MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Weiterentwicklung eines ökonomischen, validen Erhebungsinstruments zur Bestimmung der intrinsischen Motivation von Schülerinnen und Schülern“

verfasst von / submitted by

Jakob Pusch, BEd

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Education (MEd)

Wien, 2021 / Vienna, 2021

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 199 510 523 02
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Lehramt Sek (AB) UF Geographie und Wirtschaft UF Physik
Betreut von / Supervisor:	Univ.-Prof. Dr. Martin Hopf
Mitbetreut von / Co-Supervisor:	Mag. Dr. Marianne Korner

Danksagung

Ich bedanke mich bei allen Menschen, die mich bei der Erstellung dieser Arbeit begleitet und unterstützt haben.

Mein Dank gilt zuallererst Univ.-Prof. Dr. Martin Hopf und Mag. Dr. Marianne Korner für die Übernahme und Betreuung dieser Masterarbeit.

Frau Mag. Dr. Marianne Korner hatte für meine Fragen und Anliegen stets ein offenes Ohr und schuf ein angenehmes Arbeitsklima, in dem wir auf Augenhöhe über Vorgangsweisen, Formulierungen und Ergebnisinterpretationen diskutierten.

Ebenso möchte ich mich bei den zahlreichen Schülerinnen und Schülern bedanken, die die Fragebögen meiner Forschungsarbeit gewissenhaft ausgefüllt, und damit meine Forschung erst möglich gemacht haben.

Ein spezieller Dank gilt meiner Familienmitgliedern Anna Pusch, Nina Hammerschmid und Hannelore Patzelt, die sich für das Korrekturlesen dieser Arbeit Zeit nahmen und mir mit hilfreichen Feedbacks zur Seite standen.

Meinen Eltern danke ich besonders für die emotionale und finanzielle Unterstützung meines Studiums.

Danke!

Inhaltsverzeichnis

1. Einleitung: 5

2. Theoretischer Hintergrund 8

2.1. Konstruktivistische Lerntheorien 8

2.1.1. Konstruktivistische Grundlagen 8

2.1.2. Conceptual change 9

2.1.3. Conceptual change im konstruktivistisch orientierten Unterricht 10

2.2. Lernmotivation 13

2.3. Selbstbestimmungstheorie der Motivation 19

2.4. Intrinsic Motivation Inventory (IMI) 25

2.5. Test- und Fragebogenkonstruktion 27

2.5.1. Hauptgütekriterien 27

2.5.2. Nebengütekriterien 30

2.5.3. Faktoranalysen 32

2.5.4. Weitere statistische Analyse- und Testmöglichkeiten 33

2.6. Messinstrumente der Motivation 34

2.6.1. Bestehende Messinstrumente 35

2.6.2. Probleme mit dem IMI 37

2.6.3. Weiterentwicklung bestehender Skalen für das Messinstrument 38

3. Forschungsdesign 42

3.1. Forschungsfragen 42

3.2. Erstellung der Construct Maps und der Items 44

3.2.1. Construct Maps und Items zur Skala perceived competence 45

3.2.2. Construct Maps und Items zur Skala pressure tension 48

3.3. Erstellung des Fragebogens 51

3.4. Durchführung an den Schulen 52

3.5. Auswertung der Fragebögen 54

4. Ergebnisse 55

4.1. Analyse der Skalen des Fragebogens 55

4.1.1. Korrelationstabellen 56

4.1.2. Faktoranalyse der Skala perceived competence 56

4.1.3. Faktoranalyse der Skala pressure and tension 58
4.1.3. Korrelationen zwischen den Skalen perceived competence und pressure and tension 60
4.1.4. Faktoranalyse der übrigen Skalen 61
4.1.5. Faktoranalyse von Items mehrere Skalen gemeinsam 63
4.2. Inhaltliche Analyse der Ergebnisse 66
5. Diskussion 71
5.1. Diskussion der Skalen perceived competence und pressure / tension 71
5.2. Diskussion der übrigen Skalen 73
5.3. Inhaltliche Diskussion der Ergebnisse 75
5.4. Qualität des Fragebogens 76
5.5. Limitationen und Anregungen 78
6. Zusammenfassung und Ausblick 80
7. Literatur: 82
8. Abbildungsverzeichnis und Tabellenverzeichnis: 85
9. Anhänge A 86
9.1. Zusammenfassung (Deutsch) 86
9.2. Abstract (English) 87
10. Anhänge B 88
10.1.1. Fragebogen lang (41 Items) 88
10.1.2. Fragebogen kurz (34 Items) 91
10.2.1. Infoblatt zum Motivationsfragebogen lang: Zuordnung der Items 93
10.2.2. Infoblatt zum Motivationsfragebogen kurz: Zuordnung der Items 94
10.3.1. Übersichtsblatt zur Skala effort and importance (eff) 95
10.3.2. Übersichtsblatt zur Skala relatedeness (rell) 96
10.3.3. Übersichtsblatt zur Skala interest and enjoyment (int) 97
10.3.4. Übersichtsblatt zur Skala value and usefulness (val) 98
10.3.5. Übersichtsblatt zur Skala perceived choice (pch) 99
10.4. Übersichtstabelle der ausgewählten Items nach Skalen VOR der Testung 100
10.5. Übersichtstabelle der verbliebenen Items NACH der Analyse (34 Items) 100

1. Einleitung:

Im schulischen Kontext ist das Thema Motivation und deren Auswirkung auf den Lernerfolg besonders spannend. Was treibt Jugendliche an, sich intensiv einer Thematik beziehungsweise einer Aufgabenstellung zuzuwenden und sich mit dieser auseinanderzusetzen? Warum fällt es anderen schwer, selbst bei gutem Zureden und großer Anstrengung der Lehrperson sich für selbiges zu begeistern? Fragen, die neben Lehrpersonen besonders für Fachdidaktikerinnen und Fachdidaktiker von großem Interesse sind.

Um diese Fragen beantworten zu können, bedarf es einer genaueren Betrachtung des Begriffs „Motivation“. Einen spannenden Ausgangspunkt dafür liefern Deci und Ryan, die Motivation folgendermaßen beschreiben: „To be motivated means to be moved to do something“ (2000, S.1). Dabei können die Ursachen des Handelns innere und äußere Faktoren sein. Im Kontext Schule ist Lernen dann besonders nachhaltig und erfolgreich, wenn der Motivation für eine Tätigkeit intrinsische Faktoren zugrunde liegen. Dies ist dann der Fall, wenn eine Handlung ihrer selbst willen durchgeführt und an sich bereits als erfüllend wahrgenommen wird. Man könnte salopp formulieren, dass die Auseinandersetzung mit einer Thematik bei intrinsisch motivierten Personen mit einem „inneren Wollen“ einhergehen.

Um die intrinsische Motivation von Schülerinnen und Schülern zu messen, finden sich im deutschsprachigen Raum kaum geeignete Messinstrumente. Die bestehenden Instrumente sind zum einen nur speziell einsetzbar, oder weisen zum anderen nicht vernachlässigbare Schwächen auf (vgl. Korner, 2015). Eines der in der Lehr- und Lernforschung oft eingesetzten Testinstrumente zur Ermittlung der Motivation ist das von Deci und Ryan (2003) entwickelte Intrinsic Motivation Inventory (IMI). Es dient zur Messung der Motivation und umfasst sieben Subskalen. Das IMI ist in englischer Sprache verfasst und im Sport getestet.

Einen ersten Schritt, das Konstrukt ins Deutsche zu übersetzten, unternahm Korner (2015) im Rahmen ihrer Dissertation zum Thema „Cross-Age Peer Tutoring in Physik“. Dabei zeigte sich bei einer Pilotierung, dass der übersetzte Fragebogen eine nicht reproduzierbare Faktorstruktur aufweist, weswegen er in dieser Form als ungeeignet gilt. Nach diesem gescheiterten Versuch, eine deutschsprachige Version des IMI für

den Einsatz in der Sek 1, durch Übersetzen und Adaptieren der Items zu erhalten, entschloss sich Korner dazu, die Items zu den einzelnen Skalen von Grund auf neu zu konstruieren. Zur Abbildung der einzelnen Konstrukte wählte Korner den *Four Building Blocks Approach* nach Wilson (2005) und entwarf so Items zur Skala *effort and importance* (2015). Im Rahmen des *Four Building Blocks Approachs* werden bei der Entwicklung verschiedener Items der Subskalen vier Schritte durchlaufen, welche ich in Kapitel 2.6.3 näher beschreibe. Nach derselben Methode folgten weitere Neukonstruktionen der Skalen *interest and enjoyment* (Klingenböck, 2016), *perceived competence, perceived choice, value and usefulness* (Schmidt, 2017) und *relatedeness (Pieler, 2018).*

Diese bisherigen Arbeiten zu dem neu entwickelten Testinstrument zur Ermittlung der intrinsischen Motivation brachten gute Ergebnisse, zeigten jedoch noch offene Fragestellungen auf, die den Ausgangspunkt für meine Masterarbeit bilden. So weisen die beiden Skalen *perceived competence* und *pression and tension* eine hohe Korrelation auf, deren Ursache näher betrachtet werden muss.

Erster Forschungsschwerpunkt meiner Arbeit:
Ermöglicht eine Überarbeitung der Konstrukte nach dem *Four Building Blocks Approach* eine faktoranalytische Abgrenzung bezüglich verschiedener Stichproben? Können trennscharfe Skalen zu den Teilkonstrukten *perceived competence* und *pressure tension* entwickelt werden? Geht man davon aus, dass die wahrgenommene Selbstkompetenz mit Anspannung und Druck negativ korreliert, so stellt sich die Frage, ob eine der Skalen ohne weitere Bedenken weggelassen werden kann.

Im Rahmen der Bearbeitung des ersten Forschungsschwerpunkts wurden die bestehenden Konstrukte samt ihrer Items mit Expertinnen und Experten validiert und überarbeitet sowie hinsichtlich ihrer Trennschärfe analysiert. Um die erste Forschungsfrage beantworten zu können, wurden bestehende und neuentwickelte Items zu den Skalen *perceived competence* und *pressure tension* getestet und anschließend eine Korrelationsanalyse und einer konfirmatorischen Faktoranalyse unterzogen.

Ob die Skalen sich durch einen Faktor abbilden lassen, wurde mittels Kaiser-Kriterium und Scree-Plot untersucht. Weiters wurden die beiden Skalen *perceived competence* und *pressure tension* bezüglich verschiedener Stichproben näher analysiert.

Da sich die erste Forschungsfrage klären ließ, konnte ich im Anschluss der Frage nachgehen, ob sich bei gleichzeitigem Messen aller Subskalen die Faktorenstruktur reproduzieren lässt.

Zweiter Forschungsschwerpunkt:
Sind die Skalen trennscharf und stabil, wenn alle Items gleichzeitig miteinander wechselwirken?

Für die Beantwortung der Forschungsfragen wurden Schülerinnen- und Schülerbefragungen, Expertinnen und Experteninterviews, sowie eine Expertenrunde durchgeführt.

Ziel dieser Arbeit ist es, einen weiteren Schritt zur Finalisierung des ersten deutschsprachigen Testinstruments, welches die Motivation von Schülerinnen und Schüler evaluieren kann, beizutragen. Damit soll es möglich sein, in praktikabler Weise Studien über die Motivation von Lernenden im Unterricht durchzuführen.

Es gelang im Rahmen dieser Arbeit, trennscharfe Konstrukte zu den Skalen *perceived competence* und *pressure tension* zu entwickeln. Der Ausschluss einer der Skalen erwies sich als nicht zielführend, da die Konstrukte weder signifikant noch wie ursprünglich angenommen immer negativ und hoch korrelieren. Darüber hinaus konnte bei der Testung aller Skalen gezeigt werden, dass die unterschiedlichen Skalen zueinander trennscharf sind.

2. Theoretischer Hintergrund

In diesem Kapitel werden zunächst die Grundlagen der konstruktivistischen Lerntheorie dargestellt, sowie deren Kritikpunkte diskutiert. Sie bilden den theoretischen Rahmen des naturwissenschaftlichen Unterrichts. Anschließend wird ein Überblick über verschiedene Motivationstheorien und geeignete Messinstrumente der Motivation gegeben. Im Zuge dessen werden testtheoretische Grundlagen und Gütekriterien zur Evaluierung eines Testentwurfs erläutert.

2.1. Konstruktivistische Lerntheorien

Obwohl sich die Anfänge der konstruktivistischen Lerntheorien bereits im 18. Jahrhundert finden lassen, hielten sie erst seit den 1980er Jahren Einzug in die Lehr- und Lernforschung, insbesondere im naturwissenschaftlichen Bereich. Im Gegensatz zu der bis dahin vorherrschenden Lerntheorie des Kognitivismus geht man davon aus, dass nun die Lernenden selbst den Prozess des Wissenserwerbs, der als individuell gesehen wird, bestimmen. Zuvor ging man davon aus, dass das Vermitteln von Information genügt, um Lernen zu initiieren. Ähnlich einem Nürnberger Trichter wird im Rahmen der kognitivistischen Lerntheorie den Lernenden Wissen eingetrichtert, womit das Hauptaugenmerk auf dem “Lehrenden“ liegt, während die Lernenden eine passive Rolle haben (vgl. Riemeier, 2012).

Obwohl es natürlich unterschiedliche erkenntnistheoretische Strömungen innerhalb des Konstruktivismus gibt, teilen die Sichtweisen die Annahme, dass nun die Lernenden selbst den Prozess des Wissenserwerbs steuern. Damit kommt ihnen eine aktive Rolle zu, im Rahmen dessen sie zielgerichtet agieren. Wie versteht sich nun der Konstruktivismus als Erkenntnistheorie?

2.1.1. Konstruktivistische Grundlagen

Mittels unserer Sinnesorgane nehmen wir die Umwelt mit all ihren Informationen wahr, womit wir ein Abbild der Realität generieren und neues Wissen erlangen (vgl. Arnold, 2007). Zu diesen Sinneseindrücken gehören Gerüche, Laute, Schriftzeichen oder Gefühle die zunächst bedeutungslos sind und denen erst durch das Zusammentreffen mit bereits vorhandenen kognitiven Strukturen eine Bedeutung zugeschrieben werden kann. Ein Individuum konstruiert sich seine Welt infolgedessen durch das Verknüpfen

von bestehendem Vorwissen mit den aufgenommenen Informationen. Die damit erzeugten Wirklichkeiten sind subjektiv und abhängig von den jeweiligen Vorerfahrungen und bestehenden Vorstellungen. Neues Wissen wird entsprechend der Wirklichkeit durch dieses Verknüpfen mit bereits Bekanntem konstruiert. Neurowissenschaftlerinnen und Wissenschaftler gehen davon aus, dass unser Arbeitsgedächtnis etwa sieben neue Informationen gleichzeitig verarbeiten kann. Können diese mit den Informationen im Langzeitgedächtnis in Verbindung gebracht und gespeichert werden, so haben wir uns Wissen angeeignet, also etwas Neues gelernt. Kann die Information nicht mit passenden Strukturen in Verbindung gebracht werden, so gelingt es einer Person nicht, diese einzuordnen beziehungsweise ihr eine sinnvolle Bedeutung zu geben (vgl. Hopf et al., 2011).

Wie stellt sich aus konstruktivistischer Sicht die Ausgangssituation in der Schule dar? Wenn Schülerinnen und Schüler im Rahmen des Unterrichts ein Thema erarbeiten, so haben sie bereits ganz vielfältige, individuelle Erfahrungen zu dem Thema gemacht und daraus Vorstellungen zu Konzepten, Begriffen und Phänomenen entwickelt. Dieser Vorstellungen kommt beim Lernen eine Doppelrolle zu: Einerseits sind sie ein wichtiger Anknüpfungspunkt, andererseits jedoch können sie auch ein Lernhemmnis darstellen. Nach Duit (2010) stimmen häufig die aus dem Alltag gemachten Erfahrungen und Vorstellungen nicht mit den zu lernenden wissenschaftlichen Vorstellungen überein, was zu Verständnisproblemen und Lernschwierigkeiten führt. Insbesondere am Beginn eines neuen Themas führen Bemühungen, möglichst fachlich richtig zu erklären, zur Ausprägung falscher Sichtweisen. Aus konstruktivistischer Sicht muss der Unterricht an den unterschiedlichen Vorstellungen der Schülerinnen und Schüler anknüpfen und zu einem Konzeptwechsel führen.

2.1.2. Conceptual change

Damit ist gemeint, dass die Schülerinnen und Schüler von ihrem Konzept hin zu einem neuen Konzept der physikalischen Sichtweise gelangen. Ziel des Konzeptwechsels ist es, aufzuzeigen, dass die physikalische Sichtweise in vielen Situationen zielführender und widerspruchsfreier und damit fruchtbarer für „tieferes“ Lernen ist (vgl. Duit, 2010).

Studien haben gezeigt, dass an Schülerinnen- und Schülervorstellungen hartnäckig festgehalten wird und dass ein Loslassen von diesen nur in kleinen Schritten geschieht. Duit (2010) zeigt dies anhand eines Beispiels aus der Optik. Jugendliche verstanden zwar die physikalische Sicht, dass ein beleuchteter Körper Licht aussendet, fanden dies jedoch auch nach intensiven Bemühungen seitens der Lehrperson als nicht glaubhaft. Welcher Umgang ergibt sich dadurch für einen Konzeptwechsel hin zu einer physikalischen Sichtweise?

2.1.3. Conceptual change im konstruktivistisch orientierten Unterricht

Nach Jung (1986) gibt es dabei drei Lernwege: anknüpfen, umdeuten und konfrontieren. Bei Ersterem versucht man, an den Vorstellungen der Schülerinnen und Schüler anzuknüpfen, die möglichst nahe an den vorherrschenden wissenschaftlichen Sichtweisen der Physik dran sind. Der Weg des Umdeutens baut auf richtigen Ideen über physikalische Konzepte auf und zielt darauf ab, durch den Gebrauch geeigneter Sprache zu einer physikalisch sinnvollen Bedeutung zu kommen. Ein Beispiel dafür wäre die Umdeutung des Konzepts vom „Verbrauch“ elektrischen Stroms zur Umwandlung in andere Energieformen. Da diese beiden Strategien einen bruchlosen Weg verfolgen, gehören sie zu den kontinuierlichen Lernwegen (vgl. Duit, Treagust, & Widodo, 2008).

Beim Konfrontieren soll es bei den Lernenden zu einem kognitiven Konflikt kommen, weshalb diese Methode zu den diskontinuierlichen Lernwegen gehört. Dieser Konflikt kann etwa mittels eines Experiments erzeugt werden, dessen Ausgang vorhergesagt werden soll. Stimmt der Ausgang dann nicht mit der Vorhersage überein, kann dies der Ausgangspunkt einer Sichtumkehr sein. Allerdings hat sich gezeigt, dass ein kognitiver Konflikt in der Regel nicht ausreicht, um Lernende von falschen Vorstellungen wegzuführen, da sie oft nicht verstehen, worin dieser Konflikt besteht (vgl. Duit, Treagust, & Widodo, 2008).

Liegt nun ein alternatives Konzept vor, welches verständlich, ertragreich sowie nachvollziehbar ist und steigt die Unzufriedenheit der Lernenden mit den bisherigen Vorstellungen, so sind gute Bedingungen für einen erfolgreichen Konzeptwechsel geschaffen. Mit den neuen Herangehensweisen ist es möglich, neue und tiefer gehende Problemstellungen zu bewältigen und damit neue Vorstellungen auszubilden

(vgl. Posner, Strike, Hewson, & Gertzog, 1982). Nachdem wir nun bereits einige Lernwege zum Umgang mit Schülerinnen und Schülervorstellungen kennen, stellt sich die Frage, wie dazu passende Lernumgebungen aussehen müssen.

Um die Konstruktion von Wissen zu fördern, postulieren Widodo und Duit (2004, S. 237) fünf Kategorien konstruktivistischer Lernumgebungen:

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 erste Kategorie erfasst, in welchem Maß und in welcher Art an vorunterrichtliche Schülerinnen- und Schülervorstellungen angeknüpft wird. Dabei ist es wichtig, den Lernenden ihren Wissensstand bewusst zu machen und auf diesen, im Rahmen des Unterrichts, gezielt einzugehen. Daran angepasste, herausfordernde Problemstellungen können kognitive Konflikte provozieren und so zur Bildung neuer Vorstellungen beitragen. Im Rahmen der Anpassungsstrategie, die als zielführender gesehen wird, ist ein solcher kognitiver Konflikt nicht notwendig (vgl. Widodo und Duit, 2004).

Im Rahmen der Kategorie *Relevanz und Bedeutung der Lernerfahrungen* sollen die Lernbedürfnisse der Jugendlichen beachtet werden. Dabei bedarf es der Einbettung von Wissen in bestimmte soziale und den Schülerinnen und Schülern nahe Kontexte, wofür deren Interessen, Einstellungen und Emotionen erkundet und berücksichtigt werden müssen. Materialien aus dem Alltag der Jugendlichen bieten sich besonders bei der Schaffung konstruktivistisch orientierter Lernumgebungen an. Werden darüber hinaus auch Diskussionen über Alltagserlebnisse aktiv eingebaut und Anwendungsmöglichkeiten in diesen aufgezeigt, fällt die Verknüpfung des behandelten Unterrichtsthemas mit der Schülerinnen- und Schülerlebenswelt besonders leicht (vgl. Widodo und Duit, 2004).

Eine weite Kategorie nach Widodo und Duit (2004) bildet die *soziale Interaktion*, bei der davon ausgegangen wird, dass Wissen immer in einem sozialen Kontext

konstruiert wird. Dabei sollen Lernende in verschiedenen Interaktionsformen mit anderen Lernenden diskutieren, um andere Standpunkte zu verstehen und eigene Vorstellungen zu reflektieren.

Im Rahmen der *Unterstützung der Schülerinnen und Schüler beim eigenständigen Lernen* wird von der konstruktivistischen Annahme ausgegangen, dass Lernende für ihren Lernprozess selbst verantwortlich sind. Das Ergebnis des Lernprozesses kann nicht vorhergesagt und der Prozess lediglich initiiert werden. Durch Mitbestimmung, Freiräume und Ergebnisoffenheit im Lehr-Lern-Prozess kann die Unabhängigkeit gezielt gefördert werden. Dabei müssen Schülerinnen und Schüler laufend Konzeptwechsel vollziehen, weswegen sie zum Reflektieren und Hinterfragen eigener Vorstellungen befähigt und ermutigten werden müssen (vgl. Widodo und Duit, 2004).

Die letzte Kategorie *Wissenschaft, Wissenschaftler* und *wissenschaftliches Wissen* widmet sich der Annahme, dass Wissen reflektiert, revidiert und weiterentwickelt wird. Dabei können sich naturwissenschaftliche Erkenntnisse und Sichtweisen ändern, was im Englischen auch als *Nature of Science* bezeichnet wird. Ein Blick in die Geschichte zeigt, dass auch die Konstruktion von Wissen in Ideologien, Weltbildern, subjektiven Werten und Ansichten, religiösen Überzeugungen sowie politischen Strömungen eingebettet ist. Dabei haben sich auch die Arbeitsweisen und Methoden radikal verändert und weiterentwickelt, sowie gewisse Standards im wissenschaftlichen Arbeiten etabliert. Schülerinnen und Schüler sollen mit diesen Denk- und Arbeitsweisen vertraut gemacht werden, aber auch die Grenzen von Theorien und Modellen kennenlernen (vgl. Widodo und Duit, 2004).

Auf Basis dieser fünf Kategorien können konstruktivistische Lernumgebungen geschaffen werden, die bei den Lernenden einen *conceptual change* ermöglichen und fördern. Bei der Frage, wie aktiv sich die Lehrperson in den Lernprozess einschaltet und diesen initiiert, treten Clark, Kirschner und Sweller (2012) für eine *minimally guided instruction* ein. Damit meinen sie, dass die Lehrperson im Rahmen des Anfangsunterrichts explizit und aktiv Instruktionen geben soll und sich mit Fortschreiten des Lernprozesses immer weiter zurücknimmt. Inwiefern diese Strategie aus bildungswissenschaftlicher Sicht zielführend ist, ist umstritten.

Bei der Strukturierung dieses Prozesses helfen nach Duit, Treagust, & Widodo (2008) die folgenden Unterrichtsstrategien:

- Vertraut machen mit den Phänomenen
- Bewusstmachen der Schülerinnen und Schülervorstellungen
- Einführung in die physikalische Sichtweise
- Anwendung der neuen Sichtweise
- Rückblick auf den Lernprozess

Auf die besondere Relevanz der Bewusstmachung von Lernendenvorstellungen wurde im Rahmen dieser Arbeit schon hingewiesen. Nicht weniger wichtig ist der Rückblick auf den Lernprozess, auf den im Unterricht leider häufig verzichtet wird (vgl. Duit, 2010).

Abschließend lässt sich sagen, dass es aus konstruktivistischer Sicht, gerade wegen der Fokussierung auf die Lernenden, mit ihren unterschiedlichen Vorerfahrungen kein allgemein praktikables Rezept für einen erfolgreichen Unterricht gibt. Nur unter Berücksichtigung der Schülerinnen- und Schülervorstellungen, den unterschiedlichen Lernwegen im Umgang mit diesen, den Kategorien für die Schaffung einer konstruktivistischen Unterrichtsumgebung, sowie den Unterrichtsstrategien für die Unterstützung des Lernprozesses, kann es zu tatsächlichen Lerngelegenheiten und damit Bildung physikalischer Vorstellung kommen.

2.2. Lernmotivation

Ein weiterer entscheidender Faktor für den Lernerfolg und damit die Lernleistung ist die Lernmotivation, ein Forschungsgebiet mit einer großen Anzahl an Modellen. Bevor ich hier einen Überblick über die zentralen theoretischen Ansätze der Lernmotivationsforschung gebe, gilt es den Begriff „Motivation“ enger zu fassen.

Einen spannenden Ausgangspunkt dafür liefern Deci und Ryan (2000, S.1), die Motivation folgendermaßen beschreiben: „To be motivated means to be moved to do something“. Eine Person, die zu einer Handlung motiviert ist, hat dafür bestimmte Beweggründe, die unterschiedlich und sogar widersprüchlich sein können. Ausgangspunkt sind dabei psychologische Bedürfnisse, welche durch eine Handlung befriedigt werden sollen.

Um also ein Ziel zu erreichen, stehen in Situationen verschiedene Handlungsmöglichkeiten zur Wahl, deren Handlungsfolgen sich auf spätere Verhaltensweisen auswirken (vgl. Schlag, 2013).

Im Kontext Schule ist die Motivation der „Motor“ des Lernens. Lernmotivation kann als der Wunsch angesehen werden, sich bestimmte Fertigkeiten und Inhalte anzueignen. Die Theorien der Lernmotivation befassen sich mit intentionalen, bewusst gesteuerten und auf bestimmte Ziele gerichtete Lernhandlungen. Im Wesentlichen lassen sich sieben Arten der Lernmotivation unterscheiden:

1. Klassische Leistungsmotivationstheorie
2. Erweiterte Erwartungs-Wert-Theorien
3. Attributionstheorien
4. Zielorientierungskonzepte
5. Theorien intrinsischer Motivation
6. Volitionsmodelle
7. Theorien sozialer Motivation

(vgl. Urhahne, 2008: S. 151)

Im Rahmen der Leistungsmotivationstheorie werden Handlungen beschrieben, die sich an einem externen Maßstab orientieren. An ihm wird Erfolg und Misserfolg einer Tätigkeit gemessen, wobei der Maßstab eine Laufzeit bei einem Wettrennen, eine Schulnote oder auch die Beherrschung eines Instruments sein kann (vgl. Urhahne, 2008). Das Grundmodell der Leistungsmotivationstheorie geht auf das Risiko-Wahl-Modell von Atkinson (Atkinson, 1975) zurück. Nach ihm entsteht leistungsmotiviertes Handeln aus der Hoffnung auf Erfolg (Erfolgsmotiv) und dem damit einhergehenden Gefühl von Stolz und Zufriedenheit, oder aber der Furcht vor Misserfolg (Misserfolgsmotiv), begleitet von Gefühlen der Scham. Aus welchem Motiv heraus gehandelt wird, hängt von Erfahrungen und Gefühlen der Person ab. Nach Schlag (2013) ist deshalb besonders bei wenig leistungsmotivierten Schülerinnen und Schülern darauf zu achten, Erfolgserlebnisse zu ermöglichen, um damit die Motivation zu steigern. Ebenso ist eine realistische Zielsetzung notwendig, um die eigenen Erwartungen erfüllen zu können.

Die Erwartungs-Wert-Theorien gehen davon aus, dass einer Person eine Reihe alternativer Handlungen zur Verfügung stehen. Unter diesen wird jene gewählt, bei der das Produkt aus Wert und Erwartung am größten erscheint (vgl. Rudolph, 2003). Nach Urhahne (2008) können vier Arten von Erwartungen unterschieden werden: Ergebniserwartung, Selbstwirksamkeitserwartung, Erfolgserwartung und Ergebnis-Folge-Erwartung. Schätzt eine Person ein, ob eine Handlung zu einem gewünschten Ergebnis führt, spricht man von Ergebniserwartung. Schätzt eine Person ein, ob eine zielführende Handlung selbst ausgeführt werden kann, spricht man von der Selbstwirksamkeitserwartung. In der von Bandura (1977) entwickelten Theorie, erweitert er das klassische Reiz-Reaktions-Schema um kognitive Faktoren, die unsere Handlungen bestimmen. Die Wirksamkeitsüberzeugung und die Ergebniserwartung sind dabei die zwei zentralen verhaltenssteuernden Faktoren, über die eine Person im Vorfeld Einschätzungen trifft. Bei der Erfolgserwartung schätzt eine Person die Wahrscheinlichkeit des Erreichens eines Ergebnisses ein. Im Rahmen der Ergebnis-Folge-Erwartung wird davon ausgegangen, dass ein Ergebnis unterschiedliche Folgen nachziehen kann, weshalb zwischen diesen unterschieden werden muss. Die vier Erwartungen geben Aufschluss darüber, wie sicher sich eine Person über Handlungen, Ergebnisse und deren Folgen ist. Die Motivation richtet sich dabei nach der wahrgenommenen Wichtigkeit einer Handlung, wobei Vergnügen und Zukunftsrelevanz eine Rolle spielen. Außerdem werden Aspekte wie Angst zu versagen, empfundene Anstrengung, Zeit- und Energiekosten abgewogen (Urhahne, 2008). Personen entscheiden sich letztlich für Aufgaben, deren Anreizwert sie als hoch empfinden und die mit hoher Wahrscheinlichkeit erfolgreich durchgeführt werden können (Beckmann & Heckhausen, 2010).

Bei der Attributionstheorie werden Meinungen und Haltungen zur Erklärung der Ursache bestimmter Ereignisse betrachtet. Diese Ursachenzuschreibungen werden Attributionen genannt und lassen sich im Kontext Schule beispielsweise bei Prüfungen beobachten. Durch das Nachdenken und das Suchen einer Erklärung von Erfolg oder Misserfolg, können die Lernenden zukünftige Situationen besser einschätzen und ihre Handlungen anpassen. Dabei werden die Ereignisse auf Merkmale der Person oder der Situation zurückgeführt (vgl. Stiensmeier-Pelster & Heckhausen, 2010).

Nach Heider (1958) können die folgenden vier Ursachen ausfindig gemacht werden: internale und externale Gründe, sowie stabile und variable Gründe. Im Rahmen von Leistungssituationen zählen beispielsweise Anstrengung und Fähigkeit zu den der Person eigenen, internalen Gründen, während Zufall und Aufgabenschwierigkeit den externalen Gründen zugeordnet werden, da sie nicht vom Lernenden selbst beeinflusst werden können. Zu den stabilen Dimensionen gehören Fähigkeit und Aufgabenschwierigkeit, zu den zeitlich variablen Dimensionen Anstrengung und Zufall. Nach Weiner (1986) erweist sich es sich als günstig, wenn Erfolg stabilen internalen Ursachen zugeschrieben wird. Misserfolg sollte demnach variablen, externalen Faktoren zugewiesen werden. Schreiben Lernende hingegen Misserfolg stabilen internalen Gründen zu, kann es zu erlernter Hilflosigkeit kommen, was den schlechtesten Fall der Attribuierung darstellt. Attributionstheorien vermögen es also, Konsequenzen von Ursachenzuschreibungen auf Lernergebnisse und damit auch zukünftige Lernprozesse zu betrachten.

Im Rahmen von Studien zur erlernten Hilflosigkeit fanden Dweck und Reppucci (1973), dass Lernende mit gleichen Fähigkeiten verschieden auf Misserfolg reagieren. Im Rahmen ihrer Erklärungen entwickelten sie das Zielorientierungskonzept. Dabei werden das Lernziel und das Performanzziel als Grundformen der Zielorientierung unterschieden. Bei Ersterem ist die Person danach bestrebt, eigene Fähigkeiten und Kompetenzen zu entwickeln und zu verbessern. Dementsprechend werden auch die eigene Intelligenz sowie die eigenen Fähigkeiten als veränderlich aufgefasst. In ihren Untersuchungen fanden Dweck und Reppucci heraus, dass Kinder mit Lernzielen Misserfolge als eine Chance sehen und zu mehr Anstrengung bereit sind, Ziele zu erreichen. Verfolgt die Person ein Performanzziel, so will sie selbst gut dastehen und eigene Kompetenzen und Fähigkeiten gegenüber anderen beweisen beziehungsweise Defizite verbergen. Sie sind eher von der Unveränderlichkeit der eigenen Intelligenz und der eigenen Fertigkeiten überzeugt. In ihrer Studie zeigten Kinder mit Performanzzielen bei anhaltendem Misserfolg hilfloses Verhalten und konnten über Wissensdefizite nicht hinwegtäuschen (vgl. Dweck und Reppucci, 1973).

Die im Kontext dieser Arbeit wohl wichtigsten Theorien zur Lernmotivation sind die über das Konstrukt der intrinsischen Motivation. Nach Schiefele (1996, S. 52) wird intrinsische Motivation definiert als „(...) der Wunsch oder die Absicht, eine bestimmte Lernhandlung durchzuführen, weil die Handlung selbst als interessant, spannend oder sonst wie zufriedenstellend erscheint." Sie wird also um ihrer selbst willen ausgeübt und gilt deshalb als die erstrebenswerteste Lernform. Intrinsisch motivierte Lernende ergreifen von sich aus die Initiative, arbeiten freiwillig an einem Thema, suchen nach Herausforderungen und wenden viel Zeit und Mühe für deren Bewältigung auf (vgl. Lepper & Henderlong, 2000).

Unter den Theorien über die intrinsische Motivation lassen sich drei Erklärungsansätze voneinander unterscheiden:

- die Selbstbestimmungstheorie
- die Flowtheorie
- die Interessenstheorie

Da die Selbstbestimmungstheorie von Deci und Ryan die Grundlage des in dieser Arbeit weiterentwickelten Messinstruments zur intrinsischen Motivation von Schülerinnen und Schülern bildet, wird diese erst im Kapitel 2.3. näher beschrieben.

Im Rahmen der Flowtheorie von Csikszentmihalyi (2010) wird das *Flow*-Erlebnis als vollkommenes Aufgehen in einer Tätigkeit beschrieben. Handelnde fühlen sich, auch trotz hoher Belastung, optimal beansprucht, erleben positive Emotionen und das Gefühl, dass alles „glatt läuft". Ihre gesamte Aufmerksamkeit ist gefesselt, sodass dabei ihr Gefühl für Raum und Zeit verloren geht. Nach Csikszentmihalyi, Abuhamdeh & Nakamura (2005) sind drei Bedingungen notwendig, damit der Zustand des *Flows* eintritt: Die Aufgabe muss über eine klare Zielstruktur verfügen, es braucht genaue und sofortige Rückmeldung und die Anforderungen an die Person müssen im Gleichgewicht zu ihren Fähigkeiten stehen. Personen im *Flow*-Zustand sind in der Lage, Höchstleistungen zu vollbringen. Auch beim Lernen kann man diesen Zustand erleben.

Die letzte Subtheorie zur intrinsischen Motivation ist die Interessenstheorie. Von theoretischer Seite betrachtet entsteht Interesse aus einer Wechselbeziehung zwischen Person und Gegenstand (vgl. Prenzel et al., 1986). Das Interesse ist entscheidend dafür, welchen Inhalten wir uns beim Lernen zuwenden und wie gut wir uns diese merken. Es entsteht aus einer Wechselbeziehung mit dem Lerngegenstand, die durch folgende drei Zustände charakterisiert ist: Der Gegenstand des Interesses hat eine besondere Bedeutung für den Lernenden. Die Lernhandlung ist selbstbestimmt und sie ist von positiven Gefühlen begleitet (vgl. Schiefele, 1996). Weiters unterscheidet man in diesem Zusammenhang zwischen situationalem und individuellem Interesse. Das situationale Interesse ist ein momentaner Zustand und an einzelne Situation gebunden. So kann zum Beispiel jemand an einem Experiment oder einem Kapitel der Physik interessiert sein, ohne sich für die Physik im Allgemeinen zu interessieren. Hingegen besteht individuelles Interesse über die Zeit hinweg, ist nicht nur an eine Situation gebunden und ändert sich dementsprechend gering. Ersteres kann aber auch Ausgangspunkt für ein sich nach und nach entwickelndes, langfristiges Interesse sein (vgl. Urhahne, 2008).

Da eine hohe Anfangsmotivation nicht garantiert, dass ein Ergebnis beziehungsweise eine angestrebte Handlung tatsächlich realisiert wird, befassen sich Volitionstheorien mit der tatsächlichen Umsetzung und damit der Willenskraft eines Individuums. Prozesse der Volition schirmen Handlungen gegen konkurrierende Gedanken und Handlungsalternativen bis zur Zielerreichung ab. In diesem Zusammenhang beschreibt das Rubikon-Modell den vollständigen Handlungsprozess. In der ersten von vier chronologisch aufeinanderfolgenden Phasen werden verschiedene Wünsche abgewogen, wobei diese auf Realisierbarkeit und Bedeutsamkeit geprüft werden. Hier erfolgt auch die Festlegung auf ein zu erreichendes Handlungsziel. Zwischen Zielbildung (1) und Handlung (3) besteht eine Phase, in der die Realisierung (2) durch Kontrolle der Motivation, Aufmerksamkeit und einer guten Strategie zur Bewältigung von Misserfolgen vorangetrieben wird. Die dritte Phase ist charakterisiert durch die Handlung, wobei Intensität und Ausdauer abhängig von der Stärke der Volition sind. Diese wiederum ist umso höher, je größer Realisierbarkeit und Bedeutsamkeit in der ersten Phase eingeschätzt wurden. In der letzten Phase werden die eigene Leistung sowie das Handlungsergebnis analysiert und bewertet. Ist das Handlungsziel erreicht, können neue Ziele verfolgt werden (vgl. Urhahne, 2008).

Die Theorien sozialer Motivation gehen davon aus, dass Menschen durch die Anerkennung anderer Personen wie Familienangehörigen, Lehrpersonen oder Klassenkolleginnen und -kollegen, sowie Freunden motiviert sind. Der Aufbau und die Aufrechterhaltung von sozialen Beziehungen stellt für Lernende einen großen Handlungs- und Lernanreiz dar. Schülerinnen und Schüler erhalten umgekehrt, durch soziale Eingebundenheit, emotionale und kognitive Unterstützung, die dabei hilft, bessere Leistungen zu erzielen (vgl. Urhahne, 2008).

Die eben dargestellten verschiedenen Motivationstheorien zeigen, wie unterschiedlich Lernhandlungen erklärt werden können. Am Beginn scheinen die Bedürfnisse, Einstellungen, Motive und Ziele einer Person zu stehen, während das Interesse, der *Flow*-Zustand, die Selbstbestimmung und soziale Aspekte über die Intensität und Konsequenz der Lernhandlung bestimmen. Erfolg oder Misserfolg am Ende einer Handlung entscheiden über die Gefühle, die damit assoziiert werden und die von Stolz und Freude bis zu Scham und Wut reichen können. Da es meist verschiedene Gründe für eine Lernhandlung gibt und auch im Lernprozess verschiedene Lernmotivationstheorien zur Verhaltenserklärung herangezogen werden können, sollten diese nicht getrennt voneinander betrachtet werden, sondern können (teilweise) als komplementär aufgefasst werden.

2.3. Selbstbestimmungstheorie der Motivation

Wie bereits im letzten Abschnitt erwähnt, folgt nun eine genauere Betrachtung der Selbstbestimmungstheorie von Deci und Ryan (1985), auf deren Grundlage das hier überarbeitete Messinstrument zur intrinsischen Motivation von Schülerinnen und Schülern entwickelt wurde. Sie erlaubt es, Motivation nicht nur allgemein zu quantifizieren, sondern auch deren Qualität im Detail abzubilden und zu analysieren.

Die Selbstbestimmungstheorie der Motivation, kurz SDT, geht von der Annahme aus, dass Menschen den intrinsischen (also angeborenen) Wunsch haben, zu erforschen, zu verstehen und „in sich aufzunehmen“ (assimilieren) (vgl. Deci & Ryan, 1993, S. 235). Dies bildet die Grundlage für die Aneignung kognitiver Fähigkeiten.

Im Zentrum der SDT steht die Annahme eines individuellen Selbst, dessen Entwicklung es speziell im schulischen Umfeld zu fördern gilt. Dies geschieht zum einen durch die Interaktion mit und der Reaktion auf das soziale Umfeld und zum

anderen durch *organismische Integration*. Darunter verstehen Deci und Ryan die aus der Integration von Lebenserfahrung resultierende Weiterentwicklung des Selbst.

Die Autoren sprechen in diesem Zusammenhang davon, dass das individuelle Selbst dialektisch auf sein soziales Umfeld bezogen ist. Dialektisch bedeutet hier, dass Reaktionen des sozialen Umfelds die Entwicklung des Selbst fördern oder behindern können. Andererseits haben individuelle Handlungen auch nachhaltigen Einfluss auf die Umwelt. Eine weitere Grundlage für das Verständnis des individuellen Selbst bildet die Annahme angeborener psychologischer Bedürfnisse, auf die hier noch an späterer Stelle näher eingegangen wird.

Innerhalb der Selbstbestimmungstheorie können grundsätzlich zwei Arten von Motivation unterschieden werden. Die intrinsische, also selbstbestimmte Motivation, die als frei gewählt erlebt wird, und im Gegensatz dazu die extrinsische, aufgezwungene, kontrollierte Motivation. Auf dem Spektrum zwischen diesen beiden Endpunkten empfinden wir Motivation als mehr oder weniger selbstbestimmt.

Im schulischen Umfeld stehen Lehrerinnen und Lehrer oftmals vor der Herausforderung, Lerninhalte vermitteln zu müssen, die interessiertes, freudvolles Erleben und damit intrinsische Motivation nicht unbedingt nahelegen. Hier setzen Lehrpersonen meist auf extrinsische Anreize, um Schülerinnen und Schüler zu motivieren.

Den Kern der Selbstbestimmungstheorie der Motivation bilden die angeborenen psychologischen Bedürfnisse, die sowohl für die intrinsische als auch die extrinsische Motivation gleichermaßen relevant sind. Sie können uns darüber Auskunft geben, warum wir Handlungszielen motiviert nachgehen.

Nach Deci und Ryan (1985) können die folgenden drei psychologischen Grundbedürfnisse unterschieden werden:

1. **Kompetenz / Wirksamkeit**: äußert sich im natürlichen Bestreben des Individuums, sich als handlungsfähig zu erleben. Menschen nehmen sich als kompetent war, wenn sie durch die ihnen gestellten Aufgaben weder über- noch unterfordert werden. Dabei spielt regelmäßiges, positives Feedback eine entscheidende Rolle.

2. **Selbstbestimmtheit / Autonomie**: äußert sich im natürlichen Bestreben, sich als eigenständiges Handlungszentrum zu erleben. Das Erleben von Selbstbestimmtheit ist umso größer, je öfter Entscheidungsmöglichkeiten vorliegen und die Person das Gefühl hat, eine Wahlmöglichkeit zu haben sowie eine Stimme, die gehört wird.
3. **Zugehörigkeit / soziale Eingebundenheit**: äußert sich im starken Bestreben nach befriedigenden Sozialkontakten. Da Lernen immer vor einem kulturellen Kontext und einem sozialen Hintergrund stattfindet, ist eine positive Verbundenheit mit dem Umfeld ein zentraler Antrieb motivierten Verhaltens.

(vgl. Krapp, 2005)

Besonderer Bedeutung für die Motivation kommt der sozialen Umgebung zu. Speziell für den Prozess der Internalisierung und Integration extrinsischer Motivation spielt das soziale Umfeld eine wichtige Rolle. Solche Prozesse werden gefördert, wenn erwachsene Bezugspersonen das Autonomiebestreben der Heranwachsenden unterstützen und innere Anteilnahme zeigen (vgl. Grolnick & Ryan, 1989). Persönliche Anteilnahme wirkt sich also positiv auf die Selbstständigkeit der Kinder in Bezug auf ihre Schularbeiten aus. Auch Ryan und Powelson (1991) weisen auf die Bedeutung der Verhaltensweisen der Eltern auf das Lerngeschehen hin.

Intrinsische Motivation

Als intrinsische Motivation wird ein Zustand bezeichnet, der den inneren Neigungen und Wünschen der handelnden Person entspricht und für dessen Aufrechterhaltung keine extrinsischen Anreize (z.B. Lohn oder Strafe) erforderlich sind (vgl. Schiefele & Köller 1998).

Intrinsisch motivierte Handlungen sind interessensbestimmt und werden frei von äußerem Druck oder inneren Zwängen gesetzt. Das Individuum wählt was der eigenen Auffassung und den eigenen Interessen entspricht und handelt daher auch engagiert. Es sind weder äußere Anreize für das Handeln notwendig, noch ist dieses unbedingt auf ein Ergebnis hin orientiert. Die Handlung selbst ist das Motiv.

Intrinsische Motivation entsteht, wenn die drei oben ausgeführten, grundlegenden, psychologischen Bedürfnisse, also Kompetenzerleben, Autonomieerleben und soziale Eingebundenheit befriedigt werden.

Ein Zustand des allgemeinen Wohlbefindens, etwa ausgelöst durch rein spielerische Tätigkeiten, ohne den Aspekt einer Kompetenzerfahrung beziehungsweise eines durch Leistung zielgerichteten Handelns und Lernens, ist in diesem Zusammenhang nicht gemeint. Es geht also um ein positives emotionales Erleben in Erziehung und Unterricht.

Die Umfelder Schule und Familie können intrinsische Motivation durch Unterstützen von Autonomie und Kompetenz fördern, jedoch nur, wenn ein grundsätzliches intrinsisches Interesse am Lerninhalt vorhanden ist.

Extrinsische Motivation

„Extrinsic motivation is a construct that pertains whenever an activity is done in order to attain some separable outcome" (Ryan & Deci, 1999, S. 60). Extrinsische Verhaltensweisen sind beispielsweise durch Belohnung oder Bestrafung motiviert. Das handlungsleitende Motiv ist ein „ich muss". Das Maß in dem extrinsische Motivation als autonom wahrgenommen werden, variiert.

So kann eine Schülerin beziehungsweise ein Schüler ausschließlich aus Angst vor schlechten Noten, oder Sanktionen durch die Eltern, Hausaufgaben erledigen, oder aber, weil er oder sie Hausaufgaben für notwendig hält, um ein von ihm oder ihr angestrebtes Studium oder einen Berufswunsch zu verfolgen. Beides sind Formen extrinsisch motivierten Verhaltens, wobei zweiteres durchaus als vom Individuum selbstbestimmt wahrgenommen werden kann.

Eine differenzierte Betrachtung extrinsischer Motivation bietet die *Organismic Integration Theorie* (OIT) (vgl. Deci & Ryan, 1985). Sie unterteilt die extrinsische Motivation nach dem Grad der Selbstbestimmtheit und dem Ausmaß an Kontrolliertheit in 4 Typen der Verhaltensregulation:

1. **Externe Regulierung (fremdbestimmt)**
2. **Introjiziert (eher fremdbestimmt)**
3. **Identifiziert (eher autonom)**
4. **Integriert (autonom)**

Das Maß an Autonomie, mit der die motivierten Handlungen ausgeführt werden, wird hier als der „subjektiv wahrgenommene interne Anteil“ beschrieben. Deci und Ryan (1985) bezeichnen den wahrgenommenen Sitz der Entscheidung als *perceived locus of causality*, der sich nach außen oder innen verschieben kann. Extrinsisch motivierte Verhaltensweisen können durch Prozesse der Internalisierung und Integration in selbstbestimmte Handlungen überführt werden.

1. Die **externale Regulation** ist durch Verhaltensweisen gekennzeichnet, die von externen Steuerungsfaktoren wie Belohnung oder negativen Konsequenzen abhängig sind.

2. Die **introjizierte Regulation** folgt einem inneren Anstoß oder Druck. Man tut etwas, weil „es sich so gehört“, oder zum Beispiel, weil man sonst ein schlechtes Gewissen hätte, nicht aber weil es ein eigenes Verlangen ist. Soziale Anerkennung, soziale Normen und Werte werden als persönlich wichtig befunden und bilden den Antrieb der Motivation.

3. Die **identifizierte Regulation** ist gegeben, wenn eine persönliche Relevanz, ein Verhalten, als wichtig oder wertvoll erachtet und es daher in das individuelle Selbstkonzept integriert werden kann. Zum Beispiel eine Schülerin oder ein Schüler lernt für eine Prüfung, weil er einen Umstieg in eine bestimmte Schulform anstrebt, nicht aber, weil Lehrer, Eltern oder Freunde das von ihm erwarten.

4. Die **integrierte Regulation** hat den höchsten Grad an Selbstbestimmung. Das Handlungsergebnis wird subjektiv hoch bewertet. Das Maß, in dem das Handeln als selbstbestimmt wahrgenommen wird, ist hier von allen extrinsisch motivierten Verhaltensregulationen am höchsten.

Die differenzierte Sichtweise auf extrinsische Motivation zeigt, dass intrinsische und extrinsische Motivation nicht unbedingt Gegensätze sind. Da Menschen sich mit ihrem sozialen Umfeld verbunden fühlen, gleichzeitig ihr Handeln aber auch als autonom empfinden wollen, übernehmen sie Verhaltensformen oder Ziele anderer in ihr eigenes Selbstkonzept. Extrinsisch motivierte Verhaltensweisen können durch Prozesse der Internalisierung und Integration in selbstbestimmte Handlungen überführt werden (vgl. Meissner, 1981).

Extrinsische Motivation kann sich aber auch negativ auswirken. Empirische Untersuchungen haben gezeigt, dass intrinsische Motivation abnimmt, wenn man Versuchspersonen extrinsische Belohnung zum Beispiel Geld oder Auszeichnung anbietet (vgl. Deci 1971 und Ross, 1975). Spätere Studien zeigten aber auch, dass extrinsische Belohnung intrinsische Motivation aufrechterhalten kann (vgl. Harackiewicz, 1979 und Ryan, 1982)

Aus der bisherigen Betrachtung geht hervor, dass eine intrinsische Motivation die optimale Voraussetzung für Verhaltensweisen und Handlungen und somit auch für das Lernen ist. Für effektives Lernen und gute Lernergebnisse sind selbstbestimmte Formen der Handlungsregulation wichtig. Diese kann auch durch extrinsische Motivation, die integriert, beziehungsweise in eine selbstregulierte Motivation eingegliedert wird, erreicht werden.

Hingegen werden Handlungen als amotiviert bezeichnet, wenn diese keiner Intension folgen, also keinem Ziel nachgegangen wird sowie das unkontrollierte Handeln. Personen mit einer amotivierten Verhaltensweise agieren meist passiv und orientierungslos. Ziele erscheinen unerreichbar weshalb Person in einen Zustand der Motivationslosigkeit fallen.

Selbstbestimmte Motivation – Lernen

Im Rahmen von Lernumgebungen wird die Motivation durch Folgende Dinge gefördert:

- Ein optimales Anforderungsniveau, also Aufgaben, die nicht als zu leicht oder zu schwer wahrgenommen werden.
- Positives Feedback, wenn sich dieses auf Sachverhalte bezieht, die aus einer selbstbestimmten Handlung resultieren und nicht kontrollierend wirken.
- Persönliche Anteilnahme durch „signifikante“ Erwachsene, also wenn eine innere Beteiligung von Erwachsenen die Autonomiebestreben von Heranwachsenden unterstützt.
- die Qualität der sozialen Beziehungen zu Eltern und Lehrern
- Interesse am Lerngegenstand
- autonomieunterstützende Lernumgebungen

(vgl. Deci & Ryan, 1993)

Zusammenfassend begründet sich die Selbstbestimmungstheorie der Motivation auf dem angeborenen Wunsch des Menschen, seine Umwelt zu erforschen und in sich aufzunehmen und sich dabei als Handlungszentrum selbstbestimmt zu erleben. Lernmotivation sollte diesen Erfordernissen des individuellen Selbst entsprechen. Die Befriedigung psychologischer Bedürfnisse spielt eine wichtige Rolle. Intrinsische oder integriert extrinsische Motivationen unterstützen maßgeblich das Erlangen von Kompetenzen. Extrinsische Motivation ist in dem Maß, in dem sie als aufgezwungen oder als kontrolliert wahrgenommen wird, beim Lernen hinderlich. Die Autonomiebestrebungen Heranwachsender werden insbesondere durch soziale Bedingungen und anteilnehmende Bezugspersonen unterstützt.

2.4. Intrinsic Motivation Inventory (IMI)

Das von Deci und Ryan (2003) auf Basis ihrer Selbstbestimmungstheorie entwickelte *Intrinsic Motivation Inventory*, kurz IMI, ist ein multidimensionales Messinstrument zur Bestimmung der Selbstregulation und der intrinsischen Motivation von Personen. Das IMI ist in englischer Sprache verfasst und wurde bisher vor allem im Kontext sportlicher Aktivitäten, wie Mannschaftssport aber auch Einzelsport, eingesetzt und getestet. Wenig Evidenz gibt es hingegen für dessen Eignung in anderen Kontexten, etwa in naturwissenschaftlichen Unterrichtsgegenständen.

Das IMI setzt sich aus insgesamt 45 Items zusammen, wobei Befragte zu den Aussagen der Items im Rahmen einer siebenstufigen Antwortskala diesen zustimmen oder sie ablehnen können. Die Items werden in folgenden sieben Subskalen unterteilt, welche ich anschließend näher beschreibe (vgl. Deci und Ryan, 1993 & 2003):

- interest and enjoyment – Interesse und Vergnügen (7 Items)
- perceived choice – wahrgenommen Wahlfreiheit (7 Items)
- perceived competence – (selbst) wahrgenommene Kompetenz (6 Items)
- effort and importance – Einsatz und Wichtigkeit (5 Items)
- value and usefulness – Nützlichkeit und Wert (7 Items)
- (social) relatedness - Eingebundenheit (8 Items)
- felt pressure and tension – wahrgenommener Druck und Anspannung (5 Items)

Die erste Skala *interest and enjoyment* ist jene, welche nach Angabe der Autoren, die intrinsische Motivation selbst abbildet, weswegen sie meist mehrere Items umfasst als die anderen Skalen.

Die Konstrukte *perceived choice* und *perceived competence* sind positive Prädiktoren, also Vorhersagevariablen für die intrinsische Motivation und für das Selbstkonzept einer Person. Das zur Skala *perceived competence* zugehörige psychologische Grundbedürfnis ist das Kompetenzerleben. Während die Skala (felt) *pressure and tension* einen negativen Prädiktor für das Selbstkonzept bildet.

Die Skala *effort and importance* ermöglicht es, die Motivation zu quantifizieren. Hier werden Personen danach gefragt, wie wichtig eine Tätigkeit für sie ist und wieviel an Fleiß und Anstrengung sie bereit sind dafür zu investieren. Damit bildet diese Subskala ein eigenes Konstrukt ab, während die anderen Skalen die Ursache, sowie den Grad von Selbstbestimmtheit und intrinsischer Motivation messen.

Wird eine Handlung als nützlich und wertvoll erachtet, so wird diese eher verinnerlicht. Damit wird selbstgesteuertes Verhalten gefördert. Die Skala *value and usefulness* will diese Internalisierungsprozesse erfassen.

Im Rahmen der Skala *relatedness* geht es um zwischenmenschliche Beziehungen, die sich ebenfalls auf die intrinsische Motivation auswirken. Personen, die sich als sozial zugehörig beziehungsweise eingebunden fühlen, sind nach Deci und Ryan motivierter, etwas zu bewirken.

Die Items des IMIs sind nach Deci und Ryan (2003) für eine Vielzahl von Aufgaben, unter unterschiedlichen Rahmenbedingungen, faktorenanalytisch kohärent und stabil. Ihren Studien nach laden die einzelnen Items mit einem Faktor von mindestens 0,6 auf die jeweilige Skala und weisen lediglich Querladungen (das sind unerwünschte Ladungen auf andere Faktoren) kleiner 0,4 auf andere Skalen auf.

Für den Einsatz des IMIs für eigene Forschungszwecke empfehlen Deci und Ryan (2003) nur jene Subskalen zu wählen die relevant erscheinen, wobei zu eigenen Faktoranalysen geraten wird. Dabei ermutigen sie auch dazu, Items an persönliche Zwecke zu adaptieren und einzelne Items auszuschließen, da in jeder Skala auch

immer redundante Items vorhanden sind. Hier lautet die Empfehlung, etwa vier Items pro Skala zu wählen, da jedes weitere lediglich zu einer kleinen Verbesserung führt.

2.5. Test- und Fragebogenkonstruktion

Um in weiterer Folge einen Fragebogen zur Messung der intrinsischen Motivation weiterentwickeln zu können, sollen in diesem Kapitel Eigenschaften und Kriterien zur Entwicklung eines guten psychometrischen Tests, sowie Analysemethoden diskutiert werden. Nach Bühner (2011, S. 38) handelt es sich bei einem psychometrischen Test um ein wissenschaftliches Verfahren, mit dem ein oder mehrere empirisch abgrenzbare Persönlichkeitsmerkmale untersucht werden. Ziel des Tests ist es, die Ausprägung einer Eigenschaft oder einer Fähigkeit zu bestimmen und dabei genau definierten Haupt- und Nebengütekriterien zu genügen. Diese sollen nun beschrieben werden.

2.5.1. Hauptgütekriterien

Zu den Hauptgütekriterien gehören Objektivität, Reliabilität, Validität und die Skalierbarkeit eines Tests.

Objektivität

Die Objektivität gibt den Grad der Unabhängigkeit eines Tests vom Untersucher beziehungsweise der Untersucherin an. Dabei werden drei Arten von Objektivität unterschieden:

1. Damit die **Durchführungsobjektivität** gegeben ist, muss genau festgelegt werden, wie und unter welchen Bedingungen ein Test durchzuführen ist. Dies umfasst die Angabe von Zeitbegrenzung, erlaubten Hilfsmitteln und Hilfestellungen. Ein besonderes Augenmerk muss auf die Wortwahl der Instruktionen gelegt werden, denn dadurch können Rückfragen auf ein Minimum reduziert werden.
2. Um unabhängig von der auswertenden Person gleichwertige Ergebnisse zu erhalten, müssen, im Rahmen der **Auswertungsobjektivität**, genaue Vorschriften vorgegeben werden. Hier eignen sich Schablonen oder Lösungsblätter mit akzeptierten Antwortalternativen. Geschlossene Antwortformate, wie eine Likert-Skala, weisen eine höhere Auswertobjektivität als offene Formate auf.

3. Bei der **Interpretationsobjektivität** sollen die Auswerterinnen und Auswerter zur gleichen Beurteilung beziehungsweise Interpretation der Ergebnisse gelangen. Dafür sind eine ausreichend große Normstichprobe, sowie überprüfte Haupt- und Nebengütekriterien nötig, um anhand des gleichen Maßstabs zu messen. Weiters sind standardisierte Interpretationen zu empfehlen.

(vgl. Bühner, 2011)

Reliabilität

Die Reliabilität gibt den Grad der Messgenauigkeit eines Messwertes an, unabhängig davon, was er zu messen beansprucht. Nach Bühner (2011, S. 144) lässt sich Reliabilität als „(...) der Anteil der Varianz der wahren Werte an der Varianz der beobachteten Werte (...)“ definieren. Auch die Korrelationskoeffizienten können unter bestimmten Bedingungen als Basis der Reliabilitätsabschätzung dienen. Je nach Voraussetzung wird eine der folgenden Schätzmethoden gewählt, wobei hier nur die geläufigsten angeführt sind:

1. **Konsistenzmethode**

   Dabei beschäftigt man sich mit den Korrelationen beziehungsweise Kovarianzen von Items beziehungsweise von Testteilen. Man kann innerhalb der Konsistenzmethoden zwischen Testhalbierungskoeffizienten sowie inneren oder internen Konsistenzkoeffizienten unterschieden. Im Rahmen dieser Arbeit sind die Konsistenzkoeffizienten von Bedeutung, sie hängen von der Korrelation zwischen den einzelnen Items ab. Das **Cronbach-Alpha** ist dabei ein Wert, der die Itemkorrelationen quantifiziert und dabei auch die Anzahl der Items berücksichtigt.

2. **Paralleltestmethode**

   Bei der Paralleltestmethode führt man zwei Tests bei derselben Stichprobe durch, die dasselbe messen, jedoch unterschiedliche Items besitzen. Anschließend betrachtet man ihre Korrelation.

3. **Retestmethode**

   Im Rahmen der Retestmethode wird ein und derselben Stichprobe der gleiche Test nach einer bestimmten Zeit erneut vorgelegt. Anschließend wird auch hier die Korrelation zwischen den Tests betrachtet.

(vgl. Bühner, 2011)

Validität

Die Validität gibt an, ob der Test auch wirklich das misst, was er zu messen beansprucht. Man kann nach Bryant (2000) folgende drei Validitätsarten unterscheiden: Inhaltsvalidität, Kriteriumsvalidität und Konstruktvalidität.

1. **Inhaltsvalidität**

   Die Inhaltsvalidität gibt an, ob die zur jeweiligen Subskala gehörenden Items das zu messende Konstrukt hinreichend genau erfassen. Dabei müssen die gewählten Items einerseits die Breite des Konstrukts abbilden und somit repräsentativ für alle möglichen Items sein. Andererseits dürfen sie auf kein anderes Konstrukt laden, sprich kein anderes Konstrukt abbilden. Die Inhaltsvalidität lässt sich nicht nummerisch messen, sondern wird aufgrund logischer fachlicher Überlegungen ermittelt. Im Rahmen dieser Arbeit soll die Inhaltsvalidität mit Hilfe des ***Four Building Blocks Approach*** von Wilson (2005), welchen ich in Kapitels 2.6.3 vorstelle, gewährleistet werden. Es wird davon ausgegangen, dass die Durchführung der vier Schritte eine sorgfältige Repräsentation des Inhalts der Konstrukte sichert.

2. **Kriteriumsvalidität**

   Sie klärt über den Zusammenhang der Testleistung mit einem oder mehreren Kriterien, mit denen der Test aufgrund seines Messanspruchs korrelieren sollte, auf. Ermittelt wird sie auf Basis der Zusammenhänge zwischen Testkennwerten und Kriterien.

3. **Konstruktvalidität**

   Die Konstruktvalidität soll Auskunft darüber geben, ob ein Test auch die Eigenschaft misst, die er messen soll. Im Unterschied zur Inhaltsvalidität können hier homogene, konstruktnahe Inhaltsbereiche mittels konfirmatorischer Faktoranalyse zusammengefasst und von anderen

konstruktfremden Bereichen getrennt werden. Näheres zur Faktorenanalyse folgt in den Kapiteln 2.5.3. am Ende dieses Abschnitts.

(vgl. Bühner, 2011)

Skalierbarkeit eines Tests

Durch die Skalierbarkeit eines Tests soll geprüft werden, ob die Bildung eines Testwerts anhand einer gültigen Verrechnungsvorschrift vorgenommen wird. Im Rahmen dieser Arbeit sollte hier dementsprechend überprüft werden, ob das einfache Aufsummieren und Vergleichen von Itemscorewerten die dahinterliegenden empirischen Unterschiede zwischen Schülerinnen und Schülern hinsichtlich ihrer intrinsischen Motivation angemessen widerspiegelt. Dabei ist die Verrechnung der Items zu Punkten abhängig von der Itemtrennschärfe. Sie ist ein Kennwert dafür, wie viel das Item zur Subskala beziehungsweise zum Gesamtwert beiträgt (vgl. Bühner, 2011).

2.5.2. Nebengütekriterien

Neben den vier eben genannten Hauptgütekriterien gibt es noch eine Reihe weiterer Indikatoren für die Güte eines guten psychometrischen Tests.

Normierung

Ein Test wird als normiert bezeichnet, wenn es die Möglichkeit gibt, eine individuelle Leistung, beziehungsweise eine Ausprägung, anhand anderer Stichproben einzuordnen. Durch den Vergleich mit der Referenzgruppe lässt sich sagen, ob diese Leistung, beziehungsweise Ausprägung, überdurchschnittlich, durchschnittlich oder unterdurchschnittlich ist (vgl. Bühner, 2011).

Nützlichkeit

Ein Test ist nützlich, wenn er ein Persönlichkeitsmerkmal oder eine Verhaltensweise misst, beziehungsweise vorhersagt, für die ein praktisches Bedürfnis besteht. Dies ist gegeben, wenn entweder noch kein Testverfahren mit dem entsprechenden Messanspruch vorliegt oder das neue Verfahren Vorteile hinsichtlich einer Verbesserung der Gütekriterien vorweist (vgl. Bühner, 2011).

Nicht-Verfälschbarkeit

Ein Test sollte durch die getestete Person nicht willentlich oder unwillentlich in eine bestimmte Richtung verfälscht werden können. Dieses Nebengütekriterium gilt jedoch als unrealistisch, da dies immer möglich ist (vgl. Bühner, 2011).

Vergleichbarkeit

Ein Test gilt als vergleichbar, wenn Tests, die ähnliches messen sollen, existieren und ähnliche Ergebnisse liefern. Weiters ist Vergleichbarkeit gegeben, wenn eine Person bei mehrmaligem Testen ähnlich Ergebnisse erzielt. Bei Letzterem muss berücksichtigt werden, dass, abhängig vom Test, bei Wiederholung meist eine Steigerung in den Leistungen zu erwarten ist (vgl. Bühner, 2011).

Fairness

Ein Test gilt als fair, wenn die resultierenden Messwerte einzelne Gruppen nicht diskriminieren (vgl. Bühner, 2011).

Ökonomie

Damit ein Test als ökonomisch gilt, muss er in angemessener Arbeitszeit bewältigbar sein, wenig Material verbrauchen, einfach zu handhaben sein, für eine Gruppenmessung tauglich und seine Auswertung schnell und bequem sein (vgl. Bühner, 2011).

Zumutbarkeit

Ein Test ist zumutbar, wenn die Getesteten durch die Testung nicht unnötig belastet werden. Dies umfasst sowohl unverhältnismäßige körperliche als auch psychische Beanspruchung (vgl. Bühner, 2011).

Trennschärfe

Die Trennschärfe gibt an, wie gut ein Item eine gewisse Ausprägung beziehungsweise Eigenschaft misst und wird mittels einer Korrelation quantifiziert. Derartige Trennschärfen werden als Eigentrennschärfen bezeichnet. Dabei begünstigen hohe Eigentrennschärfen der Items eine hohe Reliabilität des Tests. Rechnerisch erhält man die Eigentrennschärfen als die Korrelation eines Items mit den Summenwerten der übrigen Items einer Subskala. Items mit einer Trennschärfe zwischen 0,3 und 0,5

werden als mittel eingestuft, bei Werten >0,5 spricht man von einer hohen Trennschärfe (vgl. Bühner, 2011).

2.5.3. Faktoranalysen

Im Rahmen einer **exploratorischen Faktoranalyse** geht man der Frage nach, wie viele Eigenschaften, Merkmale oder Fähigkeiten die Items messen, beziehungsweise welche Items Indikatoren bestimmter Eigenschaften, Merkmale oder Fähigkeiten darstellen.

Die sogenannten Faktoren sind in diesem Zusammenhang die latenten Variablen, also nicht direkt messbare Eigenschaften, Merkmale oder Fähigkeiten, die durch mehrere Items charakterisiert werden können. Bei der exploratorischen Faktoranalyse wird das Ziel verfolgt, Zusammenhänge der Items untereinander, durch möglichst wenige Faktoren zu erklären, wobei der Verlust an Information so gering wie möglich gehalten werden soll. Dies geschieht im Rahmen einer Hauptkomponentenanalyse, bei welcher die Items so zusammengefasst werden, dass möglichst viel Information aus den ursprünglichen Daten durch die zusammengefassten Variablen beziehungsweise Komponenten beschrieben wird. Dabei werden die Korrelationen der Items untereinander betrachtet. Die sogenannte Ladung von einem Item auf einen Faktor ist ein Maß dafür, wie eng das Item mit dem Faktor zusammenhängt (vgl. Bühner, 2011).

Um die Anzahl der sinnvoll zu extrahierenden Faktoren zu ermitteln, gibt es mehrere Methoden, wobei die folgenden zwei für den Kontext dieser Arbeit relevant sind:

Nach dem Kaiser-Gutman Kriterium werden nur jene Faktoren beibehalten, deren Eigenwert (das ist die Summe der quadrierten Ladungen auf einen Faktor über alle Items) größer eins ist. Sie erklären mehr Varianz als die ursprünglichen Faktoren. Nach Bortz und Schuster (2010) werden beim Kaiser-Gutman Kriterium meist mehr Faktoren extrahiert, als inhaltlich begründet werden können. Dieses Problem lässt sich bei der Auswertung in SPSS umschiffen, indem man im Rahmen der Hauptkomponentenanalyse eine feste Anzahl von Faktoren erzwingt.

Eine andere Methode, die Anzahl der sinnvoll zu extrahierenden Faktoren zu ermitteln, stellt der Scree-Test nach Cattell da, bei welchem man das Eigenwertdiagramm betrachtet. Dabei werden die Eigenwerte ihrem Betrag nach absteigend geordnet und

mittels Rangnummer in einem Graphen veranschaulicht. Der Scree-Test besagt, dass jene Faktoren als bedeutsam angesehen werden können, die vor dem Knick im Funktionsgraph liegen (vgl. Bortz und Schuster, 2010). Da es meist nicht eindeutig ist, welcher Knick betrachtet werden muss und diese Einschätzung subjektiv ist, dient dieser Test für diese Arbeit lediglich als zusätzliche Information beziehungsweise als Absicherung.

Nach Deci und Ryan (2003) ist ein Item aus mathematisch-statistischen Gesichtspunkten ab einer Faktorladung von ± 0,6 einer Skala zugehörig, wobei keine Querladungen (das sind unerwünschte Ladungen auf andere Faktoren) ± 0,4 auf andere Skalen auftreten sollen. Ein Blick in die naturwissenschaftsdidaktische Forschung, etwa bei Bühner (2011, S. 80), zeigt, dass die Faktorladungen ± 0,7 betragen sollten, damit ein Item ausreichend reliabel ist. Ebenso scheint der zweite als weicheres Kriterium. Nach Busker (2014) sollen bereits jene Items ausgeschlossen werden, die Querladungen von ± 0,3 aufweisen.

Die **konfirmatorische Faktoranalyse** dient dazu, eine theoretisch angenommene Faktorenstruktur zu überprüfen. Dazu werden Kovarianzmatrizen analysiert und mit der angenommen Faktorstruktur verglichen. Angenommen, das zu testende Konstrukt ist die intrinsische Motivation, welches durch mehrere Faktoren charakterisiert ist, und deren Struktur es zu kontrollieren gilt. Dieses Konstrukt stimmt gut mit den erhobenen Daten über ein, wenn inhaltlich zusammengehörige Items hoch auf denselben Faktoren laden und nur geringe Querladungen aufweisen.

2.5.4. Weitere statistische Analyse- und Testmöglichkeiten

Um Aussagen zur intrinsischen Motivation über unterschiedliche Gruppen zu machen und diese anschließend zu vergleichen, bedarf es geeigneter Test- und Analysemöglichkeiten.

Der sogenannte t-Test ist dafür geeignet, zwei unabhängige Stichproben, hinsichtlich eines Merkmals, einander gegenüberzustellen. Um ihn durchführen zu können, müssen folgende Voraussetzungen erfüllt sein: Die zwei Untergruppen müssen voneinander unabhängig sein, es muss Homogenität der Varianz vorliegen und das untersuchte Merkmal, hier die intrinsische Motivation, muss in der Grundgesamtheit normalverteilt sein (vgl. Bühner, 2011).

Mittels Levene-Tests kann die Gleichheit der Varianzen von zwei Gruppen überprüft werden, wofür jeweils sie Summe der Quadrate der Abweichungen vom Mittelwert für beiden Gruppen berechnet und verglichen wird. Er ist ein Signifikanz Test, dessen Nullhypothese davon ausgeht, dass die Fehlervarianzen über die Gruppen hinweg gleich sind (vgl. Bortz und Schuster, 2010).

Sollen mehr als zwei unabhängige Stichproben gleichzeitig, hinsichtlich einer Merkmalsausprägung, miteinander verglichen werden, so kann eine Analysis of Variances (ANOVA) durchgeführt werden. Auch hier müssen vorab bestimmte Grundvoraussetzungen überprüft werden. Die verschiedenen Gruppen müssen unabhängig voneinander sein, das untersuchte Merkmal muss in der Grundgesamtheit normalverteilt sein und die Varianz innerhalb der Stichproben muss jeweils gleich groß sein, es muss also Varianzhomogenität vorliegen. Der einfaktoriellen ANOVA liegt die F-Verteilung zugrunde. Im Vorfeld einer ANOVA kann die Varianzhomogenität mittels Levene-Tests überprüft werden.

2.6. Messinstrumente der Motivation

Nach dem nun ausführlich auf die Eigenschaften und Kriterien zur Entwicklung eines validen psychometrischen Tests, sowie auf unterschiedliche Analysemethoden eingegangen wurde, soll in diesem Kapitel geklärt werden, welche Messinstrumente es bereits im deutschsprachigen Raum zur Motivation gibt und warum mit dieser Arbeit der Versuch unternommen wurde, ein eigenes Messinstrument zu konstruieren.

Den Ausgangspunkt für diese Betrachtung stellt die Dissertation „Cross Age Peer Tutoring in Physik“, kurz CAPT, von Korner (2015) dar. Dabei wollte die Autorin Lernumgebungen hinsichtlich motivationaler und anderer nicht-kognitiver Parameter, wie zum Beispiel der Selbstwirksamkeitserwartung, analysieren. Nachdem das Konstrukt der Selbstbestimmungstheorie von Deci und Ryan sowohl intrinsische als auch extrinsische Faktoren berücksichtigt, erschien es als die geeignetste theoretische Grundlage.

Der Versuch einer Übersetzung des *Intrinsic Motivation Inventorys* ins Deutsche erwies sich als nicht zielführend, da sich zeigte, dass die Gütekriterien eines validen psychometrischen Tests selbst nach einigen Überarbeitungsschleifen nicht erfüllt werden konnten. Aus diesem Grund entschloss sich Korner (2015) dazu, die Items zu

den einzelnen Skalen von Grund auf neu zu konstruieren. Dabei wählte sie als Methode den *Four Building Blocks Approach* nach Wilson (2005). Er soll gewährleisten, dass die Merkmalsausprägung aufgrund von nach außen beobachtbaren Evidenzen eingeschätzt wird, wodurch die Objektivität gesteigert werden soll. Die zu untersuchende Ausprägung ist im Rahmen dieser Arbeit die intrinsische Motivation von Lernenden, die eine latente (= nicht messbare) Variable darstellt.

Im Anschluss an die Neukonstruktion der Skala *effort and importance* von Korner (2015) folgten Arbeiten zu den Skalen *interest and enjoyment* von Klingenböck (2016), *value and usefulness*, *perceived choice* und *perceived c*ompetence von Schmidt (2017) sowie zur Skala (*social) relatedness* von Pieler (2018). Ziel der vorliegenden Arbeit ist es, einerseits die beiden stark korrelierenden Skalen *perceived competence* und *pression and tension* näher zu analysieren und andererseits zu überprüfen, ob die Skalen bei gleichzeitiger Messung aller Items trennscharf und stabil sind. Damit soll ein weiter Schritt zur Finalisierung des ersten deutschsprachigen Testinstruments zur Messung der Motivation von Lernenden gesetzt werden.

2.6.1. Bestehende Messinstrumente

Da bereits einige Fragebögen zur Motivation im deutschsprachigen Raum bestehen, diese jedoch aus unterschiedlichen Gründen als nicht geeignet erscheinen, folgt hier ein Überblick.

Der Fragebogen von Müller et al. (2007) beruht auf der Selbstbestimmungstheorie der Motivation von Deci und Ryan. Er vermag es, unterschiedliche Formen extrinsisch motivierten Verhaltens zu beschreiben, gibt jedoch keine befriedigenden Auskünfte über den hinter einer Handlung liegenden Grund. Er besteht aus den Skalen identifizierte Regulation, intrinsische Regulation, introjizierte Regulation und externale Regulation mit jeweils vier bis 5 Items. Im Rahmen einer Studie mit über 2600 Schülerinnen und Schülern zeigte der Fragebogen sowohl eine gute Reliabilität als auch faktoranalytische Trennschärfe der Skalen.

Ein weiteres Messinstrument zur Motivation findet man bei Berger und Hänze (2004), die zur Analyse der Wirkung von Gruppenpuzzles im Physikunterricht einen Fragebogen konstruierten. Dabei interessierten sie sich vor allem für die

Motivationssteigerung, die sie durch das Erfüllen psychischer Grundbedürfnisse im Rahmen der Unterrichtsmethode des Gruppenpuzzles erwarteten. Entsprechend der Grundbedürfnisse der Selbstbestimmungstheorie nach sozialer Eingebundenheit, Kompetenzerleben und Autonomieerleben wurden drei Skalen mit jeweils zwei Items entwickelt. Im Rahmen ihrer Studien zeigten einige Items bei der Faktoranalyse mäßige Ladungen auf die jeweilige Skala, verbunden mit hohen Querladungen auf andere Skalen. Ein weiterer Kritikpunkt am Instrument von Berger und Hänze ist die Aussagekraft des Fragebogens, die durch seine Kürze nur grobe Schlüsse auf die Motivation von Schülerinnen und Schülern zulässt. Außerdem ist er für die Sekundarstufe 2 konzipiert und könnte für den hier angepeilten Einsatz in der Sekundarstufe 1 zu Verständnisschwierigkeiten führen.

Zur Evaluierung des Lernens an außerschulischen Lernorten wurde die Kurzskala intrinsischer Motivation, auch KIM genannt, von Wilde et al. (2009) entwickelt. Als theoretischen Rahmen wählten die Autoren die Selbstbestimmungstheorie von Deci und Ryan, da sie es am besten vermag Kompetenzerleben und Autonomieerleben abzubilden. Der aus zwölf Items bestehende Fragebogen ist eine Kurzversion des IMIs und umfasst die Skalen interest and enjoyment, perceived competence, perceived choice und pressure and tension. Im Rahmen einer Faktorenanalyse ließen sich die Skalen trennscharf voneinander abgrenzen und auch hinsichtlich der Reliabilität lieferte der Fragebogen ausreichende Werte. Als problematisch anzusehen ist die Übersetzung mancher Items aus dem Englischen, wodurch sie inhaltlich nicht mehr ihrer ursprünglichen Skala zugeordnet werden können. Außerdem haben sich die Autoren für das Weglassen der Skala effort und importance entschieden, welche die Anstrengungsbereitschaft erfasst. Da diese jedoch ein wesentliches Konstrukt darstellt, erscheint die KIM als unzureichendes Instrument, die intrinsische Motivation in ausreichender Güte zu erfassen.

Da keines der im deutschsprachigen Raum vorhanden Messinstrumente den Anforderungen für die Evaluation von CAPT genügt, entschloss sich Korner dafür das IMI ins Deutsche zu übersetzen. Dabei sollte das Alter der Lernenden sowie ihre Lesekompetenz berücksichtigt werden und der Fragebogen über eine klare und verständliche Sprache verfügen. Außerdem soll er den Anforderungen eines validen psychometrischen Tests genügen.

2.6.2. Probleme mit dem IMI

In einem ersten Versuch konstruierte Korner (2015) aus dem englischsprachigen IMI von Deci und Ryan (2003) ein Messinstrument bestehend aus sechs Subskalen (interest and enjoyment, perceived choice, perceived competence, effort and importance, value and usefulness, (sozial) relatedness) mit insgesamt 40 Items. Die Skala pressure and tension war für den Kontext ihrer Arbeit nicht von Relevanz und wurde deshalb ausgeschlossen. Als Ergebnisraum wählte die Autorin, anstelle einer siebenteiligen Likertskala, eine Fünfteilige, da sonst Schwierigkeiten hinsichtlich der feinen Differenzierung von Seite der Schülerinnen und Schüler zu erwarten wären. Weiters wurden einige Items an das Lese- und Sprachverständnis der Jugendlichen der Sekundarstufe 1 adaptiert, um Verständnisschwierigkeiten zu vermeiden.

Im Rahmen der ersten Pilotierung des Fragebogens mit 238 Schülerinnen und Schülern aus der Sek 1, wollte Korner mittels konfirmatorischer Faktoranalyse die Faktorstruktur reproduzieren. Dies gelang nicht, wofür die Autorin folgende mögliche Gründe ausfindig machte:

- Verständnisprobleme: Obwohl die Items sorgfältig mit Unterstützung einer Anglistin und Validierung durch zwei Lehrkräfte übersetzt wurden, kam es zu Verständnisproblemen und Nachfragen seitens der Lernenden.
- Probleme mit der Formulierung: Zum einen konnten Formulierungs-Abstufungen von Items einer Skala nicht ausreichend differenziert werden. Zum anderen finden sich Items mit widersprüchlichen Aussagen, die im Rahmen einer Analyse hoch korrelieren sollten. Korner stellte fest, dass dem nicht so ist, was vermuten lässt, dass die Befragten hier keinen Widerspruch empfanden.
- Probleme der Itempolarität: Um der Zustimmungstendenz entgegenzuwirken, verwendete Korner negativ gepolte Items. Nach Umpolung korrelierten diese Items niedrig mit den übrigen Items der jeweiligen Skala. Eine Ursache dafür könnte im nächsten Punkt liegen.
- Doppelte Verneinung: Einige Items weisen eine doppelte Verneinung auf, die die Schülerinnen und Schüler nicht gemäß ihrer Zustimmung oder Ablehnung sinnvoll beantworten konnten.

(vgl. Korner, 2015)

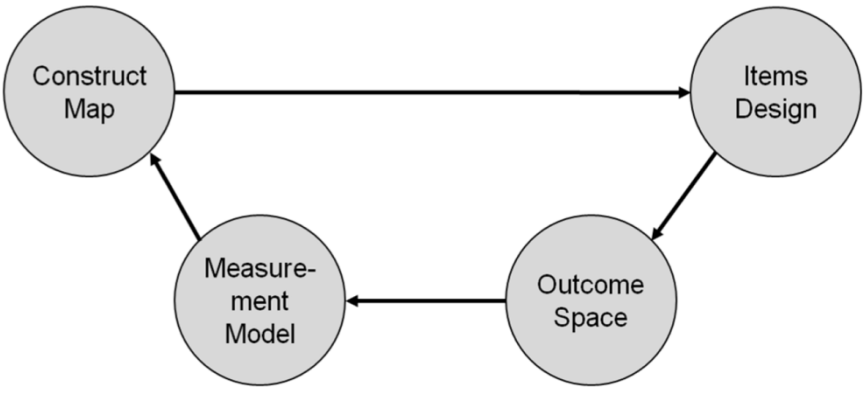
In weiterer Folge überarbeitete Korner den Fragebogen auf Grundlage inhaltlicher, sprachlicher und mathematischer Gesichtspunkte. Der daraus entstandene, deutlich verkürzte Fragebogen umfasste 23 Items. Das größte Problem bestand darin, dass die hinter den Skalen liegenden Konstrukte durch die geringe Itemanzahl nicht mehr in ihrer Breite abgebildet waren. Dieses Problem führte im Rahmen der Faktoranalyse, im Anschluss an die zweite Pilotierung des überarbeiteten Bogens dazu, dass die Faktorstruktur nicht reproduzierbar war. Zudem wies das Instrument eine Abhängigkeit von der Stichprobe auf, weswegen die Ergebnisse als nicht reliabel zu werten sind.

2.6.3. Weiterentwicklung bestehender Skalen für das Messinstrument

Da Korner mit der gewählten Vorgangsweise widersprüchliche Ergebnisse erhielt, entschied sie sich für eine Neukonstruktion der Items des IMIs. Zu Abbildung der einzelnen Konstrukte wählte Korner den *Four Building Blocks Approach* nach Wilson (2005). Diese Methode soll gewährleisten, dass die zur jeweiligen Subskala gehörenden Items das zu messende Konstrukt hinreichend genau erfassen. Somit ist Inhaltsvalidität gegeben. Damit soll der Schluss von der Beobachtung konkreter Evidenzen auf die latente Variable, hier die intrinsische Motivation, mit einem möglichst kleinen Fehler und in hoher psychometrischer Qualität erfolgen.

Der Four Building Blocks Approach

Der *Four Building Blocks Approach* nach Wilson (2005) umfasst die in der Grafik (1. Abbildung) sichtbaren vier Schritte, die zur Abbildung eins Konstrukts durchzuführen sind. Idealerweise werden sie mehrmals durchlaufen. Überlegungen zu den einzelnen Schritten anhand der Skala *perceived competence* sollen den Approach im Folgenden greifbarer machen.



1. Abbildung: Four building blocks (Wilson, 2005, S. 31)

Construct Map

Im Rahmen der *Construct Map* soll ein präzises Konzept eines Konstrukts (also eines Merkmals), etwa der *perceived competence* entwickelt und visualisiert werden. Damit wird eine inhaltliche Definition des Konstrukts vorgenommen. Zunächst werden die extremen Ausprägungen des Merkmals erfasst, sowie anschließend der Übergang von einem Extremum ins andere skizziert. Um den Standpunkt einer Person auf diesem Kontinuum zu erfassen, müssen den einzelnen Ausprägungen konkrete Evidenzen zugeordnet werden. Diese Reduktion auf Beobachtbares ist essenziell und einer der Kernpunkte beim *Four Building Blocks Approach*. Um konkrete Evidenzen zu finden, kann man das Konstrukt auf eine Kernfrage abbilden. Für das Konstrukt *perceived competence* könnte diese Frage lauten: *Was zeichnet Schülerinnen und Schüler aus, die sehr kompetent sind, also alle Aufgabenstellungen sehr gut erledigen können?* Die so entwickelten beobachtbaren Evidenzen sind dann der Ausgangspunkt für die einzelnen Items des Fragebogens, die Aufschluss über die Merkmalsausprägung einer Person zu dem beschriebenen Konstrukt geben. Im Allgemeinen muss eine gewisse Anzahl an Evidenzen vorhanden sein, damit ein Konstrukt nicht unterrepräsentiert und Inhaltsvalidität gewährleistet ist.

Item Design

Aus der *Construct Map* mit ihren konkreten Evidenzen sollen nun Items formuliert werden. Dabei versucht man, das theoretische Konstrukt in reale Begebenheiten zu transformieren. Die Items geben Auskunft über die Ausprägung des Konstrukts, weshalb sie als dessen Verwirklichung gesehen werden können. Ein Item ist in geschlossenen Formaten meist als Aussage formuliert, während man bei offenen Formaten meist Fragen wählt. Im Rahmen dieser Arbeit wurden die Items im geschlossenen Format verfasst.

Bei der Erstellung einer *Construct Map* sowie dem Item Design ist das Einbeziehen von Expertinnen und Experten hilfreich. Das können in diesem Fall entweder Fachdidaktikerinnen und Fachdidaktiker sein, oder aber auch Lernende der Sek 1, die die Zielgruppe des Fragebogens zur intrinsischen Motivation bilden.

Outcome Space

Bei diesem Schritt gilt es die Frage zu klären, welche Aspekte der Item- Antworten als Basis für einen Rückschluss auf ein bestimmtes Konstrukt verwendet werden können. Dazu werden im Rahmen von offenen Fragebögen Kategorien erstellt. Diese braucht man bei einem geschlossenen Format wie bei einer Likert Skala nicht. Im Rahmen dieser Arbeit kommt eine 5-teilige Likert Skala zum Einsatz. Dies scheint angesichts der Zielgruppe von Unterstufenschülerinnen und Schülern am geeignetsten zu sein. Der Antwort „1“ wird dann beispielsweise keine Zustimmung beziehungsweise hohe Ablehnung zugeschrieben und der Antwort „5“ hohe Zustimmung.

Measurement Model

Im letzten Schritt des *Four Building Blocks Approach* soll eine Verbindung zwischen den Antworten der befragten Person und dem Konstrukt hergestellt werden. Für den Kontext dieser Arbeit bietet es sich an, von den Antworten zu den einzelnen Items eines Konstrukts den Mittelwert zu bilden und darüber hinaus einen “Gesamtscore“ zu einer Merkmalsausprägung zu erhalten. Dabei müssen negativ gepolte Items vor der Analyse umgepolt werden. Ein hoher Gesamtscore lässt dann auf eine große Merkmalsausprägung schließen. Man könnte aber auch einzelne Items analysieren und vergleichen oder nur über bestimmte Items den Mittelwert bilden. Dies hängt von der Forschungsfrage beziehungsweise dem Ziel der Befragung ab.

Im Rahmen einer Pilotstudie sollen alle Items nach Durchlauf des *Four Building Blocks Approach* getestet, evaluiert und überarbeitet werden. Diese Arbeit ist Teil dieses zyklischen Prozesses, der die Qualität des Messinstruments gewährleisten soll.

Formulierung der Items: situiert versus verallgemeinert

Im Rahmen ihrer Diplomarbeit wirft Pieler (2018) die Frage auf, ob eine verallgemeinerte Formulierung der Items des Fragebogens sinnvoller ist, als wenn sie situiert angegeben sind. Im Kontext dieser Arbeit gilt ein Item als situiert, wenn es sich auf eine zuvor durchgeführte Tätigkeit bezieht.

Im Rahmen der Betrachtung der Motivation von Lernenden wenden Dresel und Lämmle (2017) ein, dass diese auch von Merkmalen der spezifischen Situation abhängig ist. Will man also Motivation messen, so muss man sich auf konkrete Ereignisse beziehen, da Individuen in unterschiedlichen Situationen für

unterschiedliche Dinge motiviert sind. Diese Argumentation bekräftigen Nolen et. al. (2015), indem sie postulieren, dass sich Menschen je nach sozialen Positionen und Praktiken unterschiedlich in verschiedenen Kontexten engagieren. Eine situierte Perspektive erlaubt es im Rahmen einer Analyse, nicht nur Aussagen über Lernende und ihre Motive, sondern auch über Lernende in Kontexten zu machen, was das Ziel des hier entwickelten Fragebogens darstellt. Dies bekräftigt die gewählte Vorgangsweise, grundsätzlich situiert zu bleiben. Lediglich die Items der Skala *relatedeness* sind situationsunabhängig und deshalb allgemein formuliert.

Nach Busker (2014) sollen Items Sachverhalte konkret und direkt ansprechen. Da der Fragebogen jedoch in unterschiedlichen Kontexten und unabhängig vom Unterrichtsfach eingesetzt werden soll, ist der Bezug auf eine spezifische Situation so nicht möglich. Aus diesem Grund steht einerseits bereits im Einleitungstext, dass sich die Fragen des Fragebogens auf die eben gemachte Partnerarbeit beziehen, und andererseits wurden Formulierungen wie „(...) diese Aufgabe (...)“, „(...) das Thema (...)“, „(...) die Arbeitsaufträge (...)“ oder „(...) die Aktivität (...)“ gewählt, die sich direkt auf die vorangehende Unterrichtssituation beziehen. Damit wird jedoch ein gewisses Abstraktionsvermögen von den Schülerinnen und Schülern vorausgesetzt, da diese nicht konkret benannt wird.

3. Forschungsdesign

Nachdem im vorangegangenen Kapitel die Notwendigkeit einer Neukonstruktion der Skalen des IMIs verdeutlicht wurde, liegt der Fokus dieses Kapitels auf der Vorstellung des Forschungsdesigns.

3.1. Forschungsfragen

Obwohl die bisherigen Arbeiten zu dem neu entwickelten Testinstrument zur Ermittlung der intrinsischen Motivation gute Ergebnisse brachten, zeigten sich noch offene Fragestellungen, die im Rahmen dieser Masterarbeit bearbeitet wurden.

So wiesen die beiden Skalen *perceived competence* und *pression and tension* eine hohe Korrelation auf, deren Ursache näher betrachtet wurde.

> Erster Forschungsschwerpunkt der Arbeit:
> Ermöglicht eine Überarbeitung der Konstrukte nach dem *Four Building Blocks Approach* (siehe Kapitel 2.6.3.) eine faktoranalytische Abgrenzung der beiden Skalen bezüglich verschiedener Stichproben? Können trennscharfe Skalen zu den Teilkonstrukten *perceived competence* und *pressure tension* entwickelt werden? Geht man davon aus, dass die wahrgenommene Selbstkompetenz mit Anspannung und Druck negativ korreliert, so stellt sich die Frage, ob eine der Skalen ohne weitere Bedenken weggelassen werden kann.

Im Rahmen der Bearbeitung des ersten Forschungsschwerpunkts wurden die bestehenden Konstrukte samt ihrer Items mit Expertinnen und Experten validiert und überarbeitet sowie hinsichtlich ihrer Trennschärfe analysiert. Um die erste Forschungsfrage beantworten zu können, wurden bestehende und neuentwickelte Items zu den Skalen *perceived competence* und *pressure tension* getestet und anschließend einer Korrelationsanalyse und einer konfirmatorischen Faktoranalyse unterzogen. Ob die Skalen sich durch einen Faktor abbilden lassen, wurde mittels Kaiser-Kriterium und Scree-Plot untersucht. Weiters wurden die beiden Skalen *perceived competence* und *pressure tension* bezüglich verschiedener Stichproben näher analysiert.

Da sich die erste Forschungsfrage klären ließ, konnte ich im Anschluss der Frage nachgehen, ob sich bei gleichzeitigem Messen aller Subskalen die Faktorenstruktur reproduzieren lässt.

> Zweiter Forschungsschwerpunkt:
> Sind die Skalen trennscharf und stabil, wenn alle Items gleichzeitig miteinander wechselwirken?

Dazu wurden auch die übrigen Skalen überarbeitet. Ein Hauptaugenmerk dabei lag auf einer deutlichen Reduktion der Itemanzahl. Obwohl viele Items pro Skala in der Regel eine hohe Reliabilität gewährleisten, das Konstrukt in seiner gesamten Breite abbilden und gute faktoranalytische Ergebnisse liefern, besteht die Tendenz, Aspekte mehrfach abzufragen. Außerdem ist es schwierig, bei zu vielen Items die Konstrukte unidimensional zu halten, also mit einem bestimmten Item unwillentlich Aspekte anderer Skalen mit abzufragen, die nicht zum Konstrukt passen. Für die Beantwortung der Forschungsfragen wurden Schülerinnen- und Schülerbefragungen, Expertinnen- und Experteninterviews durchgeführt sowie eine Expertenrunde gegründet.

Ziel dieser Arbeit ist es, einen weiteren Schritt zur Finalisierung eines validen, vor allem aber ökonomischen Messinstruments, welches sowohl klassischen Gütekriterien genügt als auch in der Praxis gut einsetzbar ist, beizutragen. Damit soll es möglich sein, in praktikabler Weise, Studien über die Motivation von Lernenden im Unterricht durchzuführen.

Die im Rahmen dieser Arbeit untersuchten Forschungsfragen ergaben sich aus den Arbeiten von Korner (2015), Klingenböck (2016), Schmidt (2017) und Pieler (2018), die bereits bedeutende Beiträge zur Erstellung dieses ersten deutschsprachigen Testinstruments zur Messung der Motivation von Lernenden beigetragen haben.

3.2. Erstellung der Construct Maps und der Items

Zunächst galt es, im Rahmen des ersten Forschungsschwerpunkts, die Skalen ***perceived competence*** und ***pressure and tension*** näher zu analysieren. Da angenommen werden kann, dass Lernende mit einer hohen wahrgenommenen Selbstkompetenz, sich in Lernsituationen weniger unter Druck gesetzt fühlen als diejenigen, die ihre Selbstkompetenz schwächer einschätzen, verwundert es nicht, dass die beiden Skalen hoch (negativ) miteinander korrelieren, wie es sich aus den Daten von Schmid (2017) ergab.

Diese Vermutung wirft die berechtigte Frage auf, ob man zu Gunsten eines kompakteren (und damit zumutbareren) Fragebogens eine der beiden Skalen weglassen darf. Schließlich erlaubt die Erkenntnis über die Antworten auf eine der Skalen einen Rückschluss auf die andere, nicht abgefragte Skala.

Andererseits könnte ein Fehler im Konstrukt selbst liegen, was eine Betrachtung der *Construct Maps*, sowie der daraus konstruierten Items nahelegt. Um die bestehenden Konstrukte inhaltlich genauer analysieren und sie dabei in ihrer Breite erfassen zu können, wurde in einer Runde bestehend aus Expertinnen und Experten ein sogenanntes *Itempanel* nach Wilson (2005) durchgeführt. Dabei erstellten Fachdidaktikerinnen und Fachdidaktiker des AECC Physik, im Rahmen eines Seminars, selbst zwei neue *Construct Maps* mit dazugehörigen Items nach der in Kapitel 2.6.3. vorgestellten Methode des *Four Building Blocks Approach* (2005). Anschließend konnten diese mit den bereits bestehenden verglichen werden, wodurch eine Expertenvalidierung gegeben war.

Bei der Erstellung der *Construct Maps*, sowie der Items musste die Zielgruppe des Fragebogens beachtet werden. Da das Erhebungsinstrument für die Sekundarstufe 1 (2. bis 4. Klasse) konzipiert ist, sollten Lese- und Sprachkompetenz an 11- bis 14-jährige Jugendliche adaptiert werden.

3.2.1. Construct Maps und Items zur Skala perceived competence

Zunächst wurde die Skala ***perceived competence***, welche Schmidt (2017) im Rahmen seiner Diplomarbeit neu konstruiert hat, überarbeitet. In Tab. 1 findet sich die von ihm erstellte Construct Map, und in Tab. 2 jene, die im Rahmen des *Itempanels* mit den Expertinnen und Experten neu entstand.

SchülerInnen sind bei ihrer Tätigkeit sehr kompetent.	
SchülerInnen	**Antworten zu den Items**
SchülerInnen konnten alle Aufgabenstellungen sehr gut erledigen	SchülerInnen zeigen auf und gestalten den Unterricht mit.
SchülerInnen konnten alle Aufgabenstellungen gut erledigen.	SchülerInnen halten Augenkontakt zur Lehrperson.
SchülerInnen konnten einige Aufgabenstellungen erledigen.	SchülerInnen können Fragen beantworten.
SchülerInnen konnten kaum Aufgabenstellungen erledigen.	SchülerInnen lassen sich schnell ablenken.
SchülerInnen konnten keine Aufgabenstellungen erledigen.	SchülerInnen geben unsichere Antworten und geben die Aufgaben weiter.
SchülerInnen sind mit ihrer Tätigkeit überfordert.	

Tab. 1. Construct Map zur Skala Perceived Competence (vgl. Schmidt, 2017, S. 59)

SchülerInnen sind bei ihrer Tätigkeit sehr kompetent.	
SchülerInnen	**Antworten zu den Items**
SchülerInnen sind überzeugt, immer eine souveräne Antwort geben zu können.	SchülerInnen melden sich, gestalten den Unterricht mit und sind von den eigenen Antworten überzeugt
SchülerInnen sind überzeugt, meistens eine souveräne Antwort geben zu können.	...
SchülerInnen denken, auch gute Antworten geben zu können.	Anpassungen über die Menge (meistens, manchmal, selten)
SchülerInnen bezweifeln, angemessene Antworten geben zu können.	...
SchülerInnen bezweifeln ob das was sie denken, richtig sein könnte.	SchülerInnen geben unsichere, zögerliche Antworten und warten auf Wortmeldungen der Anderen.
SchülerInnen sind mit ihrer Tätigkeit überfordert.	

Tab. 2. Construct Map zur Skala Perceived Competence aus dem Itempanel

Um in weiterer Folge zum Konstrukt passende Evidenzen zu finden, befragte Schmidt 117 Schülerinnen und Schüler. Dabei konnte er die Mitarbeit sowie das Übernehmen von Aufgaben als Indikatoren für die Skala *perceived competence* ausfindig machen. Darüber hinaus ergab die Befragung, dass kompetente Schülerinnen und Schüler in ihren Handlungen souverän wirken und auf die jeweilige Aufgabenstellung fokussiert sind. Aus diesen Erkenntnissen, sowie einer Expertenvalidierung, formulierte Schmidt die untenstehenden 11 Items (pco 1 bis pco 11).

Ebenso konstruierten die Fachdidaktikerinnen und Fachdidaktiker, im Rahmen des Seminars, Items zur Skala *perceived competence*. Im Anschluss konnte über die beiden nun vorliegenden Realisierungen des Konstrukts debattiert werden. Auf Basis dieser Diskussion, eines inhaltlichen Vergleichs mit den Items der anderen 5 Subskalen des Fragebogens, sowie unter Berücksichtigung statistisch mathematischer Gesichtspunkte, wurden die Items von Schmidt entweder beibehalten, neu formuliert oder ausgeschlossen. Die verblieben sechs Items sind in der folgenden Auflistung in **fetter Schriftstärke** dargestellt. Negativ gepolte Items wurden mit einem "(R)" gekennzeichnet. Die Begründungen für die getroffene Auswahl finden sich im Anschluss.

pco 1: Ich zeige im Unterricht oft auf.
pco 2: Ich trage etwas zum Unterricht bei.
pco 3: **Ich arbeite im Unterricht viel mit.**
pco 4: Ich halte Augenkontakt zum/zur Lehrer/Lehrerin.
pco 5: **Ich beantworte Aufgabenstellungen/Fragen schnell.**
pco 6: Bei Gruppenarbeiten übernehme ich gerne eine Aufgabe.
pco 7: **Ich weiß, dass ich Aufgaben aus dem Unterricht lösen kann.**
pco 8: Ich lasse mich im Unterricht schnell ablenken. (R)
pco 9: **Wenn ich angesprochen werde, sind meine Antworten unsicher.** (R)
pco 10: **Wenn ich eine Aufgabe bekomme, schaue ich zuerst was die anderen machen.** (R)
pco 11: **Bei Gruppenarbeiten überlasse ich die Aufgaben den anderen.** (R)

(vgl. Schmidt 2017)

Dabei galt es die Items pco 2, pco 4, pco 6 und pco 8 auf Grund ihrer schlechten Faktorladung inhaltlich zu überprüfen und gegebenenfalls auszuschließen.

Die Items pco 1: „Ich zeige im Unterricht oft auf.“ und pco 3: „Ich arbeite im Unterricht viel mit.“ sind inhaltlich sehr ähnlich, weshalb nur eines der beiden beibehalten wurde.

Bezüglich pco 4 ist zu hinterfragen, ob das Halten des Augenkontakts mit der Lehrperson ein guter Prädiktor für die wahrgenommene Selbstkompetenz ist, weswegen auch dieses Item gestrichen wurde.

Und zuletzt wurde pco 6 ausgeschlossen, da bei Gruppenarbeiten das Übernehmen von Aufgaben stark von der Partnerin / vom Partner abhängig ist.

3.2.2. Construct Maps und Items zur Skala pressure tension

Parallel zur Arbeit von Schmidt entwickelte Korner (2017) die Skala ***pressure and tension*** neu, die im Rahmen dieser Arbeit auf gleiche Weise wie die Skala *perceived competence* überarbeitet wurde. In Tab. 3 ist die *Construct Map* von Korner zu sehen und in Tab. 4 jene, die im Rahmen des *Itempanels* mit den Expertinnen und Experten neu entstand.

SchülerInnen fühlen sich stark unter Druck

SchülerInnen	**Antworten auf die Items**
SchülerInnen haben Angst vor Versagen	lehnen die Aufgabe ab, da sie sich ihr nicht gewachsen fühlen
SchülerInnen verhalten sich der Aufgabe gegenüber neutral	fühlen sich nervös und angespannt, wenn sie die Aufgaben bekommen
	fühlen sich etwas unwohl, wenn sie die Aufgabe bekommen
SchülerInnen können entspannt an die Bearbeitung der Aufgaben gehen	fühlen keinen Druck oder Anspannung, wenn sie die Aufgabe bearbeiten sollen

SchülerInnen sind entspannt

Tab. 3. Construct Map zur Skala Pressure and Tension (vgl. Korner, 2017)

SchülerInnen fühlen sich stark unter Druck

SchülerInnen	**Antworten auf die Items**
SchülerInnen empfinden aufgrund der Aufgabenstellung sehr viel Anspannung / Druck	ziehen sich vom Unterrichtsgeschehen gänzlich zurück
	...
SchülerInnen empfinden aufgrund der Aufgabenstellung weitestgehend Anspannung / Druck	beteiligen sich ein wenig am Unterrichtsgeschehen
	...
SchülerInnen empfinden aufgrund der Aufgabenstellung ein wenig Anspannung / Druck	sind entspannt und beteiligen sich aktiv am Unterrichtsgeschehen.
SchülerInnen empfinden aufgrund der Aufgabenstellung kein/e Anspannung / Druck	

SchülerInnen sind entspannt

Tab. 4. Construct Map zur Skala Pressure and Tension aus dem Itempanel

Auf Basis der *Construct Map* formulierte Korner die untenstehenden 12 Items (pres 1 bis pres 12). Ebenso erstellten die Fachdidaktikerinnen und Fachdidaktiker, im Rahmen des Seminars, Items zur Skala *pressure and tension*. Auf Grundlage der anschließenden Diskussion, eines inhaltlichen Vergleichs mit den Items der anderen 5 Subskalen des Fragebogens, sowie unter Berücksichtigung statistisch mathematischer Gesichtspunkte, wurden nur drei Items von Korner **beibehalten** sowie drei Items **neu konstruiert**. Sie sind in der folgenden Auflistung in **fetter Schriftstärke** dargestellt. Negativ gepolte Items wurden mit einem "(R)" gekennzeichnet. Die Begründungen für die getroffene Auswahl finden sich im Anschluss.

pres 1: Ich habe Angst, dass ich die Aufgabe nicht beantworten kann.

pres 2: Ich fange zu zittern an, weil ich diese Aufgabe bearbeiten soll.

pres 3: **Ich rede leise, weil ich Angst habe, etwas Falsches zu sagen.**

pres 4: **Ich schaue in der Klasse herum, ob mir jemand bei der Aufgabe helfen kann.**

pres 5: Ich hoffe, dass jemand anderer die Aufgabe übernimmt.

pres 6: Ich würde am liebsten weglaufen, wenn ich so eine Aufgabe bekomme.

pres 7: Ich fange zu schwitzen an, weil ich diese Aufgabe bearbeiten soll.

pres 8: Ich fange zu weinen an, weil ich diese Aufgabe bearbeiten soll.

pres 9: **Ich hoffe, es sieht mich niemand und ich muss die Aufgabe nicht behandeln.**

pres 10: Ich werde rot im Gesicht, weil ich Angst habe, die Aufgabe nicht lösen zu können.

pres 11: Ich beginne zu stottern, wenn ich die Antwort auf eine Aufgabe nicht weiß.

pres 12: Ich mache oft Schlampigkeitsfehler, weil ich Angst vor dem Versagen habe.

(vgl. Korner 2017)

pres 13: **Ich spüre, dass ich mich bei dieser Aufgabe verkrampfe.**

pres 14: **Ich habe Angst zu versagen.**

pres 15: **Auch wenn ich gut vorbereitet bin, habe ich Angst, dass ich die Aufgabe nicht schaffen kann.**

Das Item pres 1 wurde im Rahmen einer Neuformulierung insofern präzisiert, als dass das neue Item pres 15 nun die mangelnde Vorbereitung als Grund für die Angst ausschließt. Dabei wurde auf ein ähnliches Item aus der Pisa Befragung zurückgegriffen (vgl. PISA, 2019, S. 51).

Die Items pres 2: „Ich fange zu zittern an, weil ich diese Aufgabe bearbeiten soll." und pres 6: „Ich würde am liebsten weglaufen, wenn ich so eine Aufgabe bekomme." wurden ausgeschlossen, da sie von den Diskussionsteilnehmerinnen und Teilnehmern als übertrieben formuliert gesehen wurden.

Ebenso hinterfragt wurde, inwiefern Schülerinnen und Schüler von sich aus preisgeben würden, beim Bearbeiten von Aufgaben aus Angst zu schwitzen, zu weinen oder zu stottern. Aus diesem Grund wurde anstelle der Items pres 7, pres 8 und pres 11 das neue Item pres 13: „Ich spüre, dass ich mich bei dieser Aufgabe verkrampfe." formuliert. Neben der Vermutung, dass dieser Zustand typischer für Angstzustände ist, wird angenommen, dass Jugendliche solch einem Item eher zustimmen, als zu schwitzen, zu weinen oder zu stottern.

Das Item pres 5: „Ich hoffe, dass jemand anderer die Aufgabe übernimmt." wurde wegen seiner inhaltlichen Ähnlichkeit zu Item pco 11: „Bei Gruppenarbeiten überlasse ich die Aufgaben den anderen." gestrichen.

Da manche Personen eher die Tendenz haben, rot im Gesicht anzulaufen als andere, und die wenigstens dies von sich selbst behaupten würden, wurde das Items pres 10 ausgeschlossen.

Zuletzt wurde das Item pres 12: „Ich mache oft Schlampigkeitsfehler, weil ich Angst vor dem Versagen habe." gekürzt, da es zwei verschiedene Aspekte beschreibt. Die neue Formulierung lautet: „Ich habe Angst zu versagen." (pres 14)

Nach der eben beschriebenen Überarbeitung der Skalen *perceived competence* und *pressure tension*, wurden diese jeweils durch sechs Items repräsentiert. Die Durchführung des *Itempanels* mit den Expertinnen und Experten gewährleistet, dass die Konstrukte in ihrer Breite erfasst sind, womit Inhaltsvalidität gegeben ist. Durch die strake Reduktion der Itemanzahl soll vermieden werden, Dinge doppelt abzufragen. Die fett gedruckten zwölf Items sind in dem von mir neu erstellten und

anschließend evaluierten Fragebogen (siehe Anhang 10.1.1.) zu finden. Die Ergebnisse, Analysen und Diskussion folgen in den Kapiteln 4 und 5.

3.3. Erstellung des Fragebogens

Im Anschluss an die Überarbeitung der ersten beiden Konstrukte, im Rahmen des ersten Forschungsschwerpunkts, sollten nun auch die übrigen fünf Subskalen analysiert und bearbeitet werden. Im Sinne der Überlegungen aus der Expertenvalidierung mit den Fachdidaktikerinnen und Fachdidaktikern galt es zu überprüfen, ob die Konstrukte ausreichend in ihrer Breite abgebildet sind. Dabei wurden bestehende Items adaptiert beziehungsweise zu fehlenden Aspekten neue Items konstruiert. Außerdem wurde darauf geachtet, dass die Skalen trennscharf sind, es also keine inhaltlichen Überschneidungen von Items unterschiedlicher Skalen gibt. Weiters galt es, alle Items hinsichtlich der Lese- und Sprachkompetenz von Jugendlichen der Sek 1 zu überprüfen.

Das daraus entstandene Set an Items wurde abschließend mit Frau Dr. Korner Expertinnen validiert. Im Anhang (10.3.) findet sich zu jeder Skala ein Übersichtsblatt, auf welchem die ursprünglich bestehenden, die selektierten, sowie die neu konstruierten Items aufgelistet, sowie die angestellten Überlegungen notiert sind.

Damit konnte erstmals ein Fragebogen erstellt werden, der gleichzeitig Items zu allen sieben Subskalen umfasst, in für Schülerinnen und Schüler zumutbarer Zeit zu bearbeiten ist und demnach Auskunft über die intrinsische Motivation gibt. Anhang 10.4. zeigt die Zuordnung der 41 Items zur jeweiligen Skala, wobei diese im Rahmen des Fragebogens durchmischt wurden. Die jeweilige Zuordnung der einzelnen Aussagen (also der Items) des Fragebogens zu einer der Skalen, findet sich im Infoblatt des Anhangs (10.2.1.).

3.4. Durchführung an den Schulen

Der erstellte Fragebogen wurde nun im Anschluss an eine Intervention an verschiedenen Schulen in Wien und Niederösterreich getestet. Die Lehrpersonen wurden gebeten, eine Unterrichtsmethode einzusetzen, bei welcher die Lernenden untereinander interagieren, damit Items zu *relatedeness* sinnvoll beantwortet werden konnten. Das Thema, zu welchem die Schülerinnen und Schüler zusammenarbeiten, wurde nicht vorgegeben. Jedoch kamen nur Schulstunden in Frage, in denen naturwissenschaftliche Fächer wie Physik, Biologie oder Chemie unterrichtet werden. Das bedeutet, dass zum ersten Mal auch Stichproben bezüglich anderer Fächer als Physik hinzugenommen wurden, da das IMI prinzipiell auch für unterschiedlichste Kontexte zu verwenden ist. Für die Erhebung waren grundsätzlich alle Schülerinnen und Schüler der 6. bis 8. Schulstufe vorgesehen. Um die Stichprobe zu vergrößern, wurde der Bogen ebenso in einer zehnten Schulstufe getestet.

Die Testung erfolgte an 8 Schulen mit insgesamt 255 Schülerinnen und Schülern. Da das qualitative Ergebnis der Auswertung bei drei Schulen auffällig schlecht war, wurden diese vor der (Faktoren-) Analyse ausgeschlossen. Viele Bögen der Befragten aus den betroffenen Schulen waren entweder nur sporadisch ausgefüllt oder wiesen eine systematische Beantwortung der Fragen auf. Dies könnte mit dem hohen Anteil an Jugendlichen mit nicht-deutscher Muttersprache in Zusammenhang stehen, der jeweils über 80 Prozent betrug oder aber auch mit einer unzureichenden Administrierung der Fragebögen durch die betroffenen Lehrkräfte. Der Ausschluss betraf Jugendliche der 6. Schulstufe, die die Gruppe jüngsten Teilnehmerinnen und Teilnehmer der Befragung bilden.

Übrig blieben 158 ausgefüllte Fragebögen aus 5 Schulen, die jeweils mit einer Nummer versehen wurden. Beispielsweise wurde der erste Fragebogen mit der Bezeichnung „FB#1“ beschriftet, bei den weiteren wurde auf eine fortlaufende Nummerierung geachtet. Die einzelnen Bögen verteilen sich auf die Schulen wie in Tabelle 5 zu sehen ist, wobei wegen der anonymen Durchführung anstelle des Schulnamens die Buchstaben A, B, C, D und E gewählt wurden. Unter den Schulen befinden sich zwei Mittelschulen, ein Gymnasium, eine HTL und eine High-School, die keinem klassischen Schultyp zugeordnet werden kann.

Schule	Anzahl der Befragten	Zugehörige Fragebögen	Schulstufen
Schule A	35	FB# 1-35	6. und 8.
Schule B	42	FB# 36-77	7.
Schule C	30	FB# 95-124	6. und 7.
Schule D	25	FB# 125-149	6. und 10.
Schule E	26	FB# 150-176	7. und 8.

Tab. 5. Übersicht der Verteilung auf die unterschiedlichen Schulen

Die Beteiligung seitens der Schülerinnen und Schüler war **freiwillig** und **anonym**. Zum Zweck der Auswertung wurden lediglich die Schulform, die Schulstufe, das Unterrichtsfach in dem getestet wurde und die Muttersprache erhoben.

Wie bereits weiter oben erwähnt, war die dem Fragebogen vorangehende Intervention den Lehrpersonen überlassen. Jedoch wurden sie gebeten, den folgenden Einleitungstext den Schülerinnen und Schülern weiterzugeben:

> *Mit dem Beantworten dieser Fragen hilfst du einem Forscher an der Universität Wien, einen Fragebogen zum Thema **Motivation** zu entwickeln. Versuche die einzelnen Fragen bitte möglichst genau zu beantworten. Bitte kreuze bei JEDER Frage EINE Möglichkeit an.*
> *Die Fragen beziehen sich auf die eben gemachte Partner- oder Gruppenarbeit.*
> *Alle Antworten werden anonym behandelt und nicht benotet.*

3.5. Auswertung der Fragebögen

Die Auswertung der Fragebögen wurde mittels Statistik- und Analyseprogramm SPSS durchgeführt. In einem ersten Schritt mussten dafür die Antworten der Befragten kodiert werden.

Da das Antwortspektrum der Bögen eine fünfteilige Likertskala darstellt, wurden dafür die Werte 1 bis 5 verwendet. Stimmt die befragte Person einem Item völlig zu, so wurde die Zahl 5, bei gar keiner Zustimmung die Zahl 1 verwendet. Das bedeutet, dass eine hohe Merkmalsausprägung mit einer großen Zahl kodiert wurde. Darüber hinaus bekam „deutsche Muttersprache“ die Zahl 1, „nicht – deutsche Muttersprache“ die Zahl 0 zugewiesen. Ging dem Fragebogen eine Intervention im Fach Physik voraus, so wurde die Zahl 1, bei Biologie die Zahl 2 und bei Chemie die Zahl 3 verwendet. Blieb ein Item unausgefüllt oder war unklar beantwortet, so wurde die Zahl 99 verwendet, hingegen wurde bei Mehrfachantworten falls sinnvoll der Mittelwert gebildet. Wenn also beispielsweise 2 und 3, oder genau zwischen 2 und 3 angekreuzt war, so wurde 2,5 eingetragen.

Um Faktorenanalysen mittels SPSS ausführen zu können, mussten in einem letzten Schritt negativ gepolte Items (R) umgepolt werden. Zur Auswertung konnten alle Fragebögen der fünf Schulen herangezogen werden. Die Größe der gesamten Stichprobe umfasst 128 Schülerinnen und Schüler.

4. Ergebnisse

In diesem Kapitel präsentiere ich die Ergebnisse meiner Studie. Dabei werden zunächst die beiden Skalen *perceived competence* und *pressure tension* sowie anschließend die weiteren Skalen des IMIs betrachtet. Im Zuge dessen werden die in Kapitel 2.5.3 und 2.5.4 vorgestellten Test- und Analysemethoden angewandt. Sie sollen neben inhaltlichen Argumenten Aufschluss über die Güte und Qualität der einzelnen Skalen sowie des Fragebogens in seiner Gesamtheit geben. Darüber hinaus werden unterschiedliche Stichproben, hinsichtlich ihrer intrinsischen Motivation, miteinander verglichen.

Damit eine Faktoranalyse sinnvoll interpretierbare Ergebnisse liefern kann, muss die Stichprobengröße, nach dem Richtwert von Freiburg, mindestens dreimal, idealerweise fünfmal so groß wie die Anzahl der Items sein (vgl. Leonhart, 2009). Der im Rahmen dieser Arbeit evaluierte Fragebogen umfasst 41 Fragen, woraus sich eine Mindestanzahl von 123 Befragten ergibt. Da nach dem Ausschluss der drei Schulen die Stichprobengröße noch immer 158 Schülerinnen und Schüler umschließt, konnte eine Faktoranalyse durchgeführt werden.

4.1. Analyse der Skalen des Fragebogens

Neben dem Richtwert von Freiburg gibt es noch zwei weitere Kriterien, welche im Vorfeld einer Faktoranalyse überprüft werden müssen. Das ist zum einen die Bestimmung des Kaiser-Meyer-Olkin-Koeffizienten (kurz KMO-Koeffizient) und zum anderen die Durchführung des Bartlett-Tests.

Um zu überprüfen, ob die vorhandenen Daten für eine Faktoranalyse geeignet sind, gibt der KMO-Koeffizient die Höhe der Korrelation in der Korrelationsmatrix an (vgl. Bühner, 2011). Für die hier analysierten Daten ergibt sich dabei ein Wert von 0,839, welcher als gut klassifiziert gelten kann.

Der Bartlett-Test ist ein Signifikanztest, der überprüft, ob die Daten zufällig sind. Diese Nullhypothese kann abgelehnt werden, da die Items mit $p < 0{,}001$ signifikant korrelieren (vgl. Bühner, 2011).

4.1.1. Korrelationstabellen

Die auf den nächsten Seiten folgenden Korrelationstabellen wurden im Team erstellt. Im Rahmen der Auswertung mittels SPSS wurde zunächst jede Skala einzeln betrachtet.

Mittels Hauptkomponentenanalyse konnte überprüft werden, wie stark die einzelnen Items auf die jeweilige Skala laden. Schwach korrelierende Items, die dementsprechend eine schlechte Faktorladung (<0,6) aufweisen, können vom statistisch-mathematischen Gesichtspunkt aus, ausgeschlossen werden. Dies gilt es jedoch stets aus inhaltlicher Sicht zu überprüfen. Da die Ladung der einzelnen Items abhängig von den anderen Items der Skala ist, verändert sich bei Ausschluss eines Items die gesamte Korrelationstabelle (siehe Kapitel 2.5.3).

Um die Höhe der Itemkorrelation zu überprüfen, wurde das Cronbach-Alpha verwendet, welches die Anzahl der Items berücksichtigt.

Darüber hinaus kann mittels erklärter Gesamtvarianz überprüft werden, durch wie viele Faktoren sich die Items abbilden lassen. Nach dem Kaiser-Gutmann-Kriterium (siehe Kapitel 2.5.3.) sind das jene Komponenten, deren Eigenwert größer 1 ist.

4.1.2. Faktoranalyse der Skala perceived competence

Die sechs zur Skala *perceived competence* gehörenden Items des Fragebogens wurden für eine Faktoranalyse herangezogen. Dabei zeigte sich, dass das Item pco 5: „Ich beantworte Aufgabenstellungen/Fragen schnell.“ nur mit einem Wert von 0,447 auf die Skala lädt. Schließt man es aus, so laden die übrigen Items mit Werten über 0,651 auf die Skala (siehe Tab. 6). Da das Item pco 5 aus inhaltlichen Gründen nicht ausgeschlossen werden soll, wurde es im Rahmen der Diskussion (siehe Kapitel 5) neu formuliert.

Komponentenmatrix[a]

	Komponente 1
pco_3	,791
pco_7	,754
pco_10_u	,740
pco_11_u	,672
pco_9_u	,651

Extraktionsmethode: Hauptkomponentenanalyse.

a. 1 Komponenten extrahiert

Tab. 6. Komponentenmatrix der Skala perceived competence

Mittels Kaiser-Gutmann-Kriterium erfolgte eine Überprüfung durch wie viele Faktoren sich die Items abbilden lassen. In Tabelle 7 ist zu sehen, dass nur eine Komponente einen Eigenwert größer 1 aufweist. Diese Komponente erklärt 52,37% der Gesamtvarianz der Items.

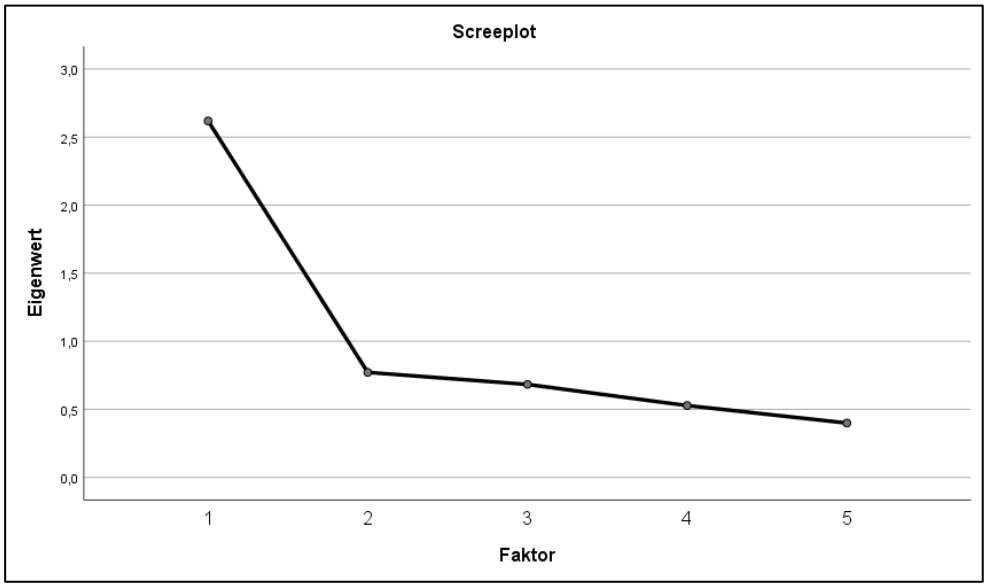
Erklärte Gesamtvarianz

	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
Komponente	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2,619	**52,372**	52,372	2,619	52,372	52,372
2	,771	15,429	67,801			
3	,683	13,660	81,461			
4	,528	10,554	92,015			
5	,399	7,985	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Tab. 7. Erklärte Gesamtvarianz der Skala perceived competence

Eine andere Methode, die Anzahl der sinnvoll zu extrahierenden Faktoren zu ermitteln, stellt der Scree-Test da, den ich in Kapitel 2.5.3 vorgestellt habe. Dabei betrachtet man das Eigenwertdiagramm, das für die Skala *perceived competence* in Abbildung 2 zu sehen ist. Man sieht deutlich, dass lediglich ein Faktor vor dem Knick im Funktionsgraphen liegt.



2. Abbildung: Screeplot zur Skala perceived competence

Darüber hinaus lässt sich aus der Item Skala-Statistik (siehe Tabelle 8) die Trennschärfe der einzelnen Items herauslesen. Sie gibt, wie im Kapitel 2.5.2. erwähnt, an, wie gut ein Item eine gewisse Eigenschaft misst. Da alle Items einen Wert von über 0,5, beziehungsweise nur knapp unter 0,5 aufweisen, kann man von einer hohen Trennschärfe sprechen.

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Quadrierte multiple Korrelation	Cronbachs Alpha, wenn Item weggelassen
pco_3	10,0260	10,748	**,614**	,427	**,700**
pco_7	9,9870	10,640	**,559**	,396	**,714**
pco_10_u	9,7273	9,533	**,566**	,327	**,713**
pco_11_u	10,2078	10,120	**,496**	,268	**,739**
pco_9_u	9,7403	11,543	**,469**	,224	**,744**

Tab. 8. Item-Skala-Statistik der Skala perceived competence

Das Cronbach-Alpha als Quantifizierung der Höhe der Itemkorrelation ergab für die 5 verbleibenden Items der Skala *perceived competence* einen Wert von 0,765. Als Maß für die Reliabilität kann ein solcher Wert als gut angesehen werden. Die rechte Spalte zeigt darüber hinaus, dass die Reliabilität der gesamten Skala schlechter wird, wenn man ein Item weglässt.

4.1.3. Faktoranalyse der Skala pressure and tension

Auf gleiche Weise folgt nun eine Betrachtung der Skala *pressure and tension*, wobei auch hier sechs Items für die Faktoranalyse herangezogen wurden.

Die Hauptkomponentenanalyse ergab, dass die Skala in zwei Komponenten zerfällt. Schließt man jedoch das Item pres 4: „Ich schaue mich in der Klasse um, ob mir jemand bei der Aufgabe helfen kann.“ aus, so sind die übrigen Items der Skala lediglich einer Komponente zuordenbar. Die Faktorladungen der einzelnen Items sind, wie in Tabelle 9 zu sehen ist, mit Werten über 0,620 ausreichend hoch.

Komponentenmatrix[a]

	Komponente 1
pres_3	,687
pres_9	,620
pres_13	,678
pres_14	,765
pres_15	,633

Extraktionsmethode: Hauptkomponentenanalyse.

a. 1 Komponenten extrahiert

Tab. 9. Komponentenmatrix der Skala pressure / tension

Nach dem Kaiser-Gutmann-Kriterium lassen sich die Items durch einen Faktor abbilden, wobei dieser 46,04% der Gesamtvarianz der Items erklärt (siehe Tab. 10). Auch das Eigenwertediagramm bestätigt dieses Ergebnis.

Erklärte Gesamtvarianz

	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
Komponente	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	2,302	**46,036**	46,036	2,302	46,036	46,036
2	,927	18,538	64,574			
3	,667	13,339	77,913			
4	,608	12,168	90,082			
5	,496	9,918	100,000			

Extraktionsmethode: Hauptkomponentenanalyse.

Tab. 10. Erklärte Gesamtvarianz der Skala pressure and tension

Weiters lässt sich aus der Item Skala-Statistik (siehe Tab. 11) die Trennschärfe der einzelnen Items herauslesen. Das Item pres 14 weist eine hohe Trennschärfe auf, während die der übrigen Items als mittel eingestuft wird.

Item-Skala-Statistiken

	Skalenmittelwert, wenn Item weggelassen	Skalenvarianz, wenn Item weggelassen	Korrigierte Item-Skala-Korrelation	Quadrierte multiple Korrelation	Cronbachs Alpha, wenn Item weggelassen
pres_3	13,611	13,065	**,474**	,226	**,650**
pres_9	13,613	13,765	**,398**	,199	**,680**
pres_13	13,645	13,721	**,459**	,252	**,658**
pres_14	14,207	11,451	**,560**	,330	**,610**
pres_15	14,513	12,815	**,421**	,233	**,674**

Tab. 11. Item-Skala-Statistik der Skala pressure and tension

Das Cronbach-Alpha als Quantifizierung der Höhe der Itemkorrelation ergab für die 6 Items der Skala *pressure and tension* einen Wert von 0,704. Als Maß für die Reliabilität kann ein solcher Wert als gut angesehen werden. Die rechte Spalte zeigt darüber hinaus, dass die Reliabilität der gesamten Skala schlechter wird, wenn man ein Item weglässt.

4.1.3. Korrelationen zwischen den Skalen perceived competence und pressure and tension

Im Anschluss an diese erste Reduktion konnte die Korrelation zwischen den Skalen *perceived competence* und *pressure tension* betrachtet werden. Tabelle 12 zeigt die bivariate Pearson-Korrelation zwischen den einzelnen Items der beiden Skalen. Dabei sieht man wie stark jedes Item mit den anderen korreliert und erhält dabei auch Auskunft über die jeweilige Signifikanz.

Betrachtet man die Items der Skala *perceived competence*, so korrelieren diese jeweils untereinander signifikant. Die Höhe dieser skaleninternen Korrelation ist überwiegend hoch und stets positiv. Greift man nun eines der Items heraus, etwa das Item pco 10, so sieht man, dass es mit den Items der *pressure* Skala negativ und niedrig korreliert. Dieses Bild ergibt sich auch für die übrigen Items der *perceived competence* Skala. In zwei Fällen (pco 7 mit pres 3 und pco 11 mit pres 15) ist die Korrelation der Items überhaupt positiv und bewegt sich auf einem sehr niedrigen Niveau (≤ 0,3). Blickt man auf die Signifikanz dieser skalenübergreifenden Item Korrelationen, so sind nur jene der Items pres 9 und pres 13 mit den Items der *perceived competence* Skala signifikant.

Korrelationen

		pco_3	pco_7	pco_10_u	pco_11_u	pres_3	pres_9	pres_13	pres_15
pco_3	Korrelation nach Pearson	1	,591**	,449**	,389**	-,118	-,470**	-,233**	-,043
pco_7	Korrelation nach Pearson	,591**	1	,441**	,296**	**,027**	-,327**	-,309**	-,058
pco_10_u	Korrelation nach Pearson	**,449****	**,441****	**1**	**,437****	**-,073**	**-,419****	**-,289****	**-,117**
pco_11_u	Korrelation nach Pearson	,389**	,296**	,437**	1	-,095	-,447**	-,191*	**,030**
pres_3	Korrelation nach Pearson	-,118	,027	-,073	-,095	1	,208*	,243**	,372**
pres_9	Korrelation nach Pearson	-,470**	-,327**	-,419**	-,447**	,208*	1	,387**	,176*
pres_13	Korrelation nach Pearson	-,233**	-,309**	-,289**	-,191*	,243**	,387**	1	,155
pres_15	Korrelation nach Pearson	-,043	-,058	-,117	,030	,372**	,176*	,155	1

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

Tab. 12. Korrelationen zwischen den Items der Skalen perceived competence und pressure and tension

4.1.4. Faktoranalyse der übrigen Skalen

Etwas weniger detailliert beschrieben folgt nun die Analyse der anderen fünf überarbeiteten Subskalen des IMIs.

Die fünf Items umfassende Skala *effort and importance* liefert im Rahmen der Faktorenanalyse gute Ergebnisse. Tabelle 13 zeigt, dass die Items eine Faktorenladung von mindestens 0,546 auf die Komponente aufweisen. Nach dem Kaiser-Gutmann-Kriterium ist nur ein Faktor zur Abbildung der Skala notwendig. Er erklärt 46,36% der Gesamtvarianz der Items. Die Items eff 4, eff 6 und eff 9 weisen eine hohe Trennschärfe auf, die der übrigen ist als mittel einzustufen. Das Cronbach-Alpha, als Quantifizierung der Höhe der Itemkorrelation, ergab für die 5 Items der Skala *effort and importance* einen Wert von 0,691. Als Maß für die Reliabilität kann ein solcher Wert als gut angesehen werden.

Komponentenmatrix[a]

	Komponente 1
eff_4	,725
eff_6	,743
eff_7	,546
eff_8	,635
eff_9	,734

Extraktionsmethode: Hauptkomponentenanalyse.

Tab. 13. Komponentenmatrix der Skala effort / importance

Rotierte Komponentenmatrix[a]

	Komponente 1	Komponente 2
rel_L_1	**,759**	,020
rel_L_6	**,723**	,214
rel_L_15	**,716**	,144
rel_S_2	,314	**,747**
rel_S_4	,190	**,790**
rel_S_16	-,027	**,536**

Extraktionsmethode: Hauptkomponentenanalyse.

Tab. 14. Komponentenmatrix der Skala relatedeness

Im Rahmen der Hauptkomponentenanalyse zerfällt die Skala *relatedeness* in zwei Faktoren. Wie in Tabelle 14 zu sehen, weisen die Items zur sozialen Interaktion der Lernenden mit der Lehrperson eine hohe Ladung auf den ersten und geringe Querladungen auf den zweiten Faktor auf. Bei den Items zur sozialen Interaktion der Lernenden untereinander zeigt sich ein umgekehrtes Bild.

Betrachtet man die beiden Teilkonstrukte für sich, so ergibt sich für das Cronbach-Alpha der Items zur sozialen Interaktion der Lernenden mit der Lehrperson ein Wert von 0,612. Jener Wert für die Items zur sozialen Interaktion der Lernenden untereinander liegt hingegen nur bei 0,514. Das Cronbach-Alpha erhöht sich bei Ausschluss des Items rel S16: „Zusammen mit anderen kann ich bessere Ergebnisse erzielen.“ auf einen Wert von 0,670. Darüber hinaus weist dieses Item mit einem Wert von 0,168 eine äußerst geringe Trennschärfe auf. Die Trennschärfe der übrigen Items beträgt zwischen 0,413 und 0,488 und kann als mittel eingestuft werden.

Die Skala *interest and enjoyment* umfasste im Rahmen des Fragebogens ursprünglich zwar 6 Items, aus inhaltlichen Gründen wurde das Item int 2 jedoch bereits vor der Faktorenanalyse ausgeschlossen. Tabelle 15 zeigt, dass die Items eine Faktorenladung von mindestens 0,647 auf die Komponente aufweisen. Nach dem Kaiser-Gutmann-Kriterium ist nur ein Faktor zur Abbildung der Skala notwendig. Er erklärt 59,78% der Gesamtvarianz der Items. Abgesehen von int 21 (0,366) weisen die übrigen Items eine hohe Trennschärfe (> 0,557) auf. Das Cronbach-Alpha als Quantifizierung der Höhe der Itemkorrelation ergab für die 5 Items der Skala einen Wert von 0,839. Als Maß für die Reliabilität kann ein solcher Wert als gut angesehen werden.

Komponentenmatrix[a]

	Komponente 1
int_3	,838
int_5	,840
int_6	,793
int_20	,827
int_21	,517

Extraktionsmethode: Hauptkomponentenanalyse.

Tab. 15. Komponentenmatrix der Skala interest and enjoyment

Eine Reduktion der Items aus inhaltlichen Gründen wurde auch bei der Skala *value and usefulness* vorgenommen. Insgesamt fünf Items standen nach Streichung des Items val 7 für eine Faktorenanalyse zur Verfügung. Die Tabelle 16 zeigt, dass die Items eine Faktorenladung von mindestens 0,588 auf die Komponente aufweisen. Nach dem Kaiser-Gutmann-Kriterium ist nur ein Faktor zur Abbildung der Skala notwendig. Er erklärt 47,23% der Gesamtvarianz der Items. Die Items weisen mit Werten zwischen 0,5 und 0,569 eine mittelhohe Trennschärfe auf. Das Cronbach-Alpha, als Quantifizierung der Höhe der Itemkorrelation, ergab für die 5 Items der Skala *value and usefulness* einen Wert von 0,737, was als gut angesehen werden kann.

Komponentenmatrix[a]

	Komponente 1
val_4	,589
val_6	,710
val_9	,766
val_13	,760
val_14	,588

Extraktionsmethode: Hauptkomponentenanalyse.

Tab. 16. Komponentenmatrix der Skala value and usefulness

Zu guter Letzt wurde die Skala *perceived choice* betrachtet. Die Hauptkomponentenanalyse ergab, dass die Skala in zwei Komponenten zerfällt. Schließt man jedoch das Item pch 13 aus, so sind die übrigen Items der Skala lediglich einer Komponente zuordenbar. Auch das Item pch 10 wurde aus inhaltlichen Gründen ausgeschlossen. Die Faktorladungen der verbleibenden vier Items sind, wie in Tabelle 17 zu sehen ist, mit Werten von zumindest 0,623 ausreichend hoch. Lediglich das Item pch 12 lädt mit einem Wert von 0,490 auf die Skala.

Komponentenmatrix[a]

	Komponente 1
pch_2	,711
pch_5	,623
pch_11	,736
pch_12	,490

Extraktionsmethode: Hauptkomponentenanalyse.

Tab. 17. Komponentenmatrix der Skala perceived competence

Nach dem Kaiser-Gutmann-Kriterium ist nur ein Faktor zur Abbildung der Skala notwendig. Dieser erklärt 41,92% der Gesamtvarianz der Items. Mit Werten zwischen 0,253 und 0,391 ist die Trennschärfe als mittel einzustufen. Das Cronbach-Alpha, als Quantifizierung der Höhe der Itemkorrelation, ergab für die 4 Items der Skala *perceived choice* einen Wert von 0,537. Als Maß für die Reliabilität kann ein solcher Wert als gut angesehen werden.

4.1.5. Faktoranalyse von Items mehrere Skalen gemeinsam

Die verbliebenen 35 Items der sieben Subskalen des Fragebogens, welche tabellarisch zusammengefasst im Anhang 10.5 zu sehen sind, standen nun für eine gemeinsame Faktorenanalyse zur Verfügung. Da die Faktorstruktur bei gleichzeitiger Analyse aller Subskalen nicht reproduzierbar war, werden im Folgenden immer einzelne Skalen miteinander betrachtet.

Zunächst wurde getestet, mit welchem Faktor die einzelnen Items auf die Komponenten laden. In Tabelle 18 ist die rotierte Komponentenmatrix der ausgewählten Items der Skalen *pressure and tension*, *perceived competence*, value and usefulness und *perceived choice* zu sehen.

Dabei wurden zur besseren Übersichtlichkeit Ladungen auf Komponenten mit Werten unter 0,25 unterdrückt.

Es ist gut ersichtlich, dass alle Items eindeutig auf die vorgesehenen Skalen laden und nur vereinzelnd geringe Querladungen zu anderen Skalen auftreten. Einzig die Items pres 9 und pres 13 weisen hohe negative Querladungen auf die Skala *perceived competence* auf. Dies war bereits in der Korrelationstabelle (siehe Tab. 12) ersichtlich.

Rotierte Komponentenmatrix[a]

	Komponente			
	1	2	3	4
pres_3			**,715**	
pres_9	-,674		**,296**	
pres_13	-,467		**,416**	
pres_14	-,291		**,750**	
pres_15			**,755**	
pco_3	**,733**			
pco_7	**,686**			
pco_10_u	**,749**			
pco_11_u	**,701**			
val_4		**,526**		
val_6	,282	**,577**		
val_9		**,754**		
val_13		**,699**		
val_14		**,595**		
pch_2				**,710**
pch_5	,261			**,533**
pch_11				**,751**
pch_12				**,417**

Extraktionsmethode: Hauptkomponentenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

Tab. 18. Faktorenanalyse der ausgewählten Items der Skalen pres, pco, val und pch

Mittels Kaiser-Gutmann-Kriterium erfolgte eine Überprüfung, durch wie viele Faktoren sich die Items abbilden lassen. Dabei wurden mehr Faktoren extrahiert als im Konstrukt angenommen, weswegen im Rahmen der Hauptkomponentenanalyse eine feste Anzahl von Faktoren vorgegeben wurde. Die vier angenommenen Komponenten erklären 50,61% der Gesamtvarianz der Items.

Das Cronbach-Alpha, als Quantifizierung der Höhe der Itemkorrelation, ergab einen Wert von 0,753. Als Maß für die Reliabilität kann ein solcher Wert als gut angesehen werden.

Darüber hinaus wurde eine Faktorenanalyse mit den Items der Skalen *interest and enjoyment*, *relatedeness*, *pressure and tension* und *perceived competence* durchgeführt. Tabelle 19 zeigt die rotierte Komponentenmatrix, aus der ersichtlich ist, mit welchen Werten die Items auf die jeweiligen Komponenten laden und ob hohe Querladungen auftreten. Auch hier wurden zwecks besserer Übersichtlichkeit Ladungen kleiner 0,25 unterdrückt.

Rotierte Komponentenmatrix[a]

	Komponente				
	1	2	3	4	5
int_3	**,734**				
int_5	**,771**				
int_6	**,681**	,331			
int_20	**,672**	,411			
int_21	**,611**				
rel_L_1					**,841**
rel_L_6		,506		,351	**,374**
rel_L_15		,386			**,644**
rel_S_2				**,788**	
rel_S_4				**,734**	
pres_3			**,722**		
pres_9	-,337	-,367	**,416**	-,356	
pres_13		-,252	**,403**	-,424	
pres_14		-,258	**,759**		
pres_15			**,749**		
pco_3	,319	**,757**			
pco_7	,257	**,754**			
pco_10_u		**,703**			
pco_11_u	,369	**,325**		,448	

Extraktionsmethode: Hauptkomponentenanalyse.

Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Die Rotation ist in 6 Iterationen konvergiert.

Tab. 19. Faktorenanalyse der ausgewählten Items der Skalen int, rel, pres und pco

Man sieht, dass zwar ein Großteil der Items eindeutig auf die vorgesehenen Skalen laden, jedoch auch vermehrt Querladungen auftreten. Das Item rel L6 lädt unerwünscht hoch auf die pco-Skala, pres 9 und pres 13 weisen allgemein hohe negative Querladungen auf andere Skalen auf und drei der *perceived competence* Items laden mittel hoch auf die *interest and enjoyment* Skala. Durch Anwenden des

Kaiser-Gutmann-Kriteriums konnte festgestellt werden, dass die fünf angenommenen Komponenten 61,08% der Gesamtvarianz der Items erklären.

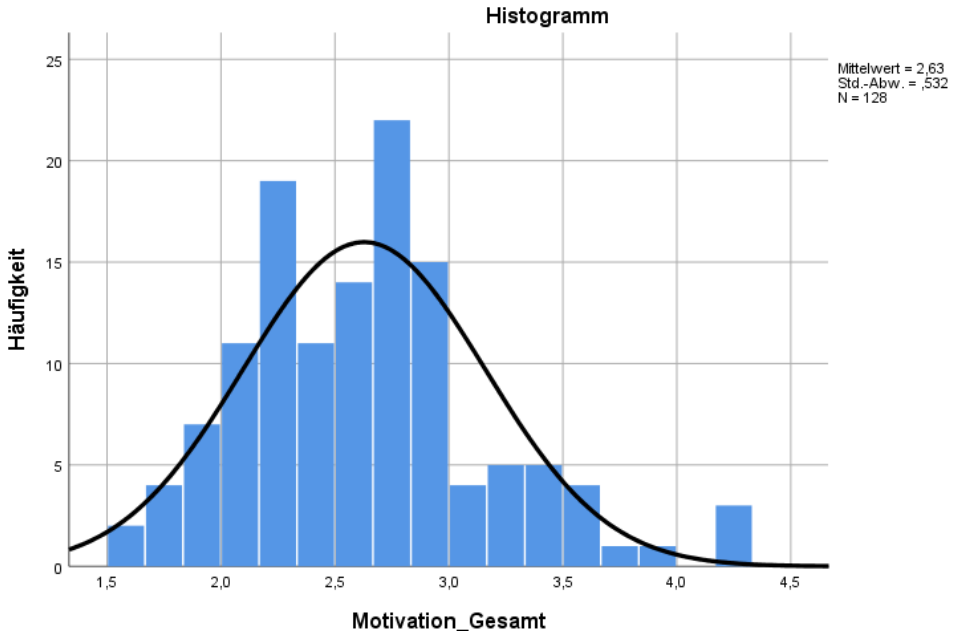
Das Cronbach-Alpha, als Quantifizierung der Höhe der Itemkorrelation, ergab einen Wert von 0,565. Als Maß für die Reliabilität kann ein solcher Wert als gut angesehen werden.

4.2. Inhaltliche Analyse der Ergebnisse

Nachdem die Items des Fragebogens sowohl hinsichtlich statistisch-mathematischer (mittels Faktorenanalyse) als auch inhaltlich-sprachlicher Gesichtspunkte analysiert wurden, verblieben 33 Items zur Messung der intrinsischen Motivation von Schülerinnen und Schülern (siehe Fragebogenversion kurz im Anhang 10.1.2.). Dabei verfügt jede Skala über mindestens 4 Items, sodass die Konstrukte ausreichend in ihrer Breite abgebildet sind. Im Anhang 10.5 ist die jeweilige Zuordnung der einzelnen verbliebenen Items zu den sieben Subskalen tabellarisch zusammengefasst.

Bevor nun im Rahmen einer ANOVA verschiedene Gruppen hinsichtlich ihrer Motivation miteinander verglichen wurden, galt es vorab einige Grundvoraussetzungen zu überprüfen, wie im Theoriekapitel 2.5.4. beschrieben.

Die Annahme, dass das untersuchte Merkmal in der Grundgesamtheit normalverteilt ist, kann in unserem Fall nur im Sample getestet werden. Die, in Abbildung 3 dargestellte Verteilung ist etwas linksschief, kann jedoch als ausreichend normalverteilt beurteilt werden.



3. Abbildung: Histogramm des Merkmals im Vergleich zur Normalverteilung

Die Gleichheit der Varianzen (Varianzhomogenität) kann mittels Levene-Tests überprüft werden. Er ist ein Signifikanz Test, dessen Nullhypothese davon ausgeht, dass die Fehlervarianzen über die Gruppen hinweg gleich sind. Wie in Tabelle 20 zu sehen ist, kann die Nullhypothese aufgrund des zu hohen Signifikanzniveaus (p = 0,216) nicht abgelehnt werden.

Levene-Test auf Gleichheit der Fehlervarianzen[a,b]

		Levene-Statistik	df1	df2	Sig.
Motivation_Gesamt	Basiert auf dem Mittelwert	1,468	4	123	**,216**
	Basiert auf dem Median	1,256	4	123	,291
	Basierend auf dem Median und mit angepaßten df	1,256	4	116,584	,292
	Basiert auf dem getrimmten Mittel	1,412	4	123	,234

Prüft die Nullhypothese, daß die Fehlervarianz der abhängigen Variablen über Gruppen hinweg gleich ist.

a. Abhängige Variable: Motivation_Gesamt

b. Design: Konstanter Term + ab_Schule

Tab. 20. Levene-Test auf Homogenität der Varianzen

Damit ist die Voraussetzung der Varianzhomogenität gegeben und eine ANOVA kann durchgeführt werden. Dabei galt es zunächst folgende Hypothese zu testen:

> *Die einzelnen **Schulen** unterscheiden sich hinsichtlich der intrinsischen Motivation ihrer Schülerinnen und Schüler*

Tabelle 21 gibt einen Überblick über die erzielten Mittelwerte im Schulvergleich, deren Höhe für die Ausprägung der Motivation steht. In der rechten Spalte ist die Anzahl der Schülerinnen und Schüler der jeweiligen Schule zu sehen.

Der Mittelwert aller Skalen für die 33 ausgewählten Items liegt bei 2,63 mit einer Standardabweichung von 0,53. Man kann erkennen, dass die Höhe der Motivation der Lernenden aus der Schule A am niedrigsten und aus der Schule D am höchsten war. Die Schulen B, C und D befinden sich hinsichtlich der Motivation in der Nähe des Mittelwerts.

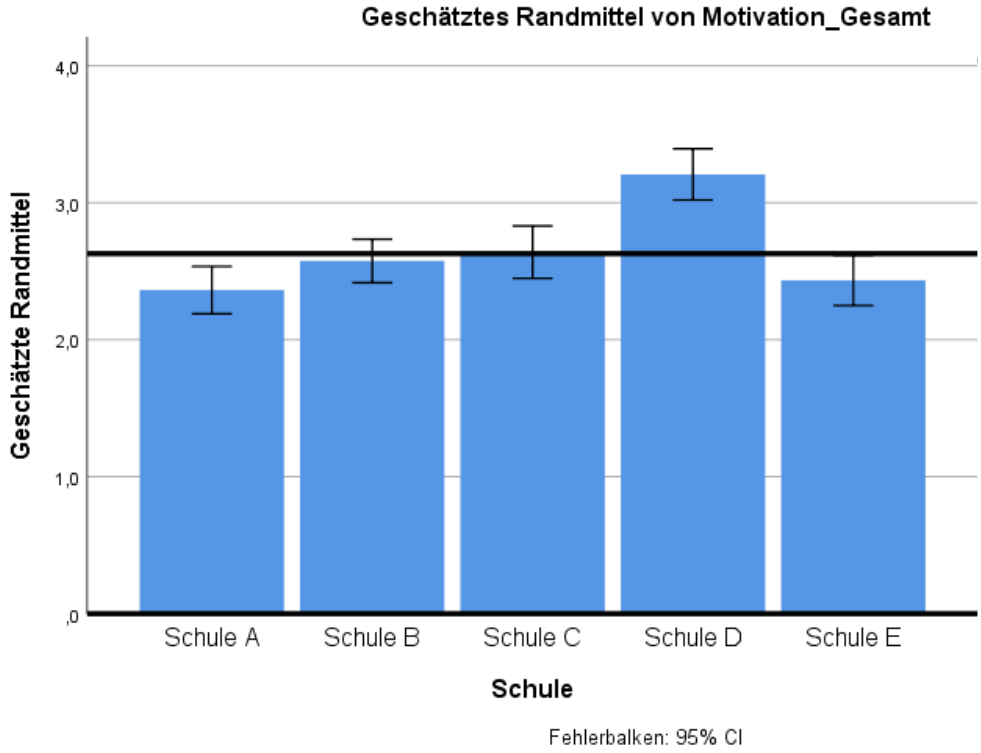
Deskriptive Statistiken

Abhängige Variable: Motivation_Gesamt

Schule	Mittelwert	Std.-Abweichung	N
Schule A	2,362	,4343	27
Schule B	2,575	,3614	32
Schule C	2,639	,4296	22
Schule D	3,207	,5680	23
Schule E	2,433	,4863	24
Gesamt	2,628	,5323	128

Tab. 21. Mittelwerte der unterschiedlichen Schulen

In Abbildung 4 findet sich eine graphische Veranschaulichung der Mittelwerte der einzelnen Schulen in Form eines Histogramms.



4. Abbildung: Mittelwerte der unterschiedlichen Schulen in einem Histogramm

Um zu überprüfen, ob sich die Schulen hinsichtlich ihrer Motivation signifikant unterscheiden, wird die folgende Nullhypothese aufgestellt, die mittels einfaktorieller Varianzanalyse getestet wird: Der Faktor Schule hat keinen Einfluss auf die intrinsische Motivation.

Wie in Tabelle 22 zu sehen ist, hat der Faktor Schule einen höchst signifikanten Einfluss ($p < 0{,}001$), weswegen die Nullhypothese verworfen werden muss. Es ist also davon auszugehen, dass sich die Lernenden der einzelnen Schulen hinsichtlich ihrer intrinsischen Motivation signifikant unterscheiden.

Einfaktorielle ANOVA

Motivation_Gesamt

	Quadratsumme	df	Mittel der Quadrate	F	Signifikanz
Zwischen den Gruppen	10,623	4	2,656	12,878	**,000**
Innerhalb der Gruppen	25,364	123	,206		
Gesamt	35,987	127			

Tab. 22. Die ANOVA zeigt signifikante Unterschiede zwischen den Schulen

Zuletzt musste die Grundgesamtheit hinsichtlich weiterer Merkmale untersucht werden. Dazu wurden einerseits die Jugendlichen nach **Schulstufe** und andererseits nach **Schulformen** hinsichtlich ihrer intrinsischen Motivation untersucht.

Tabelle 23 gibt einen Überblick über die erzielten Mittelwerte im Schulstufenvergleich, wobei deren Höhe für die Ausprägung der Motivation steht. In der rechten Spalte ist die Anzahl der Schülerinnen und Schüler pro Schulstufe zu sehen.

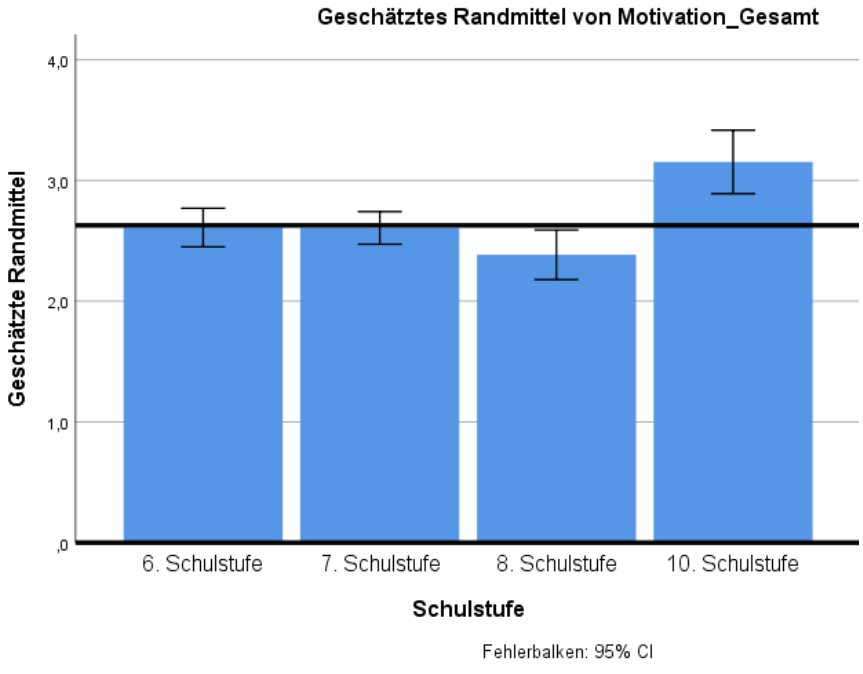
Deskriptive Statistiken

Abhängige Variable: Motivation_Gesamt

Schulstufe	Mittelwert	Std.-Abweichung	N
6. Schulstufe	2,610	,6084	38
7. Schulstufe	2,607	,3997	53
8. Schulstufe	2,385	,4554	23
10. Schulstufe	3,154	,5656	14
Gesamt	2,628	,5323	128

Tab. 23. Mittelwerte der unterschiedlichen Schulstufen

Man kann erkennen, dass die Höhe der Motivation der Lernenden aus der zehnten Schulstufe am größten und aus der achten Schulstufe am niedrigsten ist. Die intrinsische Motivation der Lernenden aus der sechsten und siebenten Schulstufe ist nahezu ident. In Abbildung 5 sind diese Ergebnisse graphisch veranschaulicht. Darüber hinaus sieht man, dass sich für die 10. Schulstufe der Fehlerbalken nicht mit den anderen überschneidet, was ein Hinweis auf eine hochsignifikante Abweichung ist.



5. Abbildung: Mittelwerte der unterschiedlichen Schulstufen in einem Histogramm

Ein Vergleich der verschiedenen Schulformen bezüglich der intrinsischen Motivation findet sich im Folgenden. Tabelle 24 gibt einen Überblick über die erzielten Mittelwerte je Schulform, wobei auch hier deren Höhe für die Ausprägung der Motivation steht. In der rechten Spalte ist die Anzahl der Jugendlichen pro Schulform zu sehen.

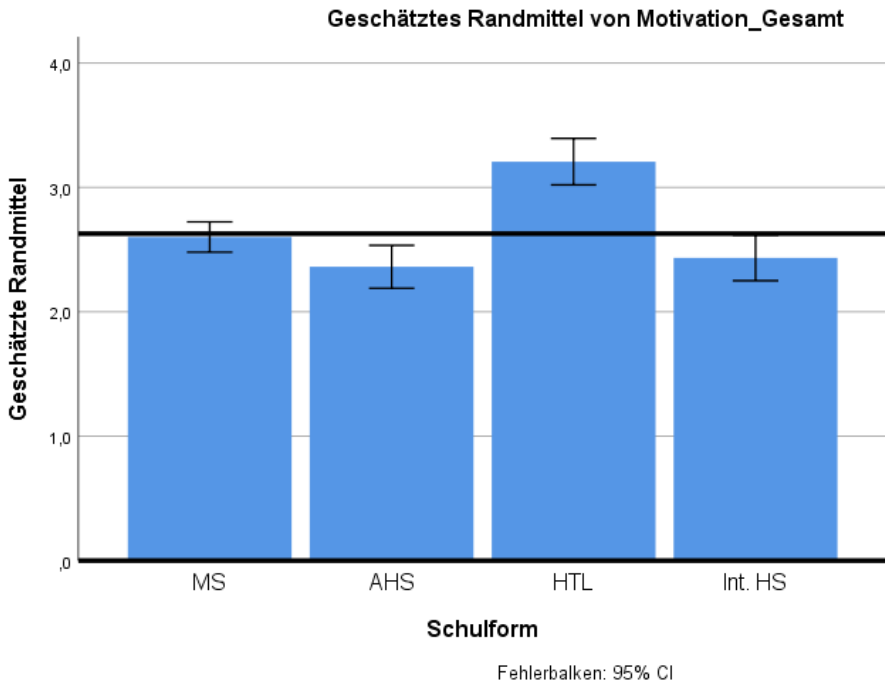
Deskriptive Statistiken

Abhängige Variable: Motivation_Gesamt

Schulform	Mittelwert	Std.-Abweichung	N
MS	2,601	,3880	54
AHS	2,362	,4343	27
HTL	3,207	,5680	23
Int. HS	2,433	,4863	24
Gesamt	2,628	,5323	128

Tab. 24. Mittelwerte der unterschiedlichen Schulformen

Man kann erkennen, dass die Höhe der Motivation der Lernenden aus der HTL am größten und aus der AHS am niedrigsten ist. Die intrinsische Motivation der Lernenden aus der Mittelschule, die die größte Stichprobe bildet, sowie der International Highschool, ist nahe am Gesamtmittelwert. In Abbildung 6 sind diese Ergebnisse graphisch veranschaulicht. Auch hier ist zu sehen, dass sich für die HTL der Fehlerbalken nicht mit den anderen überschneidet, was ein Hinweis auf eine hochsignifikante Abweichung ist.



6. Abbildung: Mittelwerte der unterschiedlichen Schulformen in einem Histogramm

5. Diskussion

Ziel dieses Kapitels ist es, die Ergebnisse und Analysen aus dem vorherigen zu diskutieren und zu interpretieren. Zunächst werden die Skalen einzeln und anschließend in ihrer Gesamtheit betrachtet und dabei auf die Forschungsfragen eingegangen. In diesem Rahmen gilt es, wenn möglich, den Bogen zum Theorieteil zu spannen und dabei Widersprüche und Übereinstimmungen aufzuzeigen. Darüber hinaus werden abschließend Güte und Qualität des Fragebogens beurteilt.

5.1. Diskussion der Skalen perceived competence und pressure / tension

Im Zuge der Beantwortung der ersten Forschungsfrage wurden in einem ersten Schritt, wie in Kapitel 3.2.1. und 3.2.2. beschrieben, die Skalen *perceived competence* und *pressure and tension* näher analysiert und überarbeitet.

Dabei stellte sich die Frage, ob, anstelle der Items zur *perceived competence* Skala, auf bestehende Items zur Selbstwirksamkeitserwartung nach Bandura (1977) zurückgegriffen werden soll. Obwohl sich die Wirksamkeitsüberzeugung, als eine der zentralen verhaltenssteuernden Faktoren Banduras Theorie, inhaltlich mit der Skala des IMIs überschneidet, wurde die Skala *perceived competence* beibehalten. Grund dafür ist der hier verfolgte, partizipative Ansatz, bei welchem die Jugendlichen beim Finden von Evidenzen und Validieren der Items eingebunden waren. Damit ist eine zielgruppengerechte Formulierung gewährleistet.

Im Anschluss an die Testung des Fragebogens an den Schulen, konnten die Faktorenanalysen und Korrelationstabellen des Abschnitts 4.1. erstellt werden, welche nun aus inhaltlicher Perspektive diskutiert werden. Die Faktorenanalyse der *perceived competence* Skala zeigt, dass das Item pco 5: „Ich beantworte Aufgabenstellungen/Fragen schnell.“ nur mit einem Wert von 0,447 auf die Skala lädt. Aus inhaltlicher Sicht kann die schlechte Faktorladung damit erklärt werden, dass die Formulierung des Items nicht zugänglich für introvertierte, schüchterne Schülerinnen und Schüler ist. Um jedoch nicht gänzlich auf den abgefragten Aspekt zu verzichten, wurde folgende Neuformulierung vorgenommen: „Wenn ich gefragt werde, fällt mir sofort eine Antwort ein.“ (pco 12, aus pco 5).

Darüber hinaus wurde das Item pco 9: „Wenn ich angesprochen werde, sind meine Antworten unsicher.“ gestrichen. Da hier neben der Kompetenz weitere Faktoren wie das soziale Umfeld, das Klassenklima und die eigene Persönlichkeit eine Rolle spielen, ist meiner Meinung nach der Rückschluss von einer unsicheren Antwort auf die wahrgenommene Kompetenz nicht zulässig.

Bei der Faktorenanalyse der *pressure and tension* Skala fiel auf, dass das Item pres 4: „Ich schaue mich in der Klasse um, ob mir jemand bei der Aufgabe helfen kann.“ mit einem Wert von -0,226 auf die Skala lädt. Das negative Vorzeichen lässt sich inhaltlich damit erklären, dass sich Jugendliche bei einem Angstzustand nicht im Klassenraum umdrehen, weshalb das Item ein negativer Prädiktor der Skala ist. Faktoranalytisch bildet die Skala bei Ausschluss des Items pres 4 nur noch in einen Faktor und die einzelnen Items zeigen mit Werten größer 0,62 gute Ladungen auf die Skala, weshalb das Item gestrichen wurde.

Da aus der Diplomarbeit von Schmidt (2017) folgen könnte, dass die *pressure tension* Skala überflüssig ist, wurde das Konstrukt zusammen mit jenem der *perceived competence* Skala genauer unter die Lupe genommen. Im Anschluss an die Überarbeitung und Präzisierung der Skalen zeigten die Korrelationstabellen des Kapitels 4.1.3., dass die *perceived competence* Items mit den *pressure tension* Items nicht signifikant, nicht immer negativ und auch nicht hoch miteinander korrelierend sind. Aus diesem Grund ist der Verzicht auf eine der Skalen nicht empfehlenswert. In weiterer Folge wurden beide Konstrukte beibehalten, um die intrinsische Motivation zu messen.

Um zu evaluieren, ob die Skalen auch bei gleichzeitiger Betrachtung mehrerer Skalen trennscharf sind, galt es eine gemeinsame Faktorenanalyse der Skalen *pressure and tension*, *perceived competence*, *value and usefulness* und *perceived choice* durchzuführen. Wie in Kapitel 4.1.5. erläutert, laden alle Items auf die vorgesehenen Skalen. Damit gelang es, trennscharfe, stabile Skalen zu den Konstrukten *pressure and tension* und *perceived competence* zu konstruieren, welche sich faktorenanalytisch abgrenzen lassen. Limitierend sei an dieser Stelle erwähnt, dass die Items pres 9 und pres 13 zwar hohe Querladungen auf andere Skalen aufweisen, jedoch aus inhaltlichen Gründen nicht auf sie verzichtet wurde.

5.2. Diskussion der übrigen Skalen

Im Rahmen der zweiten Forschungsfrage sollte untersucht werden, ob durch Überarbeitung der Konstrukte, auch die übrigen fünf Subskalen bei gleichzeitiger Testung aller Items trennscharf und stabil sind. Wie im Kapitel 3.3. erläutert, wurden zunächst die einzelnen Skalen sowohl nach statistisch-mathematischen als auch nach inhaltlichen Gesichtspunkten begutachtet und adaptiert, die Itemanzahl dabei deutlich reduziert, sowie abschließend mit Frau Dr. Korner expertenvalidiert. Im Anschluss an die Testung des Fragebogens an den Schulen konnten die Faktorenanalysen und Korrelationstabellen des Abschnitts 4.1.4. und 4.1.5. erstellt und diskutiert werden.

Betrachtet man die Analyse der *effort and importance* Skala aus Kapitels 4.1.4., so brachte diese gute Ergebnisse, weswegen alle 5 Items beibehalten wurden.

Die Skala *relatedeness* sollte nach Pieler (2018) in zwei Faktoren zerfallen, was im Rahmen der Hauptkomponentenanalyse in dieser Stichprobe bestätigt werden konnte. Auf der einen Seite wiesen die Items zur sozialen Interaktion der Lernenden mit der Lehrperson eine hohe Ladung auf den ersten und geringe Querladungen auf den zweiten Faktor auf. Bei den Items zur sozialen Interaktion der Lernenden untereinander zeigte sich andererseits ein umgekehrtes Bild. Wegen seiner geringen Faktorladung, schlechten Trennschärfe und geringen Reliabilität, sticht das Item rel S16: „Zusammen mit anderen kann ich bessere Ergebnisse erzielen.“ heraus. Bei der Suche nach einer Erklärung halfen zwei ähnliche Notizen von Lernenden, die sinngemäß die Frage „Wann?“ auf ihren Bögen notierten. Den Schülerinnen und Schülern war also nicht klar, worauf sich dieses Item bezog. Dies deutet auf eine zu unspezifische Formulierung des Items hin, sodass dieses, sowohl aus statistisch mathematischer, als auch inhaltlicher Sicht ausgeschlossen werden konnte.

Die von Klingenböck (2016) überarbeitete *interest and enjoyment* Skala konnte durch eine deutliche Itemreduktion sowie Formulierung zweier neuer Items (int 20 und int 21) trennscharf konstruiert werden. Lediglich das Item int 21: „In meiner Freizeit spreche ich über den Inhalt des Unterrichts mit meinen FreundInnen.“ wies eine etwas geringe Trennschärfe (0,366) sowie Ladung auf den Faktor (0,517) auf. Da die Werte aus statistischer Sicht toleriert werden können und mit dem Item das Konstrukt in

seiner Breite abgebildet wird, wurde das Item beibehalten. Das Item int 2: „Ich bin bereit, so etwas wieder zu machen.“ wurde hingegen gestrichen, da es aus inhaltlicher Perspektive keine neuen Erkenntnisse liefert. Darüber hinaus konnte bei der Testung aller Items die gewünschte Skalenstruktur nachgewiesen werden, womit es gelang, die Ergebnisse von Klingenböck (2016) zu erweiterten.

Obwohl die Skala *value and usefulness* faktoranalytisch gute Resultate zeigt, überschneidet sich das Item val 7: „Die Aufgabe fordert mich und bringt mich zum Nachdenken.“ inhaltlich mit anderen Skalen und ist damit nicht eindeutig zuordenbar, weswegen von ihm abgesehen wurde.

Schließlich wurde die Skala *perceived choice* analysiert. Streicht man hier das Item pch 13: „Ich kann wählen, ob ich die Aufgabe alleine oder mit MitschülerInnen bearbeite.“, so bildet die Skala nur noch in einen Faktor. Da im Unterricht die Wahl, entweder mit oder ohne Partner zu arbeiten, meist nicht gegeben ist, lässt sich dieser Schritt auch inhaltlich argumentieren. Bei Item pch 10: „Ich kann selbst entscheiden, wie genau ich die Aufgabe bearbeite.“ kann argumentiert werden, dass zu Lasten einer schlechteren Bewertung stets eine Aufgabe schnell und oberflächlich erledigt werden kann, sodass auch dieses Item ausgenommen wurde.

Um in weiterer Folge Trennschärfe und Faktorladungen der gekürzten Subskalen diskutieren zu können, wurden die in Anhang 10.5. tabellarisch zusammengefassten Items einer gemeinsamen Faktorenanalyse unterzogen. Im Rahmen der Hauptkomponentenanalyse wurden bei diesem Prozess mehr Faktoren extrahiert als inhaltlich begründet werden kann. Dies ist nach Bortz Schuster (2010) nicht verwunderlich, denn je mehr Skalen gemeinsam getestet werden, desto unschärfer ist die Faktorenstruktur. Dieser Umstand lässt sich damit erklären, dass die Konstrukte sehr fein sind und inhaltlich nahe beieinander liegen, so dass eine Reproduktion der Faktorstruktur nicht erwartet werden kann. Um dem Rechnung zu tragen, wurden bei der Hauptkomponentenanalyse lediglich 4 Subskalen gleichzeitig betrachtet und vorab die Anzahl von Faktoren festgelegt.

Wie in Kapitel 4.1.5. zu sehen, luden fast alle Items eindeutig auf die vorgesehenen Skalen und wiesen nur geringe Querladungen auf andere Skalen auf. Im Sinne einer eindeutigen Skalenstruktur sind nach Bühner (2011) Querladungen kleiner als die

Hälfte der erwünschten Ladung vertretbar, was hier überwiegend gegeben war. Lediglich die Items pres 13, int 20, rel L6, pres 13 und pco 11erfüllen diese Vorgabe nicht, wurden jedoch aus inhaltlicher Sicht beibehalten, um die Konstrukte in ihrer Breite zu erfassen.

Es fällt auf, dass drei der vier *perceived competence* Items Querladungen auf die *interest and enjoyment* Skala aufweisen. Während die *interest and enjoyment* Skala die intrinsische Motivation direkt misst, ist nach Deci und Ryan (2003) die *perceived competence* Skala ein positiver Prädiktor für die intrinsische Motivation. Damit verwundern die hohen Querladungen in der Höhe von 0,257 bis 0,369 nicht, und können als Untermauerung dieser Annahme gesehen werden.

In Anbetracht der hier diskutierten Analysen gelang es nicht nur, die Ergebnisse der Arbeiten von Korner (2015), Klingenböck (2016), Schmidt (2017) und Pieler (2018) wiederzugegeben und zu bestätigen, sondern diese sogar zu erweitern. So konnte erstmals bei Testung aller Items die Skalenstruktur reproduziert werden

5.3. Inhaltliche Diskussion der Ergebnisse

Nachdem die Items des Fragebogens, sowohl hinsichtlich statistisch-mathematischer als auch inhaltlich-sprachlicher Gesichtspunkte, analysiert und diskutiert wurden, verblieben 33 Items zur Messung der Motivation von Schülerinnen und Schülern. Dabei verfügt jede Skala über mindestens 4 Items, sodass die Konstrukte in ihrer Breite abgebildet sind. Im Anhang 10.5. ist das finale Set an Items zu den sieben Subskalen tabellarisch zusammengefasst. Im Folgenden konnten nun unterschiedliche Stichproben hinsichtlich ihrer Motivation miteinander verglichen werden. Vorab sei an dieser Stelle erwähnt, dass dem Fragebogen verschiedene Interventionen vorrausgingen.

Bei der Betrachtung aller Schülerinnen und Schüler ergab sich hinsichtlich der Motivation ein Mittelwert von 2,63 und eine Standardabweichung von 0,53. Werte nahe bei 5 stehen für eine hohe, Werte nahe bei 1 für eine niedrige Merkmalsausprägung.

Analysiert man die Grundgesamtheit hinsichtlich ihrer Motivation je Schule, so zeigt sich, dass signifikante Unterschiede bestehen (siehe Seite 67). Jugendliche der

Schule D, einer HTL, waren mit einem Mittelwert von 3,21 am motiviertesten. Mit einem Mittelwert von 2,36 schnitten die Kinder der Schule A, einem Gymnasium, hinsichtlich der Motivation am schlechtesten ab. Bei den Schulen B, C und D konnten keine signifikanten Unterschiede festgestellt werden, sie wiesen durchschnittliche Mittelwerte der Motivation auf.

Auf Grund des geringen Samples überraschte es nicht, dass sich ein ähnliches Bild beim Vergleich der Schultypen zeigte. Da die Schülerinnen und Schüler der HTL darüber hinaus einer höheren Schulstufe angehörten als die übrigen Kinder, scorten diese beim Vergleich der Schulstufen ebenfalls signifikant höher als die der übrigen Schulstufen.

5.4. Qualität des Fragebogens

Im Rahmen dieses Kapitels soll die Güte und Qualität des hier weiterentwickelten Messinstruments anhand der im Kapitel 2.5. beschriebenen Kriterien diskutiert werden.

Im Unterrichtsverlauf kann der Fragebogen im Anschluss an eine beliebige Intervention, die sich nicht nur auf Physik beschränkt, ausgeteilt werden. Diese Intervention kann nun hinsichtlich ihres Einflusses auf die Motivation der Lernenden analysiert werden. Im Vorfeld des Ausfüllens bedarf es einer kurzen Instruktion, welche immer gleich formuliert werden kann. Mit seinem geschlossenen Antwortformat, den in der Sprache der Jugendlichen formulierten Items, sowie seiner übersichtlichen Gestaltung, sollten sich kaum Rückfragen ergeben, sodass die Durchführungsobjektivität gegeben ist. Die Likert-skalierten Items gewährleisten, dass unabhängig von der auswertenden Person gleichwertige Ergebnisse erhalten werden, wodurch die Auswertungsobjektivität erfüllt ist. Da sowohl eine ausreichend große Normstichprobe als auch klare Interpretationsmöglichkeiten vorliegen, ist die Interpretationsobjektivität gewährleistet. Insgesamt kann der vorliegende Fragebogen als objektiv gewertet werden.

Um die Reliabilität abzuschätzen, wurden im Rahmen der Konsistenzmethode die Itemkorrelationen betrachtet. Das Cronbach-Alpha, als Quantifizierung dieser Korrelation, ergab einen Wert von 0,839, der als hoch angesehen werden kann.

Bei der Erstellung des Fragebogens wurde die Methode des *Four Building Blocks Approach* nach Wilson (2005) gewählt, um das zu messende Merkmal bestmöglich zu erfassen und damit Validität zu sichern.

Ebenso ist das Kriterium der Skalierbarkeit erfüllt, da die resultierenden Testwerte die Merkmalsrelation adäquat widerspiegeln.

Neben den vier eben diskutierten Hauptgütekriterien gibt es noch eine Reihe weiterer **Nebengütekriterien**. Dazu gehört die Normierung, welche gegeben ist, da die individuelle Merkmalsausprägung durch Vergleich mit anderen Lernenden eingeordnet werden kann. Außerdem ist der Fragebogen mit einer Durchführungsdauer von etwa 15 Minuten, der schnellen, einfachen Auswertung sowie dem geringen Materialeinsatz ökonomisch und damit sowohl Lernenden als auch Lehrenden zumutbar. Da keine einzelnen Gruppen diskriminiert werden, kann der Test als fair bewertet werden. Wie in Kapitel 2 bereits verdeutlich, existiert noch kein vergleichbares Testinstrument im deutschsprachigen Raum, das die intrinsische Motivation in dieser Genauigkeit messen kann, wodurch das Kriterium der Nützlichkeit erfüllt ist. Vergleichstest waren aus diesem Grund nur mit den bisherigen Arbeiten von Korner (2015), Klingenböck (2016), Schmidt (2017) und Pieler (2018) möglich. Im Rahmen der Analyse (siehe Kapitel 4.1.5.) konnten alle Skalen bis auf die *effort and importance* Skala sowie die *pressure and tension* Skala in unterschiedlichen Stichproben reproduziert werden.

Insgesamt erfüllt der Fragebogen die Haupt- und Nebengütekriterien, weshalb von einem guten psychometrischen Test gesprochen werden kann.

Zusammenfassend wurden im Rahmen der Faktorenanalyse gute Ergebnisse erzielt, sodass daraus ein Set von 33 getesteten und evaluierten, sowie einem, erst im Anschluss an die Pilotierung, neu konstruierten Item erstellt werden konnte (siehe Fragebogen kurz, Anhang 10.1.2.). Die sieben Subskalen dieser finalen Version des Fragebogens sind bei geleichzeitiger Messung aller Items trennscharf und stabil. Die Schaffung unidimensionaler Konstrukte gelang dabei vor allem durch die deutliche Reduktion der Itemanzahl pro Skala. Dies führte neben dem Effekt, Aspekte nicht doppelt abzufragen dazu, einen, in akzeptabler Zeit durchführbaren Fragebogen zu schaffen. Damit konnte das Ziel, einen bedeutenden Beitrag zur Finalisierung eines

validen, vor allem aber ökonomischen Messinstruments, welches sowohl klassischen Gütekriterien genügt, als auch in der Praxis gut einsetzbar ist, zu leisten, erreicht werden. Mit dem Fragebogen ist es nun möglich, in praktikabler Weise Studien über die Motivation von Lernenden im Unterricht durchzuführen.

5.5. Limitationen und Anregungen

Abschließend werden hier einige Limitationen diskutiert, sowie Impuls für weitere Arbeiten an dem Fragebogen gegeben.

Zunächst sei erwähnt, dass, durch die COVID Situation, die Suche nach unterstützenden Lehrpersonen schwieriger war als angenommen. Einige gaben an, aufgrund ausgefallener Stunden keine freie Unterrichtszeit zur Verfügung zu haben. Andere Lehrerinnen und Lehrer waren zwar bereit, die Befragung durchzuführen, bekamen allerdings eine Absage vonseiten der Direktion. Trotz unzähliger Anfragen führt dies dazu, dass die Gruppe der Schülerinnen und Schüler aus Mittelschulen (54 Teilnehmerinnen und Teilnehmer) sehr stark, jene aus Gymnasien (27 T. u T.) nur schwach vertreten war.

Da die Größe der Stichprobe nach Ausschluss zweier Schulen lediglich 158 Schülerinnen und Schüler umfasste, wäre eine erneute Testung des Fragebogens mit einer größeren Stichprobe zwecks Trennschärfeanalyse sinnvoll. In Anbetracht des eben erwähnten Ungleichgewichts der Befragten nach Schultyp, würde ich im Rahmen künftiger Erhebungen mehr Gymnasialklassen inkludieren.

In diesem Zusammenhang kann, zwecks größerer Stichprobe, angedacht werden, den Fragebogen neben Physik, Biologie und Chemie auch in anderen Schulfächern einzusetzen. Da sich einzelne Fragen auf Gruppenarbeiten ohne besondere Spezifizierung beziehen, sollte dies ohne weitere Adaptionen problemlos möglich sein. Damit wäre dann ein valides Testinstrument geschaffen, dass sich über Physik hinaus universell einsetzen ließe.

Als Limitation dieser Arbeit kann überdies der Umstand angesehen werden, dass die Befragung ausschließlich von den jeweiligen Fachlehrpersonen durchgeführt wurde. Trotz vorgegebenem Leitfaden zum Ablauf der Umfrage sowie einheitlichen Instruktionen, welche im Vorfeld erteilt werden sollten, könnte sich diese Gegebenheit

negativ auf die Qualität der Datenerhebung ausgewirkt haben. Hier empfehle ich, diese Rolle bei nachfolgenden Erhebungen selbst zu übernehmen, anstatt diese den Lehrpersonen zu überlassen.

Ungenaue Instruktionen und ein Abweichen von dem vorgegebenen Durchführungsablauf könnten Gründe für die unzureichenden Ergebnisse, welche zum Ausschluss zweier Schulen geführt haben, darstellen. Da es den Jugendlichen freistand, teilzunehmen oder nicht, deuten die zahlreich ausgelassenen Antworten auf den ausgeschlossenen Bögen auf Verständnisschwierigkeiten wegen mangelnder Deutschkenntnisse hin.

Ein spannender Ausgangspunkt künftiger Arbeiten wäre die Frage, ob sich durch Überarbeitung der *pressure and tension* Skala die Faktorstruktur besser reproduzieren lässt. Wie in Kapitel 4.1.5. zu sehen, wiesen speziell die Items pres 9 und pres 13 eine geringe Ladung auf die eigene, sowie hohe Querladungen auf andere Skalen auf. Darüber hinaus kann angedacht werden, die Items nach Rasch zu skalieren, um so noch einmal sicherzugehen, dass die volle Breite des jeweiligen Konstrukts abgebildet ist.

Zusätzlich kann im Rahmen künftiger Arbeit eine noch kompaktere Version entwickelt werden, die sich aus lediglich drei bis vier Items pro Skala zusammensetzt.

6. Zusammenfassung und Ausblick

Mit dieser Arbeit wurde ein weiterer Schritt zur Finalisierung des ersten deutschsprachigen Testinstruments, welches die intrinsische Motivation von Lernenden evaluieren kann, beigetragen. Damit ist es möglich, in praktikabler Weise Studien über die Motivation von Schülerinnen und Schülern im Unterricht durchzuführen.

Den Ausgangpunkt meiner Arbeit stellte das englischsprachige IMI von Deci und Ryan (2003) dar. Der Versuch einer Übersetzung ins Deutsche erwies sich als nicht zielführend, da die Gütekriterien eines validen psychometrischen Tests selbst nach einigen Überarbeitungsschleifen nicht erfüllt werden konnten. Aus diesem Grund entschied sich Korner (2015) dazu, nach der im Kapitel 2.6.3. beschrieben Methode des *Four Building Blocks Approachs* nach Wilson (2005) die Skalen des IMIs neu zu konstruieren. Im Anschluss an die Neukonstruktion der Skala *effort and importance* von Korner (2015) folgten Arbeiten zu den Skalen *interest and enjoyment* von Klingenböck (2016), *value and usefulness*, *perceived choice* und *perceived c*ompetence von Schmidt (2017) sowie zur Skala (*social) relatedness* von Pieler (2018).

Im Rahmen dieser Arbeit gelang es, trennscharfe Konstrukte zu den Skalen *perceived competence* und *pressure tension* zu entwickeln. Der Ausschluss einer der Skalen erwies sich als nicht zielführend, da die Konstrukte weder signifikant, noch wie ursprünglich angenommen, immer negativ und hoch korrelieren. Es kann daher davon ausgegangen werden, dass die beiden Konstrukte auch inhaltlich wenig korrelieren. Darüber hinaus wurden auch die übrigen fünf Subskalen überarbeitet und expertenvalidiert.

Nichtsdestotrotz wäre eine erneute Testung des Fragebogens zwecks einer Trennschärfeanalyse sinnvoll, da die Größe der Stichprobe nach Ausschluss zweier Schulen nur 158 Schülerinnen und Schüler umfasste. Hier kann angedacht werden, den Fragebogen auch in weiteren Unterrichtsfächern neben Physik, Biologie und Chemie einzusetzen. Um sicher zu gehen, dass die volle Breite des jeweiligen Konstrukts abgebildet ist, empfehlt es sich, die Items nach Rasch zu skalieren. Zusätzlich kann im Rahmen künftiger Arbeit eine Kurzversion des IMIs entwickelt werden, die sich aus lediglich drei bis vier Items pro Skala zusammensetzt.

Im Zuge der Hauptkomponentenanalyse konnte gezeigt werden, dass die Skalen bei gelichzeitiger Messung aller Items zueinander trennscharf und stabil sind. Obwohl sich bei der Analyse aller Skalen nicht die gesamte Faktorstruktur reproduzieren lies, konnte das Ziel, einen Beitrag zur Finalisierung eines validen, vor allem aber ökonomischen Messinstruments, welches sowohl klassischen Gütekriterien genügt als auch in der Praxis gut einsetzbar ist, zu leisten, erreicht werden.

7. Literatur:

Arnold, R. (2007): Ich lerne, also bin ich. Eine systemisch-konstruktivistische Didaktik.

Atkinson, J. W. (1975): Einführung in die Motivationsforschung. Stuttgart: Klett.

Bandura, A. (1977): Self-Efficacy - Toward a Unifying Theory of Behavioral Change. *Psychological Review, 84*(2), 191-215.

Beckmann, J., & Heckhausen, H. (2010): Motivation durch Erwartung und Anreiz. In J. Heckhausen & H. Heckhausen (Eds.), Motivation und Handeln. Berlin, Heidelberg: Springer Berlin Heidelberg.

Berger, R., & Hänze, M. (2004): Das Gruppenpuzzle im Physikunterricht der Sekundarstufe II–Einfluss auf Motivation, Lernen und Leistung. *Zeitschrift für Didaktik der Naturwissenschaften, 10*, 205-219.

Bortz, J., Schuster, C. (2010): Statistik für Human- und Sozialwissenschaftler. 7., vollständig überarbeitete und erweiterte Auflage. Springer-Verlag Berlin Heidelberg.

Bryant, F.B: (2000): Assessing the Validity of Measurement. In: L.G. Grimm & P.R. Yarnold (Eds.) Reading and understanding MORE multivariate statistics. Washington, DC: American Psychological Association.

Bühner, M. (2011): *Einführung in die Test-und Fragebogenkonstruktion*. München: Pearson.

Busker, M. (2014): Entwicklung eines Fragebogens zur Untersuchung des Fachinteresses. In D. Krüger, I. Parchmann & H. Schecker (Eds.), *Methoden in der naturwissenschafts-didaktischen Forschung*. Berlin, Heidelberg: Springer.

Clark, R. E., Kirschner, P. A., & Sweller, J. (2012): Putting Students on the Path to Learning. The Case for Fully Guided Instruction. *AmERIcAN EdUcATOR, 36*(1), 6-11.

Csikszentmihalyi, M. (2010): Das Flow-Erlebnis. Jenseits von Angst und Langeweile: im Tun aufgehen.

Csikszentmihalyi, M., Abuhamdeh, S. & Nakamura, J. (2005): Flow. In A. J. Elliot & C. S. Dweck (Eds.), Handbook of competence and motivation (S. 598–608). New York: Guilford.

Deci, E. L. (1971): Effects of externally mediated rewards on intrinsic motivation. *Journal of personality and Social Psychology*, *18*(1), 105.

Deci, E. L., & Ryan, R. M. (1993): Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. Zeitschrift für Pädagogik, 39(2), 223-238.

Deci, E. L., & Ryan, R. M. (2000): The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry, 11*(4), 227-268.

Deci, E. L., & Ryan, R. M. (2003): Intrinsic Motivation Inventory. IPN: http://selfdeterminationtheory.org/edu/scales/category/5-intrinsic- motivation-inventory, zuletzt geöffnet am 17.08.21

Dresel, M. & Lämmle, L. (2017): Motivation In: Götz, T. (Ed.). (2017): *Emotion, motivation und selbstreguliertes lernen* (Vol. 3481). UTB.

Duit, R. (2010). PIKO-Briefe. IPN Kiel.

Duit, R., Treagust, D., & Widodo, A. (2008): Teaching science for conceptual change: Theory and practice. In S. Vosniadou, Ed., International handbook of research on conceptual change. New York, London: Routledge.

Duit R., und Wodzinski C.T. (2010): Merkmale „guten“ Physikunterrichts. PikoBrief Nr. 4. 1. Aufl. Kiel: IPN https://www.ipn.uni-kiel.de/de/das-ipn/abteilungen/didaktik-der-physik/piko/pikobriefe032010.pdf., zuletzt geöffnet am 17.08.21

Dweck, C. S., & Reppucci, N. D. (1973): Learned helplessness and reinforcement responsibility in children. Journal of Personality ans Social Psychology, 25(1), 109- 116.

Grolnick, W. S., & Ryan, R. M. (1989): Parent styles associated with children's self-regulation and competence in school. *Journal of educational psychology*, *81*(2), 143.

Harackiewicz, J. M. (1979): The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of personality and social psychology*, *37*(8), 1352.

Hopf, M., Wiesner, H., Schecker, H. (Hg.) (2011): Physikdidaktik kompakt. Hallbergmoos: Aulis Verlag.

Jung, W. (1986): Alltagsvorstellungen und das Lernen von Physik und Chemie. Naturwissenschaften im Unterricht – Physik/Chemie

Klingenböck, A.-M. (2016): Ermittlung der intrinsischen Motivation von SchülerInnen der Unterstufe in der Elektrizitätslehre. Wien. Diplomarbeit an der Universität Wien AECC Physik.

Korner, M. (2015): Cross-age Peer Tutoring in Physik. Wien. Doktorarbeit an der Universität Wien AECC Physik.

Krapp, A. (2005): Das Konzept der grundlegenden psychologischen Bedürfnisse. Ein Erklärungsansatz für die positiven Effekte von Wohlbefinden und intrinsischer Motivation im Lehr-Lerngeschehen. *Zeitschrift für Pädagogik*, *51*(5), 626-641.

Lepper, M. R., Henderlong, J. (2000): Turning “play” into “work” and “work” into “play”. In: Intrinsic and Extrinsic Motivation: Elsevier, S. 257–307.

Leonhart, R. (2017):Multivariate Verfahren, Master Psychologie ALU Freiburg, 1. Fachsemester, WS 2016/17; https://www.ipn.uni-kiel.de/de/das-ipn/abteilungen/didaktik-der-physik/piko/pikobriefe032010.pdf, zuletzt geöffnet am 17.08.21

Lüftenegger, M. (2021): mündliches Interview vom 22.02.2021. Wien.

Meissner, W. W. (1981): Internalization in psychoanalysis. *Psychological issues*, *50*, 1-278.

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.

Nolen, S. B., Horn, I. S., & Ward, C. J. (2015): Situating motivation. *Educational Psychologist*, *50*(3), 234-247.

Pieler, J. (2018): Entwicklung der Skala relatedness des IMI zur Bestimmung der intrinsischen Motivation von SchülerInnen. Wien. Diplomarbeit an der Universität Wien AECC Physik.

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.

Prenzel, M., Krapp, A. & Schiefele, H. (1986). Grundzüge einer pädagogischen Interessentheorie. Zeitschrift für Pädagogik, 32, 163–173.

Ross, M. (1975): Salience of reward and intrinsic motivation. *journal of Personality and Social Psychology*, *32*(2), 245.

Ryan, R. M. (1982): Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of personality and social psychology*, *43*(3), 450.

Ryan, R. M., & Deci, E. L. (1999): Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions, 67, 54–67.

Schiefele, U. (1996): Motivation und Lernen mit Texten. Göttingen: Hogrefe.

Schiefele, U., & Köller, O. (1998): Intrinsische und extrinsische Motivation. In *Handwörterbuch Pädagogische Psychologie* (pp. 193-197). Beltz.

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. Diplomarbeit an der Universität Wien AECC Physik.

Riemeier, T. (2012): Moderater Konstruktivismus. In D. Krüger & H. Vogt (Eds.), Theorien fachdidaktischer Forschung.

Rudolph, U. (2003): Motivationspsychologie. Weinheim: Beltz.

Ryan, R. M., & Powelson, C. L. (1991): Autonomy and relatedness as fundamental to motivation and education. *The journal of experimental education*, *60*(1), 49-66.

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.

Weiner, B. (1986): An attributional theory of motivation and emotion. New York: Springer.

Widodo, A., & Duit, R. (2004): Konstruktivistische Sichtweisen vom Lernen und Lehren und die Praxis des Physikunterrichts. Zeitschrift für Didaktik der Naturwissenschaften, 10, 233-255.

Wilde, M., Bätz, K., Kovaleva, A., & Urhahne, D. (2009): Überprüfung einer Kurzskala intrinsischer Motivation (KIM). Zeitschrift für Didaktik der Naturwissenschaften, 15, 31-45.

Wilson, M. (2005): Constructing Measures - An Item Response Modeling Approach. New York: Taylor & Francis.

8. Abbildungsverzeichnis und Tabellenverzeichnis:

Abbildungsverzeichnis:

1. Abbildung: Four building blocks (Wilson, 2005, S. 31) 38
2. Abbildung: Screeplot zur Skala perceived competence 57
3. Abbildung: Histogramm des Merkmals im Vergleich zur Normalverteilung 66
4. Abbildung: Mittelwerte der unterschiedlichen Schulen in einem Histogramm 68
5. Abbildung: Mittelwerte der unterschiedlichen Schulstufen in einem Histogramm 69
6. Abbildung: Mittelwerte der unterschiedlichen Schulformen in einem Histogramm 70

Tabellenverzeichnis:

Tab. 1. Construct Map zur Skala Perceived Competence (vgl. Schmidt, 2017, S. 59) . 45
Tab. 2. Construct Map zur Skala Perceived Competence aus dem Itempanel 45
Tab. 3. Construct Map zur Skala Pressure and Tension (vgl. Korner, 2017) 48
Tab. 4. Construct Map zur Skala Pressure and Tension aus dem Itempanel 48
Tab. 5. Übersicht der Verteilung auf die unterschiedlichen Schulen 53
Tab. 6. Komponentenmatrix der Skala perceived competence 56
Tab. 7. Erklärte Gesamtvarianz der Skala perceived competence 57
Tab. 8. Item-Skala-Statistik der Skala perceived competence 58
Tab. 9. Komponentenmatrix der Skala pressure / tension 58
Tab. 10. Erklärte Gesamtvarianz der Skala pressure and tension 59
Tab. 11. Item-Skala-Statistik der Skala pressure and tension 59
Tab. 12. Korrelationen zwischen den Items der Skalen perceived competence und pressure and tension 60
Tab. 13. Komponentenmatrix der Skala effort / importance 61
Tab. 14. Komponentenmatrix der Skala relatedeness 61
Tab. 15. Komponentenmatrix der Skala interest and enjoyment 62
Tab. 16. Komponentenmatrix der Skala value and usefulness 62
Tab. 17. Komponentenmatrix der Skala perceived competence 63
Tab. 18. Faktorenanalyse der ausgewählten Items der Skalen pres, pco, val und pch 64
Tab. 19. Faktorenanalyse der ausgewählten Items der Skalen int, rel, pres und pco .. 65
Tab. 20. Levene-Test auf Homogenität der Varianzen 67
Tab. 21. Mittelwerte der unterschiedlichen Schulen 67
Tab. 22. Die ANOVA zeigt signifikante Unterschiede zwischen den Schulen 68
Tab. 23. Mittelwerte der unterschiedlichen Schulstufen 69
Tab. 24. Mittelwerte der unterschiedlichen Schulformen 70

9. Anhänge A

9.1. Zusammenfassung (Deutsch)

Mit dieser Masterarbeit wurde ein weiterer Schritt zur Finalisierung des ersten deutschsprachigen Testinstruments, welches die intrinsische Motivation von Lernenden evaluieren kann, beigetragen.

Den Ausgangpunkt der Arbeit stellte das englischsprachige IMI von Deci und Ryan dar. Der Versuch einer Übersetzung ins Deutsche erwies sich als nicht zielführend, weshalb Korner begann die Skalen des IMIs nach der Methode des *Four Building Blocks Approachs* nach Wilson neu zu konstruieren. In Vorarbeiten wurden bereits die Skalen *effort and importance*, *interest and enjoyment*, *value and usefulness*, *perceived choice*, *perceived c*ompetence und (*social) relatedness* entwickelt.

Im Rahmen dieser Arbeit gelang es, trennscharfe stabile Konstrukte zu den Skalen *perceived competence* und *pressure tension* zu entwickeln. Der Ausschluss der Skala *pressure tension* erwies sich als nicht zielführend, da sie mit der *perceived competence* Skala weder signifikant noch wie ursprünglich angenommen stets negativ und hoch korreliert. Darüber hinaus wurden auch die übrigen fünf Subskalen überarbeitet, in ihrem Umfang reduziert und expertenvalidiert. Im Zuge statistischer Analysen konnte gezeigt werden, dass die Skalen bei gleichzeitiger Messung aller Items zueinander trennscharf und stabil sind. Damit konnte das Ziel, einen Beitrag zur Finalisierung eines validen, vor allem aber ökonomischen Messinstruments, welches sowohl klassischen Gütekriterien genügt als auch in der Praxis gut einsetzbar ist, beizutragen, erreicht werden.

9.2. Abstract (English)

With this master thesis a further step was contributed to the finalization of the first German-language test instrument, which can evaluate the intrinsic motivation of learners.

The starting point of the work was the IMI by Deci and Ryan. Attempting to translate it into German did not prove successful, so Korner began to reconstruct the scales of the IMI using Wilson's Four Building Blocks Approach. In preliminary work, the scales effort and importance, interest and enjoyment, value and usefulness, perceived choice, perceived competence and (social) relatedness have already been developed.

In the context of this work, it was possible to develop separable and stable constructs on the scales perceived competence and pressure tension. The exclusion of the pressure tension scale proved not to be purposeful, as the constructs are neither significant nor always negatively and highly correlated as originally assumed. In addition, the other five subscales were also reconstructed, reduced in their scope and validated by experts. In the course of statistical analyses, it was shown that the scales are separable and stable when all items are measured at the same time. Thus, the goal of contributing to the finalization of a valid and, above all, economical measurement instrument that satisfies both classical quality criteria and can be used well in practice was achieved.

10. Anhänge B

10.1.1. Fragebogen lang (41 Items)

Liebe Schülerin, lieber Schüler!

Mit dem Beantworten dieser Fragen hilfst du einem Forscher an der Universität Wien, einen Fragebogen zum Thema **Motivation** zu entwickeln. Versuche die einzelnen Fragen bitte möglichst genau zu beantworten.
Bitte kreuze bei JEDER Frage EINE Möglichkeit an. Die Fragen beziehen sich auf die eben gemachte Partnerarbeit.
Alle Antworten werden anonym behandelt und nicht benotet.

1) Schulform: ..... 2) Schulstufe: .....

3) Unterrichtsfach: ..... 4) Muttersprache: .....

		stimmt völlig	**stimmt eher**	**stimmt teilweise**	**stimmt eher nicht**	**stimmt gar nicht**
1	Meine Fragen werden im Unterricht ernst genommen.	o	o	o	o	o
2	Durch die Aufgabe lerne ich etwas für die Zukunft.	o	o	o	o	o
3	Wenn ich angesprochen werde, sind meine Antworten unsicher.	o	o	o	o	o
4	Ich habe sogar auf etwas Freizeit (oder Pause) verzichtet, um bei dieser Aufgabe gut zu sein.	o	o	o	o	o
5	Ich habe in dieser Unterrichtsstunde gerne mitgemacht.	o	o	o	o	o
6	Ich habe Angst zu versagen.	o	o	o	o	o
7	Mir stehen bei der Bearbeitung einer Aufgabe unterschiedliche Hilfsmittel zur Verfügung.	o	o	o	o	o
8	Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.	o	o	o	o	o
9	Durch das Erlernte kann ich meine FreundInnen beeindrucken.	o	o	o	o	o
10	Ich rede leise, weil ich Angst habe, etwas Falsches zu sagen.	o	o	o	o	o
11	Die Lehrperson gibt mir genügend Zeit zum Nachdenken / Überlegen.	o	o	o	o	o
12	Ich habe weitergearbeitet, bis ich mit dem Ergebnis zufrieden war.	o	o	o	o	o
13	Ich fand das Thema spannend.	o	o	o	o	o
14	Auch wenn ich gut vorbereitet bin, habe ich Angst, dass ich die Aufgabe nicht schaffen kann.	o	o	o	o	o
15	Bei Gruppenarbeiten überlasse ich die Aufgabe den anderen.	o	o	o	o	o

Bitte umblättern!

		stimmt völlig	stimmt eher	stimmt teilweise	stimmt eher nicht	stimmt gar nicht
16	Ich kann selbst entscheiden, wie genau ich die Aufgabe bearbeite.	o	o	o	o	o
17	In meiner Freizeit spreche ich über den Inhalt des Unterrichts mit meinen FreundInnen.	o	o	o	o	o
18	Ich arbeite im Unterricht viel mit.	o	o	o	o	o
19	Durch die Aufgabe lerne ich über aktuelle Themen.	o	o	o	o	o
20	Ich kann zwischen mehreren Themen wählen.	o	o	o	o	o
21	Ich hoffe, es sieht mich niemand und ich muss die Aufgabe nicht behandeln.	o	o	o	o	o
22	Ich habe da viel Energie hineingesteckt.	o	o	o	o	o
23	Die Lehrperson unterstützt mich, meine Sicht / Meinung zu formulieren.	o	o	o	o	o
24	Wenn ich eine Aufgabe bekomme, schaue ich zuerst was die Anderen machen.	o	o	o	o	o
25	Ich freue mich auf den nächsten Unterricht mit diesem Thema.	o	o	o	o	o
26	Die Aufgabe fordert mich und bringt mich zum Nachdenken.	o	o	o	o	o
27	Ich kann wählen, ob ich die Aufgabe alleine oder mit MitschülerInnen bearbeite.	o	o	o	o	o
28	Ich spüre, dass ich mich bei dieser Aufgabe verkrampfe.	o	o	o	o	o
29	Die Aufgabe ist wichtig, um meine beruflichen Aussichten zu verbessern.	o	o	o	o	o
30	Zusammen mit anderen kann ich bessere Ergebnisse erzielen.	o	o	o	o	o
31	Ich kann mir frei einteilen, wann ich eine Aufgabe erledige.	o	o	o	o	o
32	Ich weiß, dass ich Aufgaben aus dem Unterricht gut lösen kann.	o	o	o	o	o
33	Ich finde die Arbeitsaufträge/ Aktivitäten haben Spaß gemacht.	o	o	o	o	o
34	Es war wichtig für mich diese Aufgabe gut zu bewältigen.	o	o	o	o	o

Bitte umblättern!

		stimmt völlig	stimmt eher	stimmt teilweise	stimmt eher nicht	stimmt gar nicht
35	Ich bin bereit, so etwas wieder zu machen.	o	o	o	o	o
36	Ich beantworte Aufgabenstellungen/Fragen schnell.	o	o	o	o	o
37	Ich schaue mich in der Klasse um, ob mir jemand bei der Aufgabe helfen kann.	o	o	o	o	o
38	Die Aufgabe ist realitätsnahe und hat Bezug zu meinem Leben.	o	o	o	o	o
39	Ich kann unterschiedliche Lösungswege wählen.	o	o	o	o	o
40	Ich habe versucht offenen Fragen auf den Grund zu gehen.	o	o	o	o	o
41	Wir hören einander zu.	o	o	o	o	o

Danke vielmals für die Beantwortung aller Fragen!

10.1.2. Fragebogen kurz (34 Items)

Liebe Schülerin, lieber Schüler!

Mit dem Beantworten dieser Fragen hilfst du einem Forscher an der Universität Wien, einen Fragebogen zum Thema **Motivation** zu entwickeln. Versuche die einzelnen Fragen bitte möglichst genau zu beantworten.
Bitte kreuze bei JEDER Frage EINE Möglichkeit an. Die Fragen beziehen sich auf die eben gemachte Partnerarbeit.
Alle Antworten werden anonym behandelt und nicht benotet.

1) Schulform: ..... 2) Schulstufe: .....

3) Unterrichtsfach: ..... 4) Muttersprache: .....

		stimmt völlig	**stimmt eher**	**stimmt teilweise**	**stimmt eher nicht**	**stimmt gar nicht**
1	Meine Fragen werden im Unterricht ernst genommen.	o	o	o	o	o
2	Durch die Aufgabe lerne ich etwas für die Zukunft.	o	o	o	o	o
3	Die Aufgabe ist realitätsnahe und hat Bezug zu meinem Leben.	o	o	o	o	o
4	Ich habe sogar auf etwas Freizeit (oder Pause) verzichtet, um bei dieser Aufgabe gut zu sein.	o	o	o	o	o
5	Ich habe in dieser Unterrichtsstunde gerne mitgemacht.	o	o	o	o	o
6	Ich habe Angst zu versagen.	o	o	o	o	o
7	Mir stehen bei der Bearbeitung einer Aufgabe unterschiedliche Hilfsmittel zur Verfügung.	o	o	o	o	o
8	Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.	o	o	o	o	o
9	Durch das Erlernte kann ich meine FreundInnen beeindrucken.	o	o	o	o	o
10	Ich rede leise, weil ich Angst habe, etwas Falsches zu sagen.	o	o	o	o	o
11	Die Lehrperson gibt mir genügend Zeit zum Nachdenken / Überlegen.	o	o	o	o	o
12	Ich habe weitergearbeitet, bis ich mit dem Ergebnis zufrieden war.	o	o	o	o	o
13	Ich fand das Thema spannend.	o	o	o	o	o
14	Auch wenn ich gut vorbereitet bin, habe ich Angst, dass ich die Aufgabe nicht schaffen kann.	o	o	o	o	o
15	Bei Gruppenarbeiten überlasse ich die Aufgabe den anderen.	o	o	o	o	o

Bitte umblättern!

		stimmt völlig	stimmt eher	stimmt teilweise	stimmt eher nicht	stimmt gar nicht
16	Ich kann unterschiedliche Lösungswege wählen.	o	o	o	o	o
17	In meiner Freizeit spreche ich über den Inhalt des Unterrichts mit meinen FreundInnen.	o	o	o	o	o
18	Ich arbeite im Unterricht viel mit.	o	o	o	o	o
19	Durch die Aufgabe lerne ich über aktuelle Themen.	o	o	o	o	o
20	Ich kann zwischen mehreren Themen wählen.	o	o	o	o	o
21	Ich hoffe, es sieht mich niemand und ich muss die Aufgabe nicht behandeln.	o	o	o	o	o
22	Ich habe da viel Energie hineingesteckt.	o	o	o	o	o
23	Die Lehrperson unterstützt mich, meine Sicht / Meinung zu formulieren.	o	o	o	o	o
24	Wenn ich eine Aufgabe bekomme, schaue ich zuerst was die Anderen machen.	o	o	o	o	o
25	Ich freue mich auf den nächsten Unterricht mit diesem Thema.	o	o	o	o	o
26	Ich habe versucht offenen Fragen auf den Grund zu gehen.	o	o	o	o	o
27	Wir hören einander zu.	o	o	o	o	o
28	Ich spüre, dass ich mich bei dieser Aufgabe verkrampfe.	o	o	o	o	o
29	Die Aufgabe ist wichtig, um meine beruflichen Aussichten zu verbessern.	o	o	o	o	o
30	Wenn ich gefragt werde, fällt mir sofort eine Antwort ein.	o	o	o	o	o
31	Ich kann mir frei einteilen, wann ich eine Aufgabe erledige.	o	o	o	o	o
32	Ich weiß, dass ich Aufgaben aus dem Unterricht gut lösen kann.	o	o	o	o	o
33	Ich finde die Arbeitsaufträge/ Aktivitäten haben Spaß gemacht.	o	o	o	o	o
34	Es war wichtig für mich diese Aufgabe gut zu bewältigen.	o	o	o	o	o

Danke vielmals für die Beantwortung aller Fragen!

10.2.1. Infoblatt zum Motivationsfragebogen lang: Zuordnung der Items

Infoblatt zum Motivationsfragebogen (lang – 41 Items)

Die Folgende Tabelle gibt eine Übersicht über die Itemnummerierung des Motivationsfragebogens (Anhang 10.1).
Dabei ist zu sehen, zu welcher Skala jedes der Items gehört.

Item / Subskala	FB Motivation	Item / Subskala	FB Motivation
eff 4	22	val 13	29
eff 6	12	val 14	9
eff 7	4	pch 2	31
eff 8	40	pch 5	39
eff 9	34	pch 10	16
rel L1	1	pch 11	20
rel L6	23	pch 12	7
rel L15	11	pch 13	27
rel S2	41	pco 3	18
rel S4	8	pco 5	36
rel S16	30	pco 7	32
int 2	35	pco 9	3
int 3	13	pco 10	24
int 5	25	pco 11	15
int 6	5	pres 3	10
int 20	33	pres 4	37
int 21	17	pres 9	21
val 4	38	pres 13	28
val 6	2	pres 14	6
val 7	26	pres 15	14
val 9	19		

10.2.2. Infoblatt zum Motivationsfragebogen kurz: Zuordnung der Items

Infoblatt zum Motivationsfragebogen (kurz – 34 Items)

Die Folgende Tabelle gibt eine Übersicht über die Itemnummerierung des Motivationsfragebogens.
Dabei ist zu sehen, zu welcher Skala jedes der Items gehört.

Item / Subskala	FB Motivation	Item / Subskala	FB Motivation
eff 4	22	val 13	29
eff 6	12	val 14	9
eff 7	4		
eff 8	26	pch 2	31
eff 9	34	pch 5	16
		pch 11	20
rel L1	1	pch 12	7
rel L6	23		
rel L15	11	pco 3	18
rel S2	27	pco 7	32
rel S4	8	pco 10	24
		pco 11	15
int 3	13	pco 12	30
int 5	25		
int 6	5	pres 3	10
int 20	33	pres 9	21
int 21	17	pres 13	28
		pres 14	6
val 4	3	pres 15	14
val 6	2		
val 9	19		

10.3.1. Übersichtsblatt zur Skala effort and importance (eff)

Von den Items der folgenden Skalen wurden jeweils die fett gedruckten behalten.

Übersichtsblatt zur Skala effort / importance (eff)

eff 2: Ich habe mich angestrengt diese Tätigkeit gut zu machen.
eff 3: Ich habe mich sofort auf meine Aufgaben gestürzt.
eff 4: **Ich habe da viel Energie hineingesteckt.**
eff 5: Ich habe bei dieser Aufgabe viel gearbeitet und bin nicht nur herumgesessen.
eff 6: **Ich habe weitergearbeitet, bis ich mit dem Ergebnis zufrieden war.**
eff 7: **Ich habe sogar auf etwas Freizeit (oder Pause) verzichtet, um bei dieser Aufgabe gut zu sein.**
eff 8: **Ich habe versucht offenen Fragen auf den Grund zu gehen.**
eff 9: **Es war wichtig für mich diese Aufgabe gut zu bewältigen.**

Begründung:

eff 2, eff 4 und eff 5 sind inhaltlich ähnlich, deshalb lediglich eines ausgewählt.

eff 3 inhaltlich ähnlich zu Items der Skala interest and enjoy

10.3.2. Übersichtsblatt zur Skala relatedeness (rell)

Übersichtsblatt zur Skala relatedeness (rell)

rel L 1:	Meine Fragen werden im Unterricht ernst genommen und die Lehrperson versucht sie immer zu beantworten.
rel L 3:	Die Lehrperson hört mir zu, wenn ich etwas im Unterricht sage.
rel L 6:	**Die Lehrperson unterstützt mich, meine Sicht / Meinung zu formulieren.**
rel L 10:	Jeder, der etwas sagen will, kommt zu Wort.
rel L 12:	Ich werde von der Lehrperson ernst genommen und respektvoll behandelt.

Neu:

rel L 1:	**Meine Fragen werden im Unterricht ernst genommen.** (gekürzt)
rel L 15:	**Die Lehrperson gibt mir genügend Zeit zum Nachdenken / Überlegen.**

rel S 2:	**Wir hören einander zu.**
rel S 3:	Wir helfen einander in der Klasse.
rel S 4:	**Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind.**
rel S 5:	Ich werde von meinen Mitschülern ausgelacht, wenn ich eine Frage stelle oder meine Meinung äußere.
rel S 6:	Meine Mitschüler behandeln mich mit Respekt.

Neu:

rel S 16:	**Zusammen mit anderen kann ich bessere Ergebnisse erzielen.**

Begründung:

rel L1 und rel L12 sind inhaltlich ähnlich.

rel L3 wurde gestrichen, da bereits durch rel S 2 mit abgefragt.

rel L 10 aus statistisch mathematischen Gesichtspunkten gestrichen.

rel S 3 neu formuliert zu rel S 16 .

rel S 5 aus statistisch mathematischen Gesichtspunkten gestrichen.
(nicht trennscharf und Cronbach Alpha besser, wenn weggelassen)

rel S 6 zu allgemein formuliert und ähnlich zu rel S 4.

10.3.3. Übersichtsblatt zur Skala interest and enjoyment (int)

Übersichtsblatt zur Skala interest / enjoyment (int)

int 1: Ich war sehr gespannt auf die Ergebnisse des Arbeitsblatts.
int 2: **Ich bin bereit so etwas wieder zu machen.**
int 3: **Ich fand das Thema spannend.**
int 4: Ich würde das behandelte Thema als interessant bezeichnen.
int 5: **Ich freue mich auf den nächsten Unterricht mit diesem Thema.**
int 6: **Ich habe bei dieser Tätigkeit gerne mitgemacht.**

Neu:
int 20: Ich finde die Arbeitsaufträge/ Aktivitäten haben Spaß gemacht.
int 21: In meiner Freizeit spreche ich über den Inhalt des Unterrichts mit meinen FreundInnen.

Begründung:

int 1 schlechte Faktorladung und bezieht sich auf eine spezifische Tätigkeit.
Hier könnte man eine allgemeinere Formulierung wie „(...) Ergebnisse dieser Tätigkeit“ vornehmen.

int 4 inhaltlich ähnlich zu int 3, deshalb gestrichen.

10.3.4. Übersichtsblatt zur Skala value and usefulness (val)

Übersichtsblatt zur Skala value / usefulness (val)

val 1: Die Aufgabe lässt mich Erfahrungen sammeln.
val 2: Die Aufgabe erleichtert mir das Verständnis des Themas.
val 3: Durch die Aufgabe lerne ich etwas dazu.
val 4: **Die Aufgabe ist realitätsnahe und hat Bezug zu meinem Leben.**
val 5: Die Aufgabe ist wichtig für die Arbeitswelt.
val 6: **Durch die Aufgabe lerne ich etwas für die Zukunft.**
val 7: **Die Aufgabe fordert mich und bringt mich zum Nachdenken.**
val 8: Die Bearbeitung der Aufgabe ist abwechslungsreich.
val 9: **Durch die Aufgabe lerne ich über aktuelle Themen.**
val 10: Ich denke, die Aufgabe ist nützlich für Tests oder Schularbeiten.
val 11: Ich finde die Aufgabe sinnlos. (R)
val 12: Ich verstehe nicht, warum ich solche Aufgaben lösen muss. (R)

Neu
val 13: **Die Aufgabe ist wichtig, um meine beruflichen Aussichten zu verbessern.**
val 14: **Durch das Erlernte kann ich meine FreundInnen beeindrucken.**

Begründung:

val 1 ist für SchülerInnen schwer verständlich, da „Erfahrungen sammeln“ abstrakt ist.

val 2 schlechte Faktorladung und inhaltlich nicht passend.

val 3 weggelassen, da ähnlich zu val 6.

val 5 neu formuliert (val 13) unter Einbeziehung einer PISA Formulierung:
(...) berufliche Aussichten verbessern (...)

val 8 schlechte Faktorladung und inhaltlich nicht eindeutig der Skala zuordenbar.

Ist val 10 ein Indikator für intrinsische Motivation?

val 11 schlechte Faktorladung, zu allgemein formuliert und deshalb nicht greifbar.

val 12 schlechte Faktorladung, negativ gepolte Items können von SchülerInnen meist nicht gemäß ihrer Zustimmung oder Ablehnung sinnvoll beantwortet werden.

10.3.5. Übersichtsblatt zur Skala perceived choice (pch)

Übersichtsblatt zur Skala perceived choice (pch)

pch 1: Ich kann mir das Thema, das ich bearbeiten soll selbst aussuchen.
pch 2: **Ich kann mir frei einteilen, wann ich eine Aufgabe erledige.**
pch 3: Ich entscheide selbst wie viel ich arbeite.
pch 4: Ich kann bei Gruppenarbeiten selbst aussuchen, mit wem ich arbeite.
pch 5: **Ich kann unterschiedliche Lösungswege wählen**.
pch 6: Ich kann den Zeitpunkt selbst bestimmen, wann ich die Aufgabe abgebe.
pch 7: Ich habe keinen Spielraum bei der Bearbeitung der Aufgabe. (R)
pch 8: Ich kann selbst keine Entscheidungen treffen, wie ich die Aufgaben bearbeite.
pch 9: Ich kann selbst bestimmen, welche Hilfsmittel ich verwenden möchte.
pch 10: **Ich kann selbst entscheiden, wie genau ich die Aufgabe bearbeite.**

Neu
pch 11: **Ich kann zwischen mehreren Themen wählen.**
pch 12: **Mir stehen bei der Bearbeitung einer Aufgabe unterschiedliche Hilfsmittel zur Verfügung.**
pch 13: **Ich kann wählen, ob ich die Aufgabe alleine oder mit MitschülerInnen bearbeite.**

Begründung:

pch 1 für den hier weiterentwickelten Fragebogen ist die uneingeschränkte Wahl des Themas nicht relevant, deshalb neu formuliert mit Item pch 11.

pch 3 das kann man immer, hat meist jedoch schlechte Note als Konsequenz – Aspekt besser mit pch 10 abgefragt.

pch 4 zeigt schlechte Faktorladung und nur in manchen Kontexten für Fragebogen relevant, deshalb neu formuliert mit Item pch 13.

pch 6 zielt inhaltlich auf ähnlichen Aspekt wie pch 2 ab, welches jedoch für viele Situationen passender scheint.

pch 7 schlechte Faktorladung, unpräzise und lässt Fragen offen: keinen Spielraum hinsichtlich was? Zeit, Genauigkeit, Themenfreiheit, hinsichtlich des Lösungsweges etc.

pch 8 wird bereits durch die verbleibenden Items gut abgedeckt.

pch 9 neu formuliert in pch 12.

10.4. Übersichtstabelle der ausgewählten Items nach Skalen VOR der Testung

Effort / importance (eff)	Relatedeness (rel)	Interest / enjoyment (int)	Value / usefulness (val)	Perceived choice (pch)	Perceived Competence (pco)	Pressure/Tension (pres)
Ich habe da viel Energie hineingesteckt. (eff 4)	Meine Fragen werden im Unterricht ernst genommen. (L 1) Neu gekürzt	Ich bin bereit, so etwas wieder zu machen. (int 2)	Die Aufgabe ist realitätsnahe und hat Bezug zu meinem Leben. (val 4)	Ich kann mir frei einteilen, wann ich eine Aufgabe erledige. (pch 2)	Ich arbeite im Unterricht viel mit. (pco 3)	Ich rede leise, weil ich Angst habe, etwas Falsches zu sagen. (pres 3)
Ich habe weitergearbeitet, bis ich mit dem Ergebnis zufrieden war. (eff 6)	Die Lehrperson unterstützt mich, meine Sicht / Meinung zu formulieren. (L 6)	Ich fand das Thema spannend. (int 3)	Durch die Aufgabe lerne ich etwas für die Zukunft. (val 6)	Ich kann unterschiedliche Lösungswege wählen. (pch 5)	Ich beantworte Aufgabenstellungen/Frag en schnell. (pco 5)	Ich schaue mich in der Klasse um, ob mir jemand bei der Aufgabe helfen kann. (pres 4)
Ich habe sogar auf etwas Freizeit (oder Pause) verzichtet, um bei dieser Aufgabe gut zu sein. (eff 7)	Die Lehrperson gibt mir genügend Zeit zum Nachdenken / Überlegen. (L 15)	Ich freue mich auf den nächsten Unterricht mit diesem Thema. (int 5)	Die Aufgabe fordert mich und bringt mich zum Nachdenken. (val 7)	Ich kann selbst entscheiden, wie genau ich die Aufgabe bearbeite. (pch 10)	Ich weiß, dass ich Aufgaben aus dem Unterricht gut lösen kann. (pco 7)	Ich hoffe, es sieht mich niemand und ich muss die Aufgabe nicht behandeln. (pres 9)
Ich habe versucht offenen Fragen auf den Grund zu gehen. (eff 8)	Wir hören einander zu. (S 2)	Ich habe in dieser Unterrichtsstunde gerne mitgemacht. (int 6)	Durch die Aufgabe lerne ich über aktuelle Themen. (val 9)	Ich kann zwischen mehreren Themen wählen. (pch 11) Neu	Wenn ich angesprochen werde, sind meine Antworten unsicher. (pco 9)	Ich spüre, dass ich mich bei dieser Aufgabe verkrampfe. (pres 13) Neu
Es war wichtig für mich diese Aufgabe gut zu bewältigen. (eff 9)	Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind. (S 4)	Ich finde die Arbeitsaufträge/ Aktivitäten haben Spaß gemacht. (int 20) Neu	Die Aufgabe ist wichtig, um meine beruflichen Aussichten zu verbessern. (val 13) Neu	Mir stehen bei der Bearbeitung einer Aufgabe unterschiedliche Hilfsmittel zur Verfügung. (pch 12) Neu (aus pres 9)	Wenn ich eine Aufgabe bekomme, schaue ich zuerst was die Anderen machen. (pco 10)	Ich habe Angst zu versagen. (pres 14) Neu gekürzt
	Zusammen mit anderen kann ich bessere Ergebnisse erzielen. (S 16)	In meiner Freizeit spreche ich über den Inhalt des Unterrichts mit meinen FreundInnen. (int 21) Neu	Durch das Erlernte kann ich meine FreundInnen beeindrucken. (val 14) Neu	Ich kann wählen, ob ich die Aufgabe alleine oder mit MitschülerInnen bearbeite. (pch 13) Neu	Bei Gruppenarbeiten überlasse ich die Aufgabe den anderen. (pco 11)	Auch wenn ich gut vorbereitet bin, habe ich Angst, dass ich die Aufgabe nicht schaffen kann. (pres 15) Neu (aus pres 1 + pisa)

Itemanzahl: 41

Pisa: https://www.pedocs.de/volltexte/2019/17449/pdf/Mang_et_al_2019_PISA_2015_Skalenhandbuch.pdf

10.5. Übersichtstabelle der verbliebenen Items NACH der Analyse (34 Items)

Effort / importance (eff)	Relatedeness (rel)	Interest / enjoyment (int)	Value / usefulness (val)	Perceived choice (pch)	Perceived Competence (pco)	Pressure/Tension (pres)
Ich habe da viel Energie hineingesteckt. (eff 4) Ich habe weitergearbeitet, bis ich mit dem Ergebnis zufrieden war. (eff 6) Ich habe sogar auf etwas Freizeit (oder Pause) verzichtet, um bei dieser Aufgabe gut zu sein. (eff 7) Ich habe versucht offenen Fragen auf den Grund zu gehen. (eff 8) Es war wichtig für mich diese Aufgabe gut zu bewältigen. (eff 9)	Meine Fragen werden im Unterricht ernst genommen. (L 1) Neu gekürzt Die Lehrperson unterstützt mich, meine Sicht / Meinung zu formulieren. (L 6) Die Lehrperson gibt mir genügend Zeit zum Nachdenken / Überlegen. (L 15) Wir hören einander zu. (S 2) Wir akzeptieren einander in der Klasse, auch wenn wir unterschiedlich sind. (S 4)	Ich fand das Thema spannend. (int 3) Ich freue mich auf den nächsten Unterricht mit diesem Thema. (int 5) Ich habe in dieser Unterrichtsstunde gerne mitgemacht. (int 6) Ich finde die Arbeitsaufträge/ Aktivitäten haben Spaß gemacht. (int 20) Neu In meiner Freizeit spreche ich über den Inhalt des Unterrichts mit meinen FreundInnen. (int 21) Neu	Die Aufgabe ist realitätsnahe und hat Bezug zu meinem Leben. (val 4) Durch die Aufgabe lerne ich etwas für die Zukunft. (val 6) Durch die Aufgabe lerne ich über aktuelle Themen. (val 9) Die Aufgabe ist wichtig, um meine beruflichen Aussichten zu verbessern. (val 13) Neu Durch das Erlernte kann ich meine FreundInnen beeindrucken. (val 14) Neu	Ich kann mir frei einteilen, wann ich eine Aufgabe erledige. (pch 2) Ich kann unterschiedliche Lösungswege wählen. (pch 5) Ich kann zwischen mehreren Themen wählen. (pch 11) Neu Mir stehen bei der Bearbeitung einer Aufgabe unterschiedliche Hilfsmittel zur Verfügung. (pch 12) Neu (aus pres 9)	Ich arbeite im Unterricht viel mit. (pco 3) Ich weiß, dass ich Aufgaben aus dem Unterricht gut lösen kann. (pco 7) Wenn ich eine Aufgabe bekomme, schaue ich zuerst was die Anderen machen. (pco 10) Bei Gruppenarbeiten überlasse ich die Aufgabe den anderen. (pco 11) Wenn ich gefragt werde, fällt mir sofort eine Antwort ein. (pco 12) wurde noch nicht getestet. Neu (aus pco 5)	Ich rede leise, weil ich Angst habe, etwas Falsches zu sagen. (pres 3) Ich hoffe, es sieht mich niemand und ich muss die Aufgabe nicht behandeln. (pres 9) Ich spüre, dass ich mich bei dieser Aufgabe verkrampfe. (pres 13) Neu Ich habe Angst zu versagen. (pres 14) Neu gekürzt Auch wenn ich gut vorbereitet bin, habe ich Angst, dass ich die Aufgabe nicht schaffen kann. (pres 15) Neu (aus pres 1 + pisa)

Itemanzahl: 33 verbleibende +1 neu konstruiertes Item

Pisa: https://www.pedocs.de/volltexte/2019/17449/pdf/Mang_et_al_2019_PISA_2015_Skalenhandbuch.pdf