überlegen
 - wenn die Lehrperson den Lernenden das Gefühl gibt, sie wären blöd oder blöder als die Lehrperson
- zwischen Lehrperson und Lernenden sollte eine nette, gute Stimmung herrschen
- man merkt, dass eine Schülerin oder ein Schüler sich wohl fühlt, durch...
 - die Mitarbeit (wie oft zeigt er/sie auf, wie viel und fleißig wird mitgearbeitet)
 - Mimik – lächeln, nicht gelangweilt dasitzen
 - bei einer Gruppenarbeit reden alle miteinander, keiner sitzt abseits
 - die Körperhaltung (offen)
- man merkt, dass eine Schülerin oder ein Schüler sich unwohl fühlt, durch...
 - dauerndes aufs Handy Schauen
 - die Person ist leise und schaut woanders hin
- über die Distanz, über die kommuniziert wird, lässt sich sagen, dass wenn sie groß ist, man sich wohl fühlt in der Gruppe und wenn sie klein ist, man eher nur mit dem Sitznachbar spricht, man sich nicht so wohl und nicht so selbstbewusst fühlt
- generell ist eine strukturierte Kommunikation wichtig
- alleine arbeiten ist auch okay, wobei es in der Klasse leise sein sollte
- respektvoller Umgang miteinander ist besonders wichtig

Befragung Nr. 3

Schule in 1050 Wien

Schulstufe 10

Datum: 14.12.2017, 9:00 Uhr

Antworten der Schülerinnen und Schüler:

- man fühlt sich miteinbezogen, wenn
 - die anderen Schüler freundlich sind
 - die Lehrperson respektvoll ist und nicht genervt ist, wenn man Fragen stellt
 - die Lehrperson einen zuhört, wenn man etwas sagt
- man fühlt sich wohl, wenn
 - das Klassenklima passt
 - die anderen nicht gemein sind
 - man gute Noten hat
- an der offenen Körperhaltung merkt man, dass sich jemand wohl fühlt
- man merkt, dass sich Leute in einer Gruppe wohl fühlen, wenn sie lachen und kommunizieren
- man kann gut lernen in der Klasse, wenn es ruhig ist und die anderen einen ausreden lassen
- Gruppenarbeit
 - ist gut, weil man sich mit den anderen austauschen kann
 - ist gut, weil man einander helfen kann

Befragung Nr. 4

Schule in 1230 Wien

Schulstufe 9

Anzahl der Schülerinnen und Schüler: 19 (m: 5, w: 14)

Datum: 20.12.2017, 13:05 Uhr

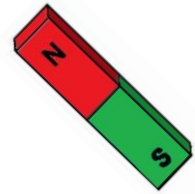
Antworten der Schülerinnen und Schüler:

- man fühlt sich eingebunden, wenn
 - man von der Lehrperson nicht ignoriert wird
 - sich die Lehrperson und die anderen SuS für einen interessieren
 - man miteinander diskutiert
 - jede/jeder drangenommen wird
 - die Unterrichtsinhalte interessant sind

- nicht nur Frontalunterricht stattfindet, sondern auch andere Methoden verwendet werden (unter anderem Gruppenarbeit)
- ein respektvoller Umgang aller herrscht
- man fühlt sich wohl, wenn ...
 - man Spaß im Unterricht hat
 - respektvoll miteinander umgegangen wird
 - einander geholfen wird
 - man Freunde in der Klasse hat
 - man so akzeptiert wird, wie man ist (seitens der Lehrperson und der SuS)
 - die Lehrperson fair beurteilt
 - man seitens der Lehrperson nicht überfordert wird und zu viel zu tun bekommt
 - das Klassenklima gut ist
 - die Lehrperson eine gute Einstellung zum Fach hat (davon begeistert ist und Motivation zeigt)
 - die Lehrperson auch andere Meinungen akzeptiert als die eigene
 - die Lehrperson zu einer guten Zeiteinteilung fähig ist (Stoffumfang, Hausübungen, ...)
- jemand der sich wohl fühlt
 - hat eine offene Körperhaltung
 - er/sie ist redefreudig
 - beteiligt sich aktiv an Diskussionen, am Bearbeiten der Aufgabe
- man merkt, dass sich jemand unwohl fühlt
 - durch eine geschlossene Körperhaltung
 - dadurch, dass die Person alleine wo sitzt oder abseits der anderen
 - indem die Person leise und unsicher spricht
- Gruppenarbeit
 - macht mehr Spaß
 - man merkt sich die Inhalte so besser
 - man kann mehr Ideen einbringen und muss keine Angst haben, dass man etwas Falsches sagt, da man sich untereinander austauschen kann
 - man kann mehr sagen

11.2.

Arbeitsblatt Magnetismus



ELEMENTARMAGNETE

1. Lest euch den folgenden Informationstext durch und markiert das Wichtigste!
2. Bearbeitet gemeinsam die Aufgaben am Ende!
3. Besprecht wie eure Ergebnisse mit dem Text zusammenpassen!



Text Elementarmagnete

Jeder Magnet hat einen Nordpol und einen Südpol. Im Bereich der Pole ist die magnetische Kraft am größten. Zwischen zwei Magneten können anziehende und abstoßende Kräfte auftreten. Gleichnamige Pole stoßen sich ab und ungleichnamige Pole ziehen einander an.

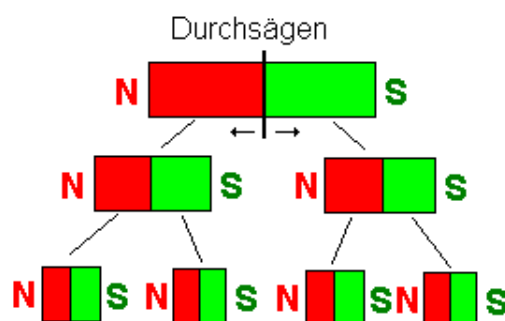
Zwischen einem Magneten und Gegenständen, die Eisen, Nickel, oder Cobalt enthalten, gibt es eine starke magnetische Anziehung. Diese Stoffe nennt man ferromagnetische Stoffe. Man kann Körper, die Eisen, Nickel oder Cobalt enthalten magnetisieren. Dazu fährt man mehrmals mit einem Pol eines Magneten über den zu magnetisierenden Gegenstand. Dabei muss man in die gleiche Richtung streichen (Barmeier, Boldt, & Ciprina, 2009).

Allgemein gilt:

„Körper, die von Magneten angezogen werden, sind auch selbst magnetisierbar.“ (Guse, Noble, Schneiderbauer, Kramer, & Liebel, 2013)

Experiment magnetisierter Nagel:

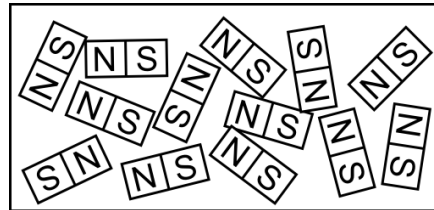
In diesem Experiment wird ein magnetisierter Nagel auseinandergeschnitten und dabei in zwei Hälften geteilt. Es stellen sich nun die Fragen, ob der Nagel noch magnetisch ist und ob die beiden Stücke noch zwei Pole haben. Macht man das Experiment, so sieht man, dass die beiden Teile des Nagels magnetisch sind und jedes Teilstück sowohl Nordpol als auch Südpol hat. Man könnte ein Teilstück des Nagels nochmals teilen. In der Abbildung unterhalb siehst du dieses Prinzip bei einem Stabmagneten, der in der Mitte geteilt wird.



(Fachportal-Physik, 2017)

Der Vorgang lässt sich beliebig oft wiederholen. Doch wie oft kann man den Nagel nun teilen, gibt es eine Grenze?

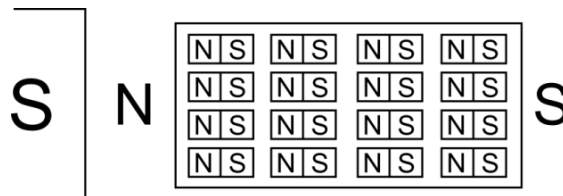
Ja, die gibt es. Man stellt sich vor, dass Magnete aus sehr vielen, sehr kleinen und nicht weiter teilbaren Magneten, den sogenannten **Elementarmagneten**, bestehen. Hier siehst du ein nicht magnetisches Material. In einem nicht magnetischen Material sind die kleinen Elementarmagnete ungeordnet und heben sich in ihrer magnetischen Wirkung auf.



(Guse et al., 2013)

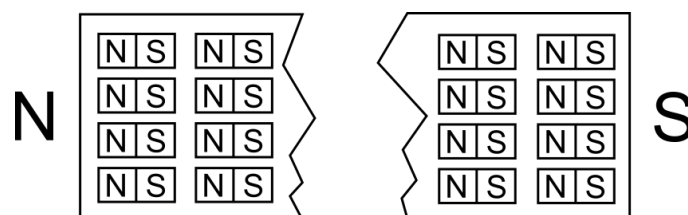
Der ferromagnetische Stoff hat hier als Ganzer keine magnetischen Pole.

Streicht man mit einem Magneten den ferromagnetischen Stoff entlang, so richten sich die Elementarmagnete größtenteils aus. Dadurch verstärken die Elementarmagnete ihre Wirkung und der ferromagnetische Stoff bildet Pole aus.



(Guse et al., 2013)

Teilt man den magnetisierten Stoff, so bleibt die Ausrichtung der Elementarmagnete bestehen. Beide Pole sind, wie in der Abbildung dargestellt, in beiden Bruchstücken vorhanden.



(Guse et al., 2013)

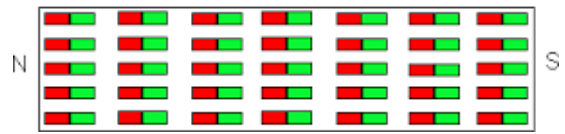
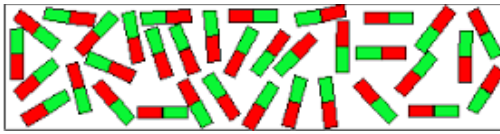
Durch Erhitzung oder starke mechanische Einwirkung wird die Ordnung der Elementarmagnete wieder zerstört. Der ferromagnetische Stoff hat dann seine Magnetisierung verloren. Man sagt er ist entmagnetisiert (Barmer et al., 2009).

Wichtig ist:

Es gibt keine magnetischen Monopole. Magnete haben immer zugleich einen Nordpol **und** einen Südpol.



Aufgabe 1:



(Brinkman)

Beschreibt in ein paar Sätzen die Situation, die in den beiden Bildern dargestellt ist. Was ist zwischen den beiden Bildern passiert?



Aufgabe 2:

Kreuze an, ob die Aussage richtig oder falsch ist!

Aussage	richtig	falsch
In einem magnetischen Material sind die Elementarmagnete größtenteils ausgerichtet.		
Es gibt Magnete, die nur einen Pol haben.		
Körper aus Eisen, Nickel und Kobalt sind magnetisierbar.		
Ist ein magnetisiertes Material einmal magnetisch, so bleibt es für immer magnetisch.		
Teilt man einen Magneten in der Hälfte, so haben beide Teile Nordpol und Südpol.		

11.3.1 Fragebogen Version 1

Liebe Schülerin, lieber Schüler!

Versuche die einzelnen Fragen bitte möglichst genau zu beantworten. Alle Antworten werden anonym behandelt und nicht benotet.

		stimmt völlig	stimmt eher	stimmt teilweise	stimmt eher nicht	stimmt gar nicht
1	Meine Fragen werden im Unterricht ernst genommen und die Lehrperson versucht sie immer zu beantworten.	☐	☐	☐	☐	☐
2	Wenn ich im Unterricht etwas sage, kann ich ausreden, ohne dass mich jemand unterbricht.	☐	☐	☐	☐	☐
3	Die Lehrperson merkt, dass ich aufzeige, auch wenn sie mich nicht dran nehmen kann.	☐	☐	☐	☐	☐
4	Wir hören einander in der Klasse zu.	☐	☐	☐	☐	☐
5	Die Lehrperson hört mir zu, wenn ich etwas im Unterricht sage.	☐	☐	☐	☐	☐
6	Wir helfen einander in der Klasse.	☐	☐	☐	☐	☐
7	Die Lehrperson gibt uns nicht genügend Zeit zum Nachdenken/Überlegen.	☐	☐	☐	☐	☐
8	Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.	☐	☐	☐	☐	☐
9	Die Lehrperson lässt uns in Kleingruppen arbeiten.	☐	☐	☐	☐	☐
10	Ich werde von meinen Mitschülern ausgelacht, wenn ich eine Frage stelle oder meine Meinung äußere.	☐	☐	☐	☐	☐
11	Die Lehrperson unterstützt mich, meine Sicht/Meinung zu formulieren.	☐	☐	☐	☐	☐
12	Meine Mitschüler behandeln mich mit Respekt.	☐	☐	☐	☐	☐
13	Wenn ich etwas nicht verstehe, traue ich mich bei der Lehrperson nachzufragen.	☐	☐	☐	☐	☐
14	Gut geplanter Unterricht ist für mich wichtig.	☐	☐	☐	☐	☐

Bitte umblättern!

		stimmt völlig	stimmt eher	stimmt teilweise	stimmt eher nicht	stimmt gar nicht
15	Ich bin mit der Beurteilung (Notengebung) durch die Lehrperson zufrieden.	☐	☐	☐	☐	☐
16	Meine Mitschüler sind gemein zueinander.	☐	☐	☐	☐	☐
17	Es kommen immer dieselben Schüler dran.	☐	☐	☐	☐	☐
18	Bei einer Gruppenarbeit teilen wir die Arbeit auf alle auf.	☐	☐	☐	☐	☐
19	Jeder, der etwas sagen will, kommt zu Wort.	☐	☐	☐	☐	☐
20	Ich kann mich nur schwer konzentrieren, wenn ich in einer Gruppe arbeite.	☐	☐	☐	☐	☐
21	Manchmal lerne ich mehr von meinen Mitschülern, als von der Lehrperson.	☐	☐	☐	☐	☐
22	Ich arbeite lieber alleine.	☐	☐	☐	☐	☐
23	Ich werde von der Lehrperson ernst genommen und respektvoll behandelt.	☐	☐	☐	☐	☐
24	In meiner Klasse lachen wir oft miteinander.	☐	☐	☐	☐	☐
25	Die Lehrperson beachtet mich nicht, wenn ich etwas zum Unterricht sagen will.	☐	☐	☐	☐	☐
26	Ich traue mich im Unterricht Fragen zu stellen.	☐	☐	☐	☐	☐
27	Es gibt Regeln, wie man sich zu Wort meldet und diese werden eingehalten.	☐	☐	☐	☐	☐
28	Die Lehrperson gibt mir genügend Zeit zum Nachdenken/Überlegen.	☐	☐	☐	☐	☐
29	In einer Gruppenarbeit stören die anderen nur und tratschen.	☐	☐	☐	☐	☐
30	Zusammen mit anderen kann ich bessere Ergebnisse erzielen.	☐	☐	☐	☐	☐

11.3.2 Fragebogen Version 2

Liebe Schülerin, lieber Schüler!

Die folgenden Fragen beziehen sich auf die eben gemachte Arbeit. Bitte versuche die einzelnen Fragen möglichst genau zu beantworten. Alle Antworten werden anonym behandelt und nicht benotet.

		stimmt völlig	stimmt eher	stimmt teilweise	stimmt eher nicht	stimmt gar nicht
1	Meine Fragen werden im Unterricht ernst genommen und die Lehrperson versucht sie immer zu beantworten.	☐	☐	☐	☐	☐
2	Wenn ich im Unterricht etwas sage, kann ich ausreden, ohne dass mich jemand unterbricht.	☐	☐	☐	☐	☐
3	Die Lehrperson merkt, dass ich aufzeige, auch wenn sie mich nicht dran nehmen kann.	☐	☐	☐	☐	☐
4	Wir hören einander in der Klasse zu.	☐	☐	☐	☐	☐
5	Die Lehrperson hört mir zu, wenn ich etwas im Unterricht sage.	☐	☐	☐	☐	☐
6	Wir helfen einander in der Klasse.	☐	☐	☐	☐	☐
7	Die Lehrperson gibt uns nicht genügend Zeit zum Nachdenken/Überlegen.	☐	☐	☐	☐	☐
8	Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.	☐	☐	☐	☐	☐
9	Die Lehrperson lässt uns in Kleingruppen arbeiten.	☐	☐	☐	☐	☐
10	Ich werde von meinen Mitschülern ausgelacht, wenn ich eine Frage stelle oder meine Meinung äußere.	☐	☐	☐	☐	☐
11	Die Lehrperson unterstützt mich, meine Sicht/Meinung zu formulieren.	☐	☐	☐	☐	☐
12	Meine Mitschüler behandeln mich mit Respekt.	☐	☐	☐	☐	☐
13	Wenn ich etwas nicht verstehe, traue ich mich bei der Lehrperson nachzufragen.	☐	☐	☐	☐	☐
14	Gut geplanter Unterricht ist für mich wichtig.	☐	☐	☐	☐	☐

Bitte umblättern!

		stimmt völlig	stimmt eher	stimmt teilweise	stimmt eher nicht	stimmt gar nicht
15	Ich bin mit der Beurteilung (Notengebung) durch die Lehrperson zufrieden.	☐	☐	☐	☐	☐
16	Meine Mitschüler sind gemein zueinander.	☐	☐	☐	☐	☐
17	Es kommen immer dieselben Schüler dran.	☐	☐	☐	☐	☐
18	Bei einer Gruppenarbeit teilen wir die Arbeit auf alle auf.	☐	☐	☐	☐	☐
19	Jeder, der etwas sagen will, kommt zu Wort.	☐	☐	☐	☐	☐
20	Ich kann mich nur schwer konzentrieren, wenn ich in einer Gruppe arbeite.	☐	☐	☐	☐	☐
21	Manchmal lerne ich mehr von meinen Mitschülern, als von der Lehrperson.	☐	☐	☐	☐	☐
22	Ich arbeite lieber alleine.	☐	☐	☐	☐	☐
23	Ich werde von der Lehrperson ernst genommen und respektvoll behandelt.	☐	☐	☐	☐	☐
24	In meiner Klasse lachen wir oft miteinander.	☐	☐	☐	☐	☐
25	Die Lehrperson beachtet mich nicht, wenn ich etwas zum Unterricht sagen will.	☐	☐	☐	☐	☐
26	Ich traue mich im Unterricht Fragen zu stellen.	☐	☐	☐	☐	☐
27	Es gibt Regeln, wie man sich zu Wort meldet und diese werden eingehalten.	☐	☐	☐	☐	☐
28	Die Lehrperson gibt mir genügend Zeit zum Nachdenken/Überlegen.	☐	☐	☐	☐	☐
29	In einer Gruppenarbeit stören die anderen nur und tratschen.	☐	☐	☐	☐	☐
30	Zusammen mit anderen kann ich bessere Ergebnisse erzielen.	☐	☐	☐	☐	☐

11.4. Übersicht der Itemzuordnung im Fragebogen

Die folgenden beiden Tabellen geben eine Übersicht, welche Nummer im Fragebogen welchem Item entspricht.

Pilotierung

Nr. im Fragebogen	Originalnr. in der Skala	Nr. im Fragebogen	Originalnr. in der Skala
1	relL1	17	relL9
2	relS1	18	relS9
3	relL2	19	relL10
4	relS2	20	relS10
5	relL3	21	relL11
6	relS3	22	relS11
7	relL4	23	relL12
8	relS4	24	relS12
9	relL5	25	relL13
10	relS5	26	relS13
11	relL6	27	relL14
12	relS6	28	relS14
13	relL7	29	relL15
14	relS7	30	relS15
15	relL8	31	relS16
16	relS8		

Tabelle 21: Zuordnung der Nummer im Fragebogen zur Itembezeichnung bei der Pilotierung

Testung

Nr. im Fragebogen	Originalnr. in der Skala	Nr. im Fragebogen	Originalnr. in der Skala
1	relL1	16	relS8
2	relS1	17	relL9
3	relL2	18	relS9
4	relS2	19	relL10
5	relL3	20	relS10
6	relS3	21	relL11
7	relL4	22	relS11
8	relS4	23	relL12
9	relL5	24	relS12
10	relS5	25	relL13
11	relL6	26	relS13
12	relS6	27	relL14
13	relL7	28	relS14
14	relS7	29	relL15
15	relL8	30	relS15

Tabelle 22: Zuordnung der Nummer im Fragebogen zur Itembezeichnung für Version 1 und 2

Rotierte Komponentenmatrix^a		
	Komponente	
	1	2
relL1	,746	
relS1	,463	
relL2	,681	
relS2	,346	,581
relL3	,719	
relS3		,619
relL4	,339	
relS4		,728
relL5	,403	
relS5		,602
relL6	,624	
relS6		,732
relL7	,417	
relS7	,395	
relL8	,628	
relS8		,731
relL9	,310	
relS9	,408	,339
relL10	,720	
relS10		,472
relL11		
relS11		-,360
relL12	,710	
relS12	,277	
relL13	-,389	
relS13	-,302	
relL14	,463	,339
relS14	,660	
relL15		,541
relS15		,351
Extraktionsmethode: Hauptkomponentenanalyse. Rotationsmethode: Varimax mit Kaiser-Normalisierung. ^a a. Die Rotation ist in 3 Iterationen konvergiert.		

Tabelle 23: Rotierte Komponentenmatrix der Vollversion der Skala relatedness