



universität
wien

DIPLOMARBEIT

Titel der Diplomarbeit

Schlussfolgerndes Denken im Zusammenhang mit dem Geschlechterstereotyp bei
fortgeschrittenen Studierenden der Psychologie und Naturwissenschaften unter
Berücksichtigung einer latenten Genderinformation in spezifischen Items

Verfasserin

Petra Niemandsfreund

Angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Wien, 2015

Studienkennzahl lt. Studienblatt: A 298

Studienrichtung lt. Studienblatt: Psychologie

Betreuerin: Univ.-Prof. DDr. Mag. Christiane Spiel

Für meine Eltern.

Vielen Dank auch für die unverzichtbare und wertvolle Unterstützung an meine Geschwister
Barbara und Andreas sowie meinen Freunden Manuela, Alice, Florian und Bernd.

INHALTSVERZEICHNIS

I.	EINLEITUNG	11
II.	THEORIE	15
1.	Schlussfolgerndes Denken als bedeutende Fähigkeit für naturwissenschaftliche Studienrichtungen	15
1.1.	Definition schlussfolgernden Denkens	15
1.2.	Piaget und die kognitive Entwicklung	17
1.3.	Vier Stufen der kognitiven Entwicklung nach Jean Piaget.....	18
1.3.1.	Sensomotorisches Stadium (Geburt bis 2 Jahre)	19
1.3.2.	Präoperationales Stadium (2 bis 7 Jahre)	19
1.3.3.	Konkret-operationales Stadium (7 bis 12 Jahre)	20
1.3.4.	Formal-operationales Stadium (ab 12 Jahren).....	21
1.4.	Kritik an Piagets Theorie	22
2.	Übergang vom konkret-operationalen zum formal-operationalen Denken.....	23
2.1.	Syllogismen.....	24
2.2.	Entwicklungspsychologische Ergebnisse zur Schlussfolgerung mit Syllogismen	26
3.	Entwicklungsfortschritt	27
3.1.	Leistungsabfall bei validen Formen.....	27
3.2.	Effekte von Moderatorvariablen auf die Leistung	28
4.	Geschlechtsunterschiede in kognitiven Fähigkeiten	29
4.1.	Geschlechterstereotype	30
4.2.	Theorien zu Geschlechterstereotypen	31
4.2.1.	Theorie der sozialen Rollen	32
4.2.2.	Das Stereotypinhaltsmodell	33
4.3.	Der Einfluss stereotyper Erwartungen	34
4.3.1.	Self-fulfilling prophecy oder der Pygmalion-Effekt	34
4.3.2.	Stereotype – Threat.....	35

5.	Einfluss von Note und Fähigkeitsselbstkonzept auf die Leistung im schlussfolgernden Denken und geschlechtsspezifische Differenzen.....	36
5.1.	Zusammenhang von Note und Leistung	36
5.2.	Fähigkeitsselbstkonzept	37
5.2.1.	Fähigkeitsselbstkonzept und Leistung.....	38
5.2.2.	Geschlechtsspezifische Leistungsunterschiede und Fähigkeitsselbstkonzept.....	39
6.	Zielsetzung vorliegender Arbeit und Forschungsfragen	41
6.1.	Zusammenhang der SDV-Testleistung mit Studienrichtung, Geschlecht und Studienleistung	42
6.2.	Kompetenz im schlussfolgernden Denken bei Studierenden im zweiten Studienabschnitt bzw. Masterstudiengang	43
6.3.	Zusammenhang von Genderinformation mit Lösungswahrscheinlichkeit	44
6.4.	Zusammenhang von Fähigkeitsselbstkonzept mit der Testleistung.....	45
6.5.	Übersicht der Forschungshypothesen	45
III.	EMPIRIE	47
7.	Methode und Verfahren	47
7.1.	Untersuchungsdesign und Stichprobe.....	47
7.2.	Erhebungsinstrument	48
7.2.1.	SDV	48
7.2.2.	Erweiterung um ein Genderitem.....	50
7.3.	Untersuchungsdurchführung.....	50
7.4.	Spezifisches Auswertungsverfahren	51
7.5.	Stichprobe	52
8.	Ergebnisse	54
8.1.	Unterschiede im schlussfolgernden Denken bezüglich Studienrichtung und Geschlecht	54
8.2.	Zusammenhang der Studienleistung mit der Lösungswahrscheinlichkeit im SDV....	55
8.3.	Extraktion der Lösungsklassen	56

8.3.1. Lösungswahrscheinlichkeit in Abhängigkeit der WINMIRA – Klassen	58
8.3.2. Lösungswahrscheinlichkeit SDV unter Berücksichtigung der Klassenzugehörigkeit und Negation.....	63
8.4. Zusammenhang der Genderinformation im Genderitem mit der Testleistung	67
8.5. Einfluss des Fähigkeitsselbstkonzepts weiblicher Studierender auf das Lösen der Genderitems	71
9. Diskussion	73
10. Zusammenfassung	79
IV. LITERATURVERZEICHNIS	81
V. ANHANG.....	90

Tabellenverzeichnis

Tabelle 1: <i>Aufbau der vier Argumentformen</i>	25
Tabelle 2: <i>Vierfeldertafel zu Geschlecht und Studienrichtung</i>	53
Tabelle 3: <i>Lösungswahrscheinlichkeit im SDV in Abhängigkeit der Erstsprache</i>	54
Tabelle 4: <i>Kennwerte SDV Leistung in Abhängigkeit von Studienrichtung und Geschlecht</i>	55
Tabelle 5: <i>BIC-Werte der Mixed Rasch-Modelle mit und ohne Genderitems mit unterschiedlicher Klassenanzahl</i>	57
Tabelle 6: <i>Mittlere Zuordnungswahrscheinlichkeit gereiht nach Klasse mit Genderitems</i>	57
Tabelle 7: <i>Mittlere Lösungswahrscheinlichkeiten Klasse 1</i>	58
Tabelle 8: <i>Mittlere Lösungswahrscheinlichkeiten Klasse 2</i>	58
Tabelle 9: <i>Mittlere Lösungswahrscheinlichkeiten Klasse 3</i>	58
Tabelle 10: <i>Mittlere Lösungswahrscheinlichkeiten Klasse 4</i>	58
Tabelle 11: <i>Mittlere Lösungswahrscheinlichkeiten Klasse 5</i>	59
Tabelle 12: <i>Lösungswahrscheinlichkeiten (M, SD) in den Argument- und Inhaltsbedingungen je Klasse</i>	62
Tabelle 13: <i>Lösungswahrscheinlichkeit der vier Argumentformen (n = 391)</i>	62
Tabelle 14: <i>Deskriptivstatistische Kennwerte zu den mittleren Lösungswahrscheinlichkeiten</i>	63
Tabelle 15: <i>Kennwerte der Differenz der Lösungswahrscheinlichkeit von unnegiert zu negiert mit Prüfgrößen</i>	64
Tabelle 16: <i>Kontingenztafel mit Häufigkeiten und Anteilswerten der Lösungskategorie je Studienrichtung</i>	65
Tabelle 17: <i>Kontingenztafel mit Häufigkeiten und Anteilswerten der Lösungskategorie je Studienrichtung und Geschlecht</i>	66
Tabelle 18: <i>Kodierung der Items</i>	67
Tabelle 19: <i>Item 1, Genderitem nach Geschlecht</i>	68
Tabelle 20: <i>Item 11, Genderitem nach Geschlecht</i>	69
Tabelle 21: <i>Item 21, Genderitem nach Geschlecht</i>	70
Tabelle 22: <i>Mittelwerte für beide Testversionen A und B mit 35 und 32 Items</i>	71
Tabelle 23: <i>SDV Lösungswahrscheinlichkeiten, korrigierte Itemtrennschärfen und Cronbach's Alpha bei Auslassung des entsprechenden Items (k = 35, n = 391)</i>	91

Abbildungsverzeichnis

<i>Abbildung 1.</i> Sechs Beispielitems konditionaler Syllogismen aus dem ursprünglichen SDV.....	48
<i>Abbildung 2.</i> Neue hinzugefügte SDV-Items.....	49
<i>Abbildung 3.</i> Bivariates Streudiagramm für Studienleistung und Lösungswahrscheinlichkeit im SDV unter Berücksichtigung der beiden Studienrichtungen.	56
<i>Abbildung 4.</i> Mittlere Lösungswahrscheinlichkeiten der 35 Items getrennt über die vier Argument- und drei Inhaltsformen pro Klasse.....	59
<i>Abbildung 5.</i> Lösungswahrscheinlichkeiten SDV in Abhängigkeit von Lösungsklasse und Negation.	65
<i>Abbildung 6.</i> Selbsteinschätzung in Abhängigkeit von Erhebungszeitpunkt und Geschlecht.....	72

I. EINLEITUNG

Um sich in unserer Gesellschaft zurechtzufinden, ist es relevant, logische Schlüsse ziehen zu können. Zudem spielt schlussfolgerndes Denken eine zentrale Rolle bei der Gewinnung neuer Erkenntnisse (Bortz & Döring, 2006). Es ist beispielsweise die Grundlage, Konsequenzen von Annahmen und Hypothesen zu bestimmen, Instruktionen, Regeln und allgemeine Gesetzmäßigkeiten zu verstehen sowie Aussagen zu prüfen. Ohne logische Schlussfolgerungen sind Wissenschaft, Technologie, Gesetze, soziale Grundsätze und Kultur nicht vorstellbar (Johnson-Laird & Byrne, 1991).

Schlussfolgerndes Denken stellt demnach eine zentrale intellektuelle Fähigkeit dar (Johnson-Laird & Byrne, 1991) und ist zudem eine notwendige Kompetenz für das erfolgreiche Absolvieren naturwissenschaftlicher Studienrichtungen (OECD, 2014; Spiel & Glück, 2008). In zahlreichen Untersuchungen zeigt sich jedoch, dass sich Frauen und Männer im schlussfolgernden Denken dahingehend unterscheiden, dass Männer höhere Fähigkeiten aufweisen (z.B. Hyde, 1990; Hyde, Fennema, & Lamon, 1990; Kiper & Kattmann, 2003). In naturwissenschaftlichen Studienrichtungen liegt der Anteil männlicher Studierender nach wie vor deutlich über jenem weiblicher Studierender. Im österreichischen Frauenbericht 2010 (Gärtner, Schwabe, & Sommer-Binder, 2010) wurde erhoben, dass im Studienjahr 2007/2008 der Frauenanteil für die Studienrichtung Psychologie 77,6% und für Physik 23,1% betrug. Ein ähnliches Bild zeigte sich im Wintersemester 2011 an der Universität Wien und der Technischen Universität Wien. In den Studienrichtungen Mathematik, Physik und Chemie waren insgesamt 71,40% (N = 1234) Männer und 28,59% Frauen (N = 494) inskribiert; in der Studienrichtung Psychologie waren es 23,39% Männer (N = 838) und 76,61% Frauen (N = 2744) (Bundesministerium für Wissenschaft Forschung und Wirtschaft, 2011; Technische Universität Wien, 2011). Diese Daten spiegeln die klassischen Verteilungsunterschiede wider, wonach Männer häufiger in „konventionellen“ Studienrichtungen (Mathematik, Chemie, Physik) vorzufinden sind als Frauen.

Diese in den naturwissenschaftlichen Studienrichtungen vorzufindenden Genderdifferenzen (durch soziokulturelle Interaktionen konstruiertes Geschlecht) beruhen unter anderem auf konventionellen Einstellungen und Verhaltensweisen, die als jeweils charakteristisch und adäquat für Frauen und Männer gelten (Caplan, Crawford, Hyde, & Richardson, 1997). Neben soziokulturellen Geschlechtsunterschieden werden in der Forschung auch kognitive Differenzen zwischen Frauen und Männern untersucht. Diese Ungleichverteilung von Frauen und Männern in

naturwissenschaftlichen Studienrichtungen könnte demnach als Folge vermittelter soziokultureller und kognitiver Geschlechterstereotype aufgefasst werden (Rost, 2010). Ob und inwiefern geschlechtsspezifische Unterschiede im schlussfolgernden Denken bei fortgeschrittenen Studierenden in den Studienrichtungen Physik, Mathematik, Chemie und Psychologie vorhanden sind, soll in vorliegender Untersuchung erhoben werden.

Zur Erfassung schlussfolgernden Denkens werden gewöhnlich Syllogismen (verbal vorgegebene logische Argumente, bei denen aus einer gegebenen Prämisse eine Konklusion zu ziehen ist) herangezogen (Beller & Spada, 1999). In den Untersuchungen von Spiel, Glück & Gößler (2001, 2004) wurde mittels des Syllogismustests SDV („Schlussfolgerndes Denken Verbal“; Spiel, Glück, und Gössler (2003)) die Kompetenz des schlussfolgernden Denkens im Sinne der Stadientheorie Piagets von Schülerinnen und Schülern zwischen der 7. und 12. Schulstufe erhoben. Mittels Einsatz des Mixed Rasch-Modells konnten qualitative Kompetenzunterschiede im schlussfolgernden Denken ermittelt werden. Gemäß Piaget (Piaget, 1976b; Piaget & Inhelder, 1977) vollzieht sich die Entwicklung des deduktiven Denkens in vier Stadien, wobei das formal-operationale Stadium alle möglichen Denkprozesse impliziert und somit als das höchste zu erreichende Stadium gilt. Der Übergang vom konkret-operationalen Stadium zum formal-operationalen Stadium bildet die letzte kognitive Entwicklungsstufe und sollte folglich in der Adoleszenz abgeschlossen sein. Spiel, et al. (2004) identifizierten zwischen dem konkreten und formalen Stadium zwei Übergangsstadien. Klopff (2011) und Eder (2012) ermittelten in ihren erweiterten Untersuchungen bei Studienanfängerinnen und Studienanfängern der Studienrichtungen Physik, Chemie, Mathematik und Psychologie drei Übergangsstadien. Ein Stadium der formalen Operationen konnte nicht bestimmt werden.

Ziel vorliegender Arbeit ist, fortgeschrittene Studierende der Studienrichtungen Physik, Chemie, Mathematik und Psychologie, im Sinne der Stadientheorie Piagets, hinsichtlich ihrer Performanz im deduktiven Schlussfolgern zu untersuchen. Es ist von Interesse, ob das Stadium der formalen Operationen erreicht wird. Basis hierzu bilden die Untersuchungen von Spiel, et al. (2001, 2004), Klopff (2011) und Eder (2012), welche hier erweitert werden sollen. In der gegenständlichen Untersuchung wird zur Erhebung des schlussfolgernden Denkens, ebenso wie in den Vorgängerstudien, der Test „Schlussfolgerndes Denken Verbal“ (SDV, Spiel, et al., 2003) herangezogen. Dieser wurde für die vorliegende Untersuchung um drei spezifische Items, sog. Genderitems, welche zur Erfassung von Geschlechterstereotypen eine latente Genderinformation beinhalten, erweitert. Es wird geprüft, ob dermaßen transportierte latente

Geschlechterstereotypen schließlich unterschiedliche Lösungswahrscheinlichkeiten bei weiblichen und männlichen Studierenden im SDV bewirken. Ebenso wird die Auswirkung des Fähigkeitsselbstkonzeptes der Studierenden auf die SDV-Testleistung untersucht.

Im ersten Abschnitt des Theorieteils werden zunächst allgemeine Definitionen und ein Überblick zum schlussfolgernden Denken gegeben sowie die kognitive Entwicklung nach Piaget beschrieben. Anschließend folgt im zweiten Abschnitt eine detaillierte Beschreibung des Übergangs vom konkret-operationalen zum formal-operationalen Stadium. Der Aufbau und Einsatz von Syllogismen wird erläutert und entwicklungspsychologische Ergebnisse zum schlussfolgernden Denken mittels Einsatz von Syllogismen erörtert. Der dritte Abschnitt analysiert den kognitiven Entwicklungsfortschritt sowie den Einfluss von Moderatorvariablen, welche die Performanz im deduktiven Denken beeinflussen können. Im Abschnitt vier wird auf Geschlechtsunterschiede in kognitiven Fähigkeiten eingegangen. Darüber hinaus wird der Einfluss stereotyper Erwartungen auf die Leistung kognitiver Fähigkeiten in naturwissenschaftlichen Fächern beschrieben. Weitere Einflussfaktoren auf die Leistung im schlussfolgernden Denken und geschlechtsspezifische Differenzen werden im fünften Abschnitt veranschaulicht. Anschließend folgt die Ausarbeitung der einzelnen Forschungsfragen. Der empirische Teil vorliegender Arbeit befasst sich in Abschnitt sieben mit den herangezogenen Methoden und Verfahren sowie der Beschreibung der Untersuchungsdurchführung. In Abschnitt acht folgt die Ergebnisdarstellung und wird im folgenden neunten Abschnitt diskutiert. Abschließend werden im Abschnitt zehn die wesentlichen Erkenntnisse und Ergebnisse zusammengefasst.

Die vorliegende empirische Untersuchung stellt eine Erweiterung der Arbeiten von Klopff (2011) und Eder (2012) dar. Die Datenerhebung wurde in Zusammenarbeit mit einer weiteren Diplomandin, Sabine Sinzinger, durchgeführt.

II. THEORIE

1. Schlussfolgerndes Denken als bedeutende Fähigkeit für naturwissenschaftliche Studienrichtungen

Die Fähigkeit des schlussfolgernden Denkens ist eine notwendige Voraussetzung, um ein naturwissenschaftliches Studium erfolgreich zu absolvieren (OECD, 2014). Bevor im zweiten Abschnitt dieser Arbeit auf den Zusammenhang von Geschlechterstereotypen auf die Leistung im schlussfolgernden Denken eingegangen wird, soll zunächst eine Definition über den Begriff des schlussfolgernden Denkens gegeben und anschließend die kognitive Entwicklungstheorie von Jean Piaget beschrieben werden.

1.1. Definition schlussfolgernden Denkens

Systematisches Hypothesenprüfen und wissenschaftliches Ableiten bedürfen der Fähigkeit schlussfolgernder Denkleistungen (Gauffroy & Barrouillet, 2011). Hinter dieser Fähigkeit steht ein Denkprozess, der von einer Reihe an Aussagen zu einer neuen führt (Johnson-Laird, 1983). In der Literatur werden überwiegend zwei Formen schlussfolgernden Denkens erörtert. Zum einen *induktives* Schlussfolgern, das die Ableitung einer spezifischen Erfahrung zu einer allgemeinen Gesetzmäßigkeit und die Anwendung auf neue Ergebnisse erlaubt. Derartige Verallgemeinerungen müssen jedoch nicht zwingend wahr sein. (Waldmann & Weinert, 1990). In Einzel- bzw. Subtests wird induktives Denken vor allem durch Zahlen- oder Wortanalogien, Matrizenaufgaben oder Ergänzungen von Buchstaben- oder Zahlenreihen gemessen (Beckmann & Guthke, 1999). Zum anderen *deduktives* Schlussfolgern, das uns erlaubt Folgerungen (Konklusionen) aus vorhergehenden Erkenntnissen (Prämissen) zu ziehen. Diese Konklusionen regen Bemühungen zur Problemlösung und Entscheidungsfindung an und werden letztlich gezogen um Ziele zu erreichen (Leighton, 2004). Nach Kubinger (2006) etwa ist schlussfolgerndes Denken (*deductive reasoning*) „die Fähigkeit, Gesetzmäßigkeiten oder logisch zwingende Zusammenhänge erkennen und zweckentsprechend verwerten zu können.“ (Kubinger, 2006, S. 198).

Im täglichen Leben hängen zahlreiche Situationen mit der Fähigkeit zusammen, deduktiv zu denken. Trotzdem fällt es uns teilweise schwer, Deduktionen zu lösen, was in weiterer Folge zu Fehlern führt. Individuen unterscheiden sich deutlich in diesem Vermögen. Menschen mit geringeren Fähigkeiten im deduktiven Denken neigen im täglichen Leben eher dazu, Fehlschlüsse zu ziehen (Johnson-Laird, 1999). Deduktives Schlussfolgern wird auch als eine bedeutende Grundlage der formalen Wissenschaft aufgefasst, d.h. es werden Strukturen untersucht, welche von der Wirklichkeit losgelöst und unabhängig (z.B. Mathematik, Sprachwissenschaften, Logik), aber auch für das tägliche Denken von Relevanz sind (Reverberi et al., 2010). Aufgrund der Bedeutung des deduktiven Denkens für vorliegende Arbeit wird im Folgenden speziell auf dieses eingegangen.

Bei der Deduktion handelt es sich um einen systematischen Prozess, dessen Ziel es ist, valide (gültige) Folgerungen aus einer Reihe von Prämissen (Aussagen) zu ziehen. Dabei werden die Prämissen als wahr angesehen, und es wird rückgeschlossen, welche Konklusionen (Folgerungen), wenn überhaupt, folgen (Bara, Bucciarelli, & Lombardo, 2001). Dieser Denkprozess basiert auf Gesetzmäßigkeiten, die einen logischen Zusammenhang zwischen den Prämissen und der Konklusion aufbauen (Johnson-Laird, 1983).

Konditionale Aussagen sind prototypisch für deduktives Schlussfolgern (Evans, Clibbens, & Rood, 1995) und werden auf Basis von „wenn-dann“ Beziehungen gezogen (Janveau-Brennan & Markovits, 1999). In der Aussagenlogik drückt beispielsweise p die verbale Aussage (Proposition bzw. Behauptung) „Es regnet“ und q „Alicia wird nass“ aus. Der logische Operator *wenn . . . dann* setzt diese beiden Aussagen miteinander in Beziehung und bildet die Hauptprämisse einer konditionalen Aussage: „*wenn p dann q* “ (Eysenck & Keane, 2000). Der Inhalt von p und q ist dabei nicht relevant. Wichtig ist die Relation zwischen dem Antezedenten (Vordersatz) „ p “ und dem Konsequenten (Nachsatz) „ q “ einer konditionalen Prämisse (Beller & Kuhn Münch, 2007). Die beiden Behauptungen p und q können entweder wahr oder falsch sein. Durch Kombination entstehen vier mögliche Aussageformen: sowohl p als auch q sind wahr; p ist wahr, wenn q falsch ist und umgekehrt; beide, p und q , sind falsch. Wenn wir beispielsweise wissen „es regnet“ (p ist wahr) und „Alicia wird nass“ (q ist wahr), kann davon ausgegangen werden, dass die Aussage „Wenn es regnet, dann wird Alicia nass“ wahr ist (p und q sind wahr). Regnet es hingegen (p ist wahr) und Alicia ist nicht nass (q ist falsch), dann ist die Aussage „Wenn es regnet, dann wird Alicia nass“ falsch. „Wenn es nicht regnet (p ist falsch), dann wird Alicia nass“ (q ist wahr) ist eine Behauptung die wahr ist. Alicia könnte in diesem Fall durch einen anderen Umstand, etwa einem auf sie geschütteten Wassereimer nass geworden sein. Sind sowohl p als auch q falsch, z.B. „Es regnet nicht“ und „Alicia ist nicht nass“, so führt diese zu

einer wahren Aussage (Eysenck & Keane, 2000). Zusammenfassend führen konditionale Schlussfolgern dazu, Schlüsse über Situationen zu ziehen, bei denen das Auftreten eines Ereignisses bedingt oder abhängig vom Auftreten eines anderen Ereignisses ist (Thompson, 1994).

Jean Piaget versuchte, als ein Vorreiter der kognitiven Entwicklung, Aufschluss über die menschliche Fähigkeit des deduktiven Schlussfolgerns zu geben (Johnson-Laird, 1999). Er entwickelte die „genetischen Epistemologie“, eine umfassende Erkenntnistheorie. Diese beeinflusste die Lehre der kognitiven Entwicklungspsychologie entscheidend. Sein Fokus lag auf der Entwicklung höherer kognitiver Fähigkeiten, wie dem Erwerb des begrifflichen Denkens sowie dem logischen, kausalen und wissenschaftlichen Denken (Sodian, 2007). Im Folgenden wird daher die kognitive Entwicklung nach Piaget näher beschrieben.

1.2. Piaget und die kognitive Entwicklung

Piaget (1975b) interessierte sich vor allem für die Erforschung der Erkenntnisentwicklung eines Menschen und wie dieser seine Weltanschauung konstruiert. Gemäß Piaget (2003) ist es erforderlich, dass ein Individuum auf ein Objekt einwirkt, um es zu erkennen und folglich zu transformieren. Demnach ist die Erkenntnis laufend mit *Transformationen* durch Handlungen (wie etwa Stoßen oder Ziehen als fundamentale sensomotorische Handlungen) oder Operationen (komplexere, intellektuelle Handlungen wie etwa Gruppierungen, Zuordnungen, etc.) verbunden. Operationen beschreibt Piaget als „verinnerlichte Handlungen, die nicht tatsächlich durchgeführt werden müssen, sondern denkend vollzogen werden können, und die sich vor allem auf Transformationen und Relationen, also dynamische und komplexe Aspekte der Wirklichkeit beziehen“ (Scharlau, 1996, S. 48).

Ein Kind entwickelt sich, indem seine Handlungen erstmals nicht zum angestrebten Ergebnis führen, was eine Umwandlung respektive Änderung der Handlungs- bzw. Denkstrukturen erfordert. Diese Umwandlungen bzw. Änderungen ermöglichen die Entwicklung von leistungsfähigeren Strukturen im Kind, welche laut Piaget durch selbstständiges sich Ausprobieren und aktives Handeln erfolgt (Örter & Montada, 2008). Piaget beschreibt (kognitive) Strukturen als geordnete Verbindungen von Schemata (Flammer, 2009). Unter Struktur versteht Piaget Basisaktivitäten, sogenannte Operationen. Diese Operationen stellen

Handlungen dar, welche verinnerlicht, in ein System eingeordnet und reversibel sind (Scharlau, 1996).

Die kognitive Entwicklung veranschaulicht Piaget anhand von vier Stufen (Piaget, 1976b).

1.3. Vier Stufen der kognitiven Entwicklung nach Jean Piaget

Das Stufenmodell der kognitiven Entwicklung basiert auf einem strukturalistischen Ansatz, indem Piaget Entwicklung als einen Konstruktionsprozess postuliert (Örter & Montada, 2008). Piaget (1976b) beschreibt den Aufbau kognitiver Entwicklung anhand von vier Stufen, wobei die höchste Gleichgewichtsform die operative Gruppierung bildet. Diese nach Alter zugeordneten Stufen sind universell und werden immer in der gleichen Abfolge durchlebt, wobei die Geschwindigkeit, wie lange man sich in einer Stufe befindet, variieren kann (Piaget, 1984). Die Integration und Transformation von Strukturen, welche in einem vorherigen Stadium konstruiert wurden, bildet auf der nächsthöheren Stufe die Basis für neue kognitive Leistungen (Örter & Montada, 2008). Jede einzelne Stufe ist als in sich differenziert zu betrachten. Die Gesamtstruktur einer Stufe kommt erst gegen Ende, in Form der *Äquilibration*, zur Vollendung. An der *Äquilibration* (kognitives Gleichgewicht) sind zwei Grundprozesse beteiligt, nämlich die *Assimilation* und die *Akkommodation* (Piaget, 1976a). Als *Assimilation* bezeichnet Piaget (1976b) vereinfacht jede Handlung und Wirkung eines Menschen auf die Umwelt. Es handelt sich demnach um eine „Anpassung“ (Assimilierung) von neuen Elementen an schon vorher aufgebaute bzw. erworbene Strukturen (Piaget, 2003), wie etwa das kindliche Symbolspiel und mathematisches Denken (Scharlau, 1996). Die *Akkommodation* versucht die Eigenschaften von zu assimilierenden Elementen zu berücksichtigen. Bei diesem Prozess werden bereits wohlbekannte Schemata erweitert, um neue Erkenntnisse verarbeiten zu können. Das Rollenspiel bei kleinen Kindern oder das Interesse an Neuem, Unbekannten sind typische Beispiele für akkommodative Aktivitäten (Scharlau, 1996).

Dennoch bilden die Stufen kein so einheitliches Bild, wie es vielleicht den Anschein macht. Fähigkeiten, die auf einer Stufe erworben wurden, müssen auf der nächsthöheren aktiv und neu rekonstruiert werden. Oftmals wird dieses Wiederherstellen bzw. Rekonstruieren als vermeintlicher Rückschritt angesehen. Es gilt daher zu bedenken, dass es sich um einen kontinuierlichen Entwicklungsprozess handelt, welcher Schritt für Schritt durchlaufen wird (Scharlau, 1996).

Im Folgenden werden die vier Entwicklungsstadien beschrieben. Für eine eingehende Darstellung der Stadientheorie sei auf die Werke von Piaget (1972a, 1972b, 1975a, 1975b, 1976a, 1976b, 1984, 2003) sowie Piaget und Inhelder (1977) verwiesen.

1.3.1. Sensomotorisches Stadium (Geburt bis 2 Jahre)

In diesem Stadium ist die Entwicklung des Säuglings eng mit einer direkten und aktiven Auseinandersetzung mit der Umwelt verbunden. Aus angeborenen Reflexen bilden sich sensomotorische Schemata, welche die kognitive Basis für seine Handlungen darstellen und sich zunehmend differenzieren und integrieren. Ein Beispiel bildet das Saugschema, welches charakterisiert, wie ein Säugling ein Objekt in den Mund nimmt und daran saugt. Die Intelligenz des Säuglings ist handlungsgebunden (Örter & Montada, 2008). Die Fähigkeit zur Objektpermanenz wird gegen Ende dieses Stadiums ausgebildet. Die Objektpermanenz ist die Erkenntnis des Kindes, dass, sobald sich ein Gegenstand nicht mehr im Blickfeld befindet, dieser dennoch existieren kann. Das Kind entwickelt praktische Begriffe von Ding, Raum, Kausalität und Zeit (Piaget, 1972b), welche sich aber vorerst nur auf Handlungen und nicht auf das Denken beziehen. Erste Problemlösestrategien, wie etwa das Beseitigen eines Hindernisses um an ein gewünschtes Objekt zu gelangen, werden entwickelt (Örter & Montada, 2008).

1.3.2. Präoperationales Stadium (2 bis 7 Jahre)

In diesem Stadium kommt es zur Ausbildung der Symbolfunktion. Handlungen werden zu mentalen Repräsentationen verinnerlicht (z.B. im symbolischen Spiel) und ermöglichen damit auch den Spracherwerb (Piaget, 1984). Für den Erwerb physikalischer und numerischer Grundkonzepte fehlen die notwendigen logischen Voraussetzungen. Piaget führt dieses mangelnde Kausalverständnis auf das Fehlen logischer Operationen zurück (Örter & Montada, 2008). Die Fähigkeit erster Schlussfolgerungen findet in der Transduktion statt. Das Kind erlangt die Fähigkeit, von Einzelfall auf Einzelfall zu schließen (Scharlau, 1996). In diesem Stadium werden bereits ausgebildete Vorstellungen vermehrt koordiniert, logisch miteinander verbunden und „verbegrifflicht“ (Scharlau, 1996, S. 45). Schwierigkeiten bei der Erhaltung von Masse und Zahlen entstehen aufgrund der Vorstellung von Zuständen anstatt von Transformationen (Örter & Montada, 2008). Beim Umschütten von Perlen von einem breiten in ein hohes, schmales Glas

nimmt das Kind an, dass sich ohne Veränderung der Perlenanzahl nun mehr Perlen im neuen Glas befinden (Piaget, 1976b). Piaget bezeichnet diese Art des Denkens als *anschauliches Denken*, da das Kind hier noch stark im Anschaulichen selbst verhaftet ist. Das Kind ist zur Ausführung von Umformungen, s.g. Transformationen, noch nicht fähig (Scharlau, 1996). Da der Erhaltungsbegriff fehlt, gelingt Denken noch sehr eingeschränkt. Das Denken ist charakterisiert durch Egozentrismus und fehlende Reversibilität (Umkehrbarkeit) (Örter & Montada, 2008).

1.3.3. Konkret-operationales Stadium (7 bis 12 Jahre)

Dieses Entwicklungsstadium ist davon gekennzeichnet, dass konkrete Operationen mit einer Handlung verbunden sind, welche gemeinsam mit der Benutzung der Sprache zu einem logischen Aufbau führen. Die Fähigkeit, eine unabhängige Schlussfolgerung zu ziehen, welche nicht an eine konkrete Handlung gebunden ist, ist hier noch nicht gegeben (Piaget, 1976b). Ein wesentliches Merkmal des konkret-operationalen Stadiums ist die Entwicklung der Unveränderbarkeit (*Invarianz*) von Gewicht, Volumen und Masse, welche über die Transformation hinaus fort dauert. Es kommt zur Entwicklung erster (vor-)wissenschaftlicher Begriffe der Mathematik, Physik, natürlicher Zahlen sowie messbarer Zeit und Raums (Scharlau, 1996). Aufgrund einer zunehmenden *Reversibilität* (Umkehrung) des Denkens bilden sich erstmals Klassifikations- und Ordnungssysteme aus (Piaget, 1984). Reversibilität beinhaltet die Fähigkeit, Handlungen systematisch umzukehren und wieder zum Ausgangspunkt zurückzukehren (Scharlau, 1996). Piaget (1984) beschreibt zwei Formen der Reversibilität, welche bei konkreten Operationen noch voneinander getrennt angewendet werden. Zum einen spricht er von *Inversion* und *Negation*, bei der es darum geht die Wirkung einer Operation aufzuheben. Zum anderen von *Reziprozität*, bei der es um die Operation von Relationen geht (Piaget, 1984). Die Reversibilität stellt eine wichtige Rolle für die Äquilibration dar. Dadurch, dass ein Kind nun in der Lage ist, Transformationen rückgängig zu machen, besteht die Fähigkeit der Wiederherstellung eines Gleichgewichts (Scharlau, 1996).

Durch die Entwicklung der *operativen Gruppierungen* (d.h. wenn die Erhaltung eines Ganzen möglich ist) kommt es auch zur Ausbildung der *Klasseninklusion* und zum Aneinanderreihen asymmetrischer Beziehungen. Dadurch entwickelt sich die *Transitivität*, welche die Basis der Deduktionen $A = B, B = C$ somit $A = C$, oder $A < B, B < C$, somit $A < C$ darstellt. Können

erstmalig additive Gruppierungen gebildet werden, so können auch multiplikative Gruppierungen, im Sinne von Zuordnungen, verstanden werden (Piaget, 1976b).

1.3.4. Formal-operationales Stadium (ab 12 Jahren)

Die bis zu dieser Stufe erworbenen „konkreten“ Gruppierungen werden im formal-operationalen Stadium, welches in der Adoleszenz beginnt, auf eine neue Stufe des Denkens übertragen. Das formale Denken beginnt sich nun auszubilden, indem der Jugendliche nicht mehr auf Aufgaben in der Gegenwart fixiert sondern fähig ist, hypothetisch-deduktiv zu denken (Piaget, 1976b). Aussagen die verbal formuliert werden, können nun hypothetisch, entsprechend abgehoben von konkreten Inhalten, gelöst werden (Piaget, 1984). Formales Denken besteht aus einer Art Reflexion über Operationen; genauer aus Implikationen und daraus, dass zwischen Aussagen Kontroversen bestehen (Piaget, 1976b). Als bedeutendstes Merkmal des formalen Denkens wird die Änderung der Auffassung von dem was *wirklich* und von dem was *möglich* ist angesehen (Piaget & Inhelder, 1977).

Im Stadium des formalen Denkens kommt es auch zur Ausbildung von zwei neuen Strukturen. Die erste Struktur bildet das *kombinatorische System*. Voraussetzung für dieses System ist das Denken in einfachen Hypothesen, was erfordert, dass alle möglichen Kombinationen in Betracht gezogen werden können (Piaget, 1984). Diese Struktur ist ein Indiz für den Wechsel von der Gruppierung zur Gruppe (Scharlau, 1996). Die zweite Struktur bildet die *propositionale Logik*, welche einen wesentlichen Erwerb des formalen Denkens darstellt. Aus zwei Propositionen (Aussagen) p und q und deren Negation können nicht nur die vier Grundassoziationen (p und q , p und $\neg q$, $\neg p$ und q und $\neg p$ und $\neg q$) gebildet werden, auch 16 Kombinationen die sich aus dem Verbinden dieser vier grundlegenden Assoziationen ergeben, können nun berücksichtigt werden (Piaget, 1984). Dieses System aus 16 Kombinationen, bildet den Beginn für die binären 16 Operationen (Piaget & Inhelder, 1977). In diesem Zusammenhang beschreibt Piaget die sogenannte INRC-Gruppe, welche das formal-operative Denken charakterisiert und in den Werken von Piaget (1972b, 1984), Piaget und Inhelder (1977) und Scharlau (1996) detailliert beschrieben wird. Dabei geht es um die Ausführung der Kleinschen Gruppe, einer Vierergruppe (Piaget, 1984). Die vier abgeleiteten Elemente dieser Gruppe können wie folgt veranschaulicht werden: Fasst man beispielsweise die Proposition $p \vee q$ als eine konstant bleibende Implikation auf, so stellt diese eine identische Transformation I dar („entweder p richtig oder q richtig oder beide richtig“). Kommt es zur Negation N dieser Aussage (durch die Operation Reversibilität

aufgrund von Inversion), so erhält man $N(p \vee q) = \neg p \wedge \neg q$ („weder p noch q “). Ändert man dieselbe Aussage reziprok, so entsteht daraus $R = \neg q \vee \neg p$ („entweder nicht p oder nicht q “). Durch die Bildung des korrelativ C dieser Feststellung bildet sich daraus $C: p \wedge q$ („sowohl p also auch q “). Derart kommt man auf vier kommutative Transformationen, sodass $CN = R$, $CR = N$, $RN = C$ und $RNC = I$ (Piaget, 1972b, S. 307) ist. Nun ist es möglich, gültige (deduktive) Schlüsse zu ziehen, indem Aussagen aus anderen Aussagen abgeleitet werden können, wobei es nicht davon abhängt, ob die ursprüngliche Aussage selbst zutrifft. Vor allem in den formalen Wissenschaften (wie etwa der Mathematik oder Physik) werden Deduktionen angewandt (Scharlau, 1996). Zusammenfassend bilden für Piaget die formalen Operationen „nur die Struktur der höchsten Gleichgewichtsformen, zu denen die konkreten Operationen streben, wenn sie sich in allgemeine Systeme, die die verschiedenen Aussagen miteinander kombinieren, reflektieren.“ (Piaget, 1976b, S. 170).

1.4. Kritik an Piagets Theorie

Piaget war der Annahme, dass das Denken eines Kindes über die Stadien hinweg zusammenhängt. Empirische Studien zeigten jedoch, dass es zu den unterschiedlichen Entwicklungszeitpunkten zu größeren Schwankungen kommen kann, als Piaget angenommen hatte. Ein weiterer Kritikpunkt ist, dass er vor allem bei Säuglingen und jüngeren Kindern deren kognitive Fähigkeiten unterschätze (Örter & Montada, 2008). Ebenso wurden Piaget Mängel in seiner Arbeitsweise zugeschrieben. Neben dem Fehlen bedeutsamer Eigenschaften seiner Untersuchungspopulationen, wie etwa Angaben über die Zugehörigkeit zur jeweiligen sozialen Schicht, genaue Anzahl der Kinder und Jugendlichen, Schulbildung, etc., wurde ihm auch die meist fehlerhafte statistische Datenverarbeitung vorgeworfen (Buggle, 1997).

2. Übergang vom konkret-operationalen zum formal-operationalen

Denken

Ein bedeutender Schritt in der kognitiven Entwicklung findet im Übergang zwischen der Kindheit und Jugend statt (Örter & Montada, 2008). Dabei kann man den Übergang vom konkret-operationalen Stadium zum formal-operationalen Stadium nicht strikt abgrenzen, da er kontinuierlich verläuft. Im konkret-operationalen Stadium werden Aussagen in ihre entsprechenden Klassen und Verbindungen zerlegt und wieder zusammengesetzt und orientieren sich am „Wirklichen“. Hingegen können Aussagen im formal-operationalen Stadium „Möglichkeiten“ zum Ausdruck bringen und stützen sich dabei nicht mehr auf konkrete Inhalte, sondern auf den Grad der Wahrheit, respektive Falschheit der Kombinationen. Das Denken wird formal. (Piaget & Inhelder, 1977). Diesen Übergang vom konkreten zum formalen Stadium untersuchte Piaget (1984) an einer sozial höheren Bevölkerungsschicht. Anhand weiterer Forschungsergebnisse, welche sich auf andere soziale Bevölkerungsgruppen stützten, musste er erkennen, dass seine Ergebnisse nicht generalisiert werden konnten. Diese Resultate zeigten, dass das Erreichen des formal-operationalen Stadiums keineswegs als Schlusspunkt im Rahmen einer Entwicklung anzusehen ist, vielmehr wird oftmals nur das konkret-operationale Stadium erreicht. Eine mögliche Ursache sah er in der Entwicklungsgeschwindigkeit, mit der die Stadien durchlebt werden, wobei die Reihenfolge unverändert bleibt. Die unterschiedliche Qualität und Häufigkeit an intellektuellen Anreizen, die ein Kind erhält, könnte eine mögliche Ursache für diese Geschwindigkeitsdifferenz sein. Piaget vermutete, dass die Stufe des formalen Denkens verzögernd erst im Alter zwischen 15 und 20 Jahren (und nicht wie angenommen von 11 bis 15 Jahren) und unter besonders ungünstigen Umständen vielleicht sogar nie erreicht wird. Gemäß Piaget könne jedoch jeder die Stufe des formalen Denkens erreichen, wenn das soziale Umfeld und die erlernten Erfahrungen ausreichend geistige Reize bieten. Als eine zweite Erklärung nennt Piaget altersabhängige Unterschiede hinsichtlich der Fähigkeiten des Denkens. Die Ausbildung konkreter Strukturen scheint in der Gruppe von 7- bis 10/11jährigen homogener zu sein als die Entwicklung formaler Strukturen in der Gruppe der 14- bis 15jährigen. Die kognitiven Fähigkeiten differenzieren sich scheinbar zunehmend bei 12- bis 15jährigen und noch mehr bei den 15- bis 20jährigen. Diesem Umstand zufolge würden aber nur jene Personen formale Strukturen ausbilden, die Begabung in Bereichen der Logik, Mathematik und Physik zeigen. Piaget spricht hier nicht mehr von einer eigenständigen vierten Stufe, sondern von einer differenzierten Ansammlung an Entwicklungsschritten. Eine dritte und für Piaget sehr wahrscheinliche Vermutung wäre, dass jedes Individuum zwischen 15 und 20 Jahren das formal-

operationale Stadium erreicht, jedoch in unterschiedlichen Bereichen, je nach beruflichem Fachgebiet und individueller Fähigkeit (Piaget, 1984). Das hieße, Erwachsene und Jugendliche denken besonders in jenen Bereichen formal-operational, in denen sie sich bereits wohlbekanntes Wissen angeeignet haben (Piaget, 1972a).

2.1. Syllogismen

Zur Evaluierung aussagelogischen Denkens formaler Operationen im Rahmen der Entwicklungstheorie von Piaget werden häufig Syllogismen herangezogen (Byrnes & Overton, 1986; Schröder, 1989; Spiel, Gittler, Sirsch, & Glück, 1997). Syllogismen sind Aussagen über Elemente aus Mengen (Khemlani & Johnson-Laird, 2012). Ein Syllogismus baut auf einem Konditionalsatz auf. Dieser setzt sich aus einer Hauptprämisse und dazugehöriger Unterprämisse zusammen (Evans, Newstead, & Byrne, 1993; Schröder, 1989). Je nachdem, ob es sich um eine Affirmation (Bejahung) oder Negation (Verneinung) des Antezedenten (Vordersatz) bzw. Konsequenten (Nachsatz) in der Unterprämisse handelt, entstehen daraus vier Argumentformen. Tabelle 1 skizziert ein Itembeispiel aus dem „Leistungsprofiltest Schlussfolgerndes Denken-Verbal“ (SDV; Spiel, et al., 2003). Der SDV-Test ist ein Fragebogen, dessen Items aus konditionalen Syllogismen und den jeweils zugehörigen vier Argumentformen gebildet werden. Zur Erhebung der Fähigkeit schlussfolgernden Denkens kam der SDV bereits in einigen Untersuchungen zur Anwendung (z.B. Eder, 2012; Gößler, 2001; Klopff, 2011; Spiel, et al., 2004) und wird auch in vorliegender Arbeit vorgegeben.

Tabelle 1

*Aufbau der vier Argumentformen****Affirmation des Antezedenten bzw. Modus Ponens (MP):***

Wenn es regnet, dann ist die Straße nass (Hauptprämisse).

Es regnet (Unterprämisse).

Eindeutige Konklusion: Die Straße ist nass.

Negation des Antezedenten (NA):

Wenn es regnet, dann ist die Straße nass.

Es regnet nicht.

Keine eindeutige Konklusion: Vielleicht ist die Straße nass, vielleicht aber auch nicht.

Affirmation des Konsequenten (AK):

Wenn es regnet, dann ist die Straße nass.

Die Straße ist nass.

Keine eindeutige Konklusion: Vielleicht regnet es, vielleicht aber auch nicht.

Negation des Konsequenten bzw. Modus Tollens (MT):

Wenn es regnet, dann ist die Straße nass.

Die Straße ist nicht nass.

Eindeutige Konklusion: Es regnet nicht.

Die Argumentformen MP und MT stellen bikonditionale Konklusionen dar (Spiel & Glück, 2008); „wenn und nur wenn p wahr ist, dann ist q wahr“. In obigem Beispiel drückt die konditionale Aussage „*Wenn es regnet, dann ist die Straße nass*“ aus, dass die Aussage „*es regnet*“ eine notwendig Bedingung für den Konsequent „*die Straße ist nass*“ ist (Beller & Kuhn münchen, 2007). In der Argumentform MP führt dies zu der Schlussfolgerung „*regnet es folglich ist die Straße nass*“, bei der Argumentform MT zu „*die Straße ist nicht nass, folglich regnet es nicht*“. Bei MP und MT handelt es sich um valide Argumentformen, da sie zu einem eindeutigen Schluss führen.

Die beiden Formen AK und NA werden häufig als *fallacies* (Täuschungen) bezeichnet, da sie keine logischen Folgerungen nach sich ziehen (Evans, et al., 1995). Bei der Form NA kommt es

häufig zu der logisch falschen Folgerung „*die Straße ist nicht nass*“ aus „*es regnet nicht*“ und bei der Form AK „*es regnet*“ aus „*die Straße ist nass*“ zu schlussfolgern (Spiel & Glück, 2008). Da die beiden Argumentformen NA und AK zu keinem eindeutigen Schluss führen, spricht man von invaliden Argumentformen.

2.2. Entwicklungspsychologische Ergebnisse zur Schlussfolgerung mit Syllogismen

Erst beim Erreichen des formal-operationalen Stadiums mit etwa 12 Jahren ist es gemäß Piaget (1984) möglich sowohl die validen, als auch die invaliden Argumentformen zu lösen. Als Begründung führt Piaget (1984) an, dass erst im formal-operationalen Stadium die Fähigkeit entwickelt wird, in verbal formulierten Hypothesen zu denken. „Denn hypothetisch zu denken und die Konsequenzen zu ziehen, die die Hypothesen notwendigerweise implizieren (unabhängig von der tatsächlichen Wahrheit oder Falschheit der Prämissen), ist ein formaler Denkprozeß“ (Piaget, 1984, S. 49) und ermöglicht somit auch das Lösen der invaliden Formen. Demgemäß ist die Unterscheidung zwischen einer hinreichenden und notwendigen Bedingung unabdingbar, d.h. ein Verständnis für die Asymmetrie des Konditionals (Byrnes & Overton, 1986; Staudenmeyer & Bourne, 1977), welches nach Piaget (Piaget & Inhelder, 1977) im formal-operationalen Stadium auftritt.

Markovits et al. (1996) zeigten, dass bereits Kinder im Alter von 7 oder 8 Jahren, denen syllogistische Aufgaben vorgelegt wurden, valide Formen richtig lösen können. Demnach können Aufgaben der Form MP bereits mit 6 Jahren gelöst werden, MT Aufgaben mit etwa 9 Jahren. Es handelt sich um Argumentformen, die zu einem eindeutigen Schluss führen. Die Lösung des MP fällt deutlich leichter als die des MT (Evans, et al., 1995; Kodroff & Roberge, 1975; Schröder, 1989). Wildman und Fletcher (1977) legen dar, dass Kinder die Argumentform MT nur aufgrund einer Fehldeutung richtig lösen können, indem sie die Aufgaben bikonditional interpretieren und nicht deduktiv schlussfolgern. Argumentformen, die zu keinem eindeutigen Schluss gelangen, folglich AK und NA, können im Alter zwischen 13 und 15 Jahren gelöst werden (Byrnes & Overton, 1986; Kodroff & Roberge, 1975). Jüngere Kinder geben für die Formen AK und NA oft fälschlicherweise bikonditionale Antworten (z.B. „*P impliziert Q, Q ist wahr, daher ist P wahr*“), d.h. sie lösen invalide Formen häufig nicht (Janveau-Brennan & Markovits, 1999). Die Argumentform AK scheint schwieriger zu lösen als die Argumentform NA (Müller, Overton, & Reese, 2001; Newstead, Handley, Harley, Wright, & Farrelly, 2004).

Nach Schröder (1989) können invalide Argumentformen mit zunehmendem Alter kontinuierlich besser gelöst werden.

3. Entwicklungsfortschritt

Wie sich ein Entwicklungsfortschritt im deduktiven Denken äußert und welche Moderatorvariablen sich auf die deduktive Leistung auswirken, wird im folgenden Kapitel beschrieben.

3.1. Leistungsabfall bei validen Formen

Studien haben gezeigt, dass die Lösungswahrscheinlichkeit der Argumentform MT mit dem Alter sogar abnimmt, während sie sich bei den invaliden Formen erhöht (z.B. Byrnes & Overton, 1986; Janveau-Brennan & Markovits, 1999; Markovits & Vachon, 1990; Spiel & Glück, 2008; Spiel, et al., 2004). In einigen Fällen wurde auch beobachtet, dass es zu einem Leistungsabfall bei der Argumentform MP kommt, während sich die Lösungswahrscheinlichkeit der invaliden Formen erhöht (z.B. Janveau-Brennan & Markovits, 1999; Markovits, Fleury, Quinn, & Venet, 1998; Spiel & Glück, 2008; Spiel, et al., 2004). Dieser Umstand wird damit erklärt, dass die Lösungsunsicherheit der invaliden Argumentformen auf die validen Formen übergeneralisiert wird (Kuhn, 1977). Wildman und Fletcher (1977) demonstrierten, dass Studierende, die jenes Alter erreichen, welches im Allgemeinen dem Entwicklungsmaximum im formalen Schlussfolgern zugeordnet wird, zunehmend zu fehlerhaften Konklusionen der Argumentform MT kommen. Sie führen zudem an, dass MP-Aufgaben normalerweise für Jugendliche ab der 8. Schulstufe zu einfach sind, um Entwicklung deutlich zu machen. Aufgaben zu den beiden Argumentformen NA und AK erfordern jeweils nicht eindeutige Konklusionen und sind daher besonders sensitiv, um Entwicklungsverbesserungen aufzuzeigen. Fragestellungen zur Argumentform MT scheinen fälschlicherweise sensitiv für den kognitiven Prozess zu sein, da er während der Weiterentwicklung zu einer höheren Stufe des deduktiven Denkens zu fehlerhaften Lösungen führt. Diese Lösungsunsicherheit der validen Formen im fortgeschrittenen Alter spiegelt sich laut zahlreichen wissenschaftlichen Untersuchungen darin, dass es sich hierbei um einen Entwicklungsfortschritt im deduktiven Schlussfolgern handelt (z.B. Byrnes & Overton,

1986; Janveau-Brennan & Markovits, 1999; Markovits & Vachon, 1990; Spiel & Glück, 2008; Spiel, et al., 2004).

3.2. Effekte von Moderatorvariablen auf die Leistung

In der Literatur finden sich hauptsächlich zwei empirisch untersuchte Typen aufgabenbezogener Moderatorvariablen. Zum einen ist das der semantische Aufgabeninhalt und zum andern der Einsatz von Negationen in der Hauptprämisse (Spiel, et al., 1997; Spiel & Glück, 2008; Spiel, et al., 2001, 2004). Gemäß Piaget (1984) ist es im formal-operationalen Stadium möglich, losgelöst vom Inhalt eines Problems zu denken. Das Mögliche kann dem Wirklichen übergeordnet werden. Demzufolge sollten Personen im formal-operationalen Stadium fähig sein, unabhängig vom Aufgabeninhalt, richtige deduktive Schlüsse zu ziehen. Vorwissen übt jedoch einen deutlichen Einfluss darauf aus, welche Schlussfolgerungen Menschen ziehen (Thompson, 1994). Thematische bzw. vertraute Aufgaben, wie z.B. „Wenn es regnet, dann ist die Straße nass“ (Markovits & Vachon, 1990), können durch direkten Bezug auf Erfahrungen gelöst werden. Somit eignen sich derartige Aufgabenstellungen nicht, um schlussfolgerndes Denken zu erheben (Evans, et al., 1993). Konkrete Aufgabeninhalte hingegen stehen abseits von direkten Erfahrungen und faktischem Wissen (z.B., „Wenn die Sonne scheint, trägt Michaela ein rotes T-Shirt“ (Spiel, et al., 1997)). Somit kann ausgeschlossen werden, dass es zu einem direkten Rückgriff auf Erfahrungen bzw. allgemeinen Erkenntnissen kommt (Spiel, et al., 2004). Die konditionale Prämisse „Wenn da ein X ist, dann ist dort ein Y“, steht für einen abstrakten Aufgabeninhalt, bei der sämtlicher Konnex zum verfügbaren Allgemeinwissen ausgeschaltet wird (Roberge & Paulus, 1971). Im täglichen Leben sind Personen derartigen kontext-unabhängigen Interpretationen sehr selten ausgesetzt. Stattdessen interpretieren sie die konditionale Prämisse kontextspezifisch, in Form einer Ursache-Wirkungs-Aussage. Die eindeutig vorgegebene Aussage wird oftmals übergangen, indem nach alternativem Wissen oder Einflussfaktoren gesucht wird, die ausgeschlossen sein müssen, damit eine Wirkung auftritt (Beller & Kuhn Münch, 2007). Bei kontrafaktischen Prämissen werden Antezedent und Konsequent derart zusammengesetzt, dass die Aussage der Erfahrung widerspricht (Schröder, 1989; Spiel, et al., 2004). „Wenn Mäuse fliegen können, dann sind sie größer als Pferde“ ist ein Beispiel einer kontrafaktischen Prämisse (Roberge & Paulus, 1971).

Studien zeigen, dass es aufgrund der Schwierigkeit des Aufgabeninhalts zu unterschiedlichen Resultaten in der Performanz des schlussfolgernden Denkens kommt. Einigkeit besteht darin,

dass thematische Aufgaben leichter gelöst werden als Aufgaben mit abstraktem oder kontrafaktischem Inhalt (Spiel, et al., 2004). In der Studie von Overton, Ward, Black, Noveck, und O'Brien (1987) wird betont, dass thematische Kontexte die Performanz steigern. Hier erhöhen vor allem vertraute Aufgabeninhalte die Performanz etwas mehr als konkrete Aufgabeninhalte. In einer Längsschnittstudie kam Schröder (1989) zu dem Ergebnis, dass die Performanz bei kontrafaktischen Aufgaben kontinuierlich, im Vergleich zu erfahrungsnahen Inhalten jedoch zeitlich verspätet zunimmt. Abstrakte Aufgaben können allerdings früher gelöst werden als kontrafaktische Aufgabeninhalte. Erst mit zunehmendem Alter kommt es zu einer Angleichung dieser beiden letztgenannten Inhaltsformen. Außerdem stellte er Interaktionseffekte zwischen Aufgabeninhalt und Aussageform (MP, MT, AK, NA) fest. Nach Piaget sollte im formalen Stadium der Inhalt die Performanz nicht mehr beeinflussen (Piaget & Inhelder, 1977). Dabei gleicht sich die Auswirkung des Inhalts mit zunehmender Kompetenz im schlussfolgernden Denken an (Markovits & Vachon, 1990).

Negationen stellen eine weitere Moderatorvariable dar. Diese sollten bereits im konkret-operationalen Stadium gelöst werden können. Es konnte gezeigt werden, dass durch systematisches Variieren von Negationen die Performanz beim syllogistischen Schlussfolgern beeinflusst wird. Dabei scheinen vor allem Negationen im Antezedenten die Aufgabenschwierigkeit zu erhöhen und von Versuchspersonen weniger häufig gelöst zu werden (Moshman, 1977; Roberge & Mason, 1978; Wildman & Fletcher, 1977).

4. Geschlechtsunterschiede in kognitiven Fähigkeiten

Von Interesse für vorliegende Studie ist, ob sich weibliche Studierende hinsichtlich ihrer Fähigkeit im schlussfolgernden Denken von ihren männlichen Kommilitonen unterscheiden. Zahlreiche Studien zu kognitiven Geschlechtsunterschieden zeigen, dass bei sprachlichen Fähigkeiten vor allem das weibliche Geschlecht überlegen ist, Männer dagegen eine höhere Kompetenz in mathematischen bzw. naturwissenschaftlichen Bereichen aufweisen (z.B. Hyde, 1990; Hyde, et al., 1990; Kiper & Kattmann, 2003; Knight, Fabes, & Higgins, 1996). Derart geschlechtsspezifische Unterschiede wurden auch in der PISA-Studie 2012 für Österreich erhoben.

In der internationalen Schulleistungsstudie PISA (*Program for International Student Assessment*) werden die Kenntnisse und Fähigkeiten von Schülerinnen und Schülern aus 65

Teilnehmerstaaten gegen Ende ihrer Pflichtschulzeit in den Bereichen Mathematik, Lesen und Naturwissenschaften erhoben. In Österreich zeigte sich in der PISA-Studie 2012 im Vergleich zu 2003 (in beiden Jahren war Mathematik Erhebungsschwerpunkt) ein deutliches Geschlechtergefälle im Bereich Mathematik. Burschen zeigten dabei deutlich höhere mathematische Leistungen als Mädchen. Mehr Burschen als Mädchen bekundeten zudem größeres Interesse und Spaß am Unterrichtsfach Mathematik zu haben. In den Naturwissenschaften (Physik, Chemie, Biologie und Geowissenschaften) wiesen beide Geschlechter ähnliche Leistungsniveaus auf (OECD, 2014).

Bevor auf weitere geschlechtsspezifische Unterschiede eingegangen wird, soll zum besseren Verständnis der Begriff Geschlechterstereotype definiert sowie entsprechende Theorien vorgestellt werden, welche Geschlechterdifferenzen in naturwissenschaftlichen Fächern zu erklären versuchen.

4.1. Geschlechterstereotype

„Geschlechterstereotypen sind kognitive Strukturen, die sozial geteiltes Wissen über die charakteristischen Merkmale von Frauen und Männern enthalten“ (Eckes, 2008, S. 171). Dabei setzen sie sich aus einem deskriptiven und einem präskriptiven Anteil zusammen. Der deskriptive Anteil beschreibt „typische“ Eigenschaften und Verhaltensweisen von Frauen und Männern. Frauen *sind* typischerweise emotional und einfühlsam, Männer stark und dominant. Der präskriptive Anteil beschreibt, wie sich Frauen und Männer traditionell verhalten *sollen*. Demnach sollen Frauen einfühlsam und emotional sein, Männer hingegen stark und dominant. Überraschung wird erlebt, wenn die deskriptive Annahme verletzt wird. Wird die präskriptive Annahme verletzt, so kann dies Ablehnung oder gar Sanktionen zur Folge haben. Eine Verletzung stereotyper Erwartungen führt nur selten zu einer Modifizierung von Geschlechterstereotypen, da diese gegen Veränderung äußerst resistent sind (Eckes, 2008).

Analog zum Geschlechterstereotyp-Konzept existiert das Konzept der *Geschlechterrolle*. Das Gewicht liegt hier „auf den sozial geteilten *Verhaltenserwartungen*, die sich auf Individuen aufgrund ihres sozial zugeschriebenen Geschlechts richten.“ (Eckes, 2004, S. 165).

Diese Geschlechterrollenzuschreibungen sind mitunter ein Grund für sogenannte Orientierungsschemata im Bereich der Berufswahl bei Frauen und Männern; demgemäß sich Frauen für

„typische“ Frauenberufe und Männer für „typische“ Männerberufe entscheiden (Rustemeyer, 1988). Folgende Kurzgeschichte soll diese klassischen Geschlechterrollenzuschreibungen verdeutlichen:

Ein Vater und sein Sohn fahren zusammen im Auto und haben einen schrecklichen Unfall. Der Vater stirbt bei diesem Unfall. Der Sohn wird mit einem Rettungsauto ins Spital gebracht und auf den Operationstisch gelegt. Der Arzt sieht ihn an und sagt, dass hier ein Spezialist benötigt wird. Der Spezialist kommt herein, sieht den Jungen auf dem Operationstisch an und verkündet: Ich kann diesen Jungen nicht operieren, er ist mein Sohn. (sinngemäß übersetzt nach Stöger, Ziegler, & David, 2004). Auf den ersten Blick scheint diese Kurzgeschichte absurd. Die Lösung ist jedoch denkbar einfach. Der Spezialist ist eine Frau und gleichzeitig die Mutter des Jungen auf dem Operationstisch. In der Studie von Stöger, et al. (2004) konnten mehr als die Hälfte aller Teilnehmerinnen und Teilnehmer dieses Spezialisten-Rätsel nicht lösen. Verdrehten sie jedoch das elterliche Geschlecht, sodass die Mutter bei dem Unfall ums Leben kam und der Vater der Spezialist war, wurde dieses Rätsel von den Teilnehmerinnen und Teilnehmern fast zur Gänze gelöst. Diese Kurzgeschichte zeigt, dass automatisch angenommen wird, dass der Spezialist ein Mann ist.

Diese traditionellen und sozialen Normen, wie sie in einer Gesellschaft vorliegen, werden von Individuen internalisiert (Steins, 2010) und bereits früh in der Kindheit erlernt (Metz-Göckel, 1988).

4.2. Theorien zu Geschlechterstereotypen

Forschungsarbeiten, die sich mit den Inhalten von Geschlechterstereotypen beschäftigen zeigen, dass sich im Wesentlichen zwei Merkmalsdimensionen von kulturell geteilten Annahmen über Frauen und Männer subsumieren lassen (Abele & Wojciszke, 2007; Eckes, 2004). Derart stereotype Vorstellungen über Geschlechter werden den beiden Merkmalen *Wärme* oder *Expressivität* (auch Femininität, Gemeinschaftsorientierung oder „*communion*“ genannt) und (aufgabenbezogene) *Kompetenz* oder *Instrumentalität* (auch Maskulinität, Selbstbehauptung oder „*agency*“ genannt) zugeordnet (Eagly & Karau, 2002). Unter dem Konzept Wärme verbinden sich Merkmale die vielfach mit Frauen assoziiert werden. Frauen werden etwa Eigenschaften wie Einfühlungsvermögen, Sanftmütigkeit, Anpassungsfähigkeit und Verständnis zugeschrieben. Hingegen werden im Konzept Kompetenz/Instrumentalität Merkmale gebündelt, die häufig mit Männern in Verbindung gebracht werden, wie etwa Unabhängigkeit, Dominanz, Mut und

Durchsetzungsstärke (Eckes, 2004). Solche Zuweisungen klassischer Geschlechtsmerkmale scheinen zeitlich stabil und kulturell konstant zu sein (Eckes, 2008).

Einen theoretischen Erklärungsansatz für derartige Zuschreibungen klassischer Geschlechtsmerkmale versucht die Theorie der sozialen Rollen von Eagly zu geben, welche im Folgenden dargestellt wird.

4.2.1. Theorie der sozialen Rollen

In der *Theorie der sozialen Rollen* von Eagly (1987) tendieren Menschen dazu, Frauen und Männern typische Merkmale ihrer jeweiligen sozialen Rolle zuzuweisen. So etwa besteht der typische Frauenstereotyp in der Annahme, dass Frauen hauptsächlich als Hausfrauen tätig oder in Berufen mit geringerem Status zu finden sind (z.B. Lehrerinnen, Krankenschwestern oder soziale Berufe). Männer hingegen bekommen mehrheitlich die Rolle des Ernährers zugeschrieben und werden häufiger mit einem höheren Berufsstatus assoziiert (z.B. in Führungspositionen, Rechtsanwälte oder technische Berufe) (Eckes, 2008). Diese Zuschreibungen werden gemäß der Theorie der sozialen Rollen anhand psychologischer und sozialer Prozesse vermittelt. Bedeutend dabei ist, dass jedes Geschlecht erwartet, jene Merkmale bzw. Eigenschaften von sich zu haben, die für die jeweiligen Geschlechterrollen bezeichnend sind und mit denen sie auch in typischer Weise ausgestattet sind. Durch Verhaltensbeobachtung der jeweiligen Geschlechterrolle wird auf Eigenschaften des Inhabers dieser Rolle geschlossen (Eagly, Wood, & Diekmann, 2000).

Einen alternativen Erklärungsansatz bietet das Stereotypinhaltsmodell von Susan Fiske (Fiske, Cuddy, Glick, & Xu, 2002). Dieses Modell geht nicht nur von einem Globalstereotyp (Merkmalszuschreibungen, die als jeweils typisch für Frauen und Männer gelten) aus, sondern berücksichtigt auch sog. Substereotype, welche die Globalstereotype inhaltlich präzisieren oder gar widersprüchlich zu diesen sein können (Eckes, 2004).

4.2.2. Das Stereotypinhaltsmodell

Die Inhalte solcher Stereotypen determinieren sich im Stereotypinhaltsmodell („*stereotype content model*“, Fiske, et al., 2002) anhand der Einordnung von Gruppen auf den beiden Dimensionen *Wärme* und *Kompetenz*. Der relative *Status* einer Gruppe, der entweder hoch oder niedrig ist, bestimmt die Einordnung auf der Kompetenzdimension. Besitzt eine Gruppe einen hohen Status, wird dieser Kompetenz zugeschrieben; eine Gruppe wird als inkompetent erachtet, wenn diese einen niedrigen Status besitzt. Die Einordnung auf der Wärmedimension erfolgt anhand der Art der *Interdependenz* (kooperativ oder kompetitiv) zwischen Gruppen. Kooperative Gruppen werden als warm bzw. unbedrohlich beurteilt, da sie deren gemeinsame Bestrebungen nicht gefährden. Als kalt bzw. bedrohlich werden hingegen kompetitive Gruppen eingeschätzt. Werden diese beiden Dimensionen kombiniert, erhält man folgende vier Stereotype: 1) *paternalistischer* Stereotyp (warmherzig/inkompetent), 2) *neidvoller* Stereotyp (kalt/kompetent), 3) *verachtender* Stereotyp (kalt/inkompetent) und 4) *bewundernder* Stereotyp (warm/kompetent) (Fiske, et al., 2002).

Gemäß dem Stereotypinhaltsmodell wird die traditionell stereotype Frau als paternalistisch, d.h. mit relativ niedrigem sozialen Status, verbunden mit einer kooperativen Interdependenz mit Männern charakterisiert. Der traditionelle Männerstereotyp wird als neidvoll, d.h. mit relativ hohem gesellschaftlichen Status, kombiniert mit einer kompetitiven Ausrichtung gegenüber Frauen bestimmt. Diese auf den jeweiligen Dimensionen zugewiesenen traditionellen Geschlechtersubstereotypen konnten in der Studie von Eckes (2002) bestätigt werden.

Das Stereotypinhaltsmodell zeigt demnach neben einem strukturell heterogenen Globalstereotyp auch Substereotype auf, die den weit gefassten Globalstereotypen spezifizieren.

Spezifische Geschlechterstereotype konnten bereits bei Kindergartenkindern gefunden werden. Kanka, Wagner, Schober, und Spiel (2011) untersuchten Kindergartenkinder im Alter von etwa 2 bis 6 Jahren. Burschen wiesen im Vergleich zu den Mädchen sowohl eine vermehrt genderstereotype Einstellung, als auch ein häufigeres genderstereotypes Verhalten auf. Mädchen waren hingegen aufgeschlossener hinsichtlich der Bereitschaft unterschiedliche Geschlechtsrollen zu übernehmen. Außerdem zeigte sich, dass, je älter die Kinder (sowohl Burschen als auch Mädchen) werden, desto gefestigter auch deren genderstereotypes Verhalten wird (vgl. Kanka, Wagner, Schober, & Spiel, 2013). Burschen kamen zudem den Erwartungen ihrer Eltern hinsichtlich genderstereotypen Verhaltens mehr nach als Mädchen. Welchen Einfluss stereotype Erwartungen auf Kinder bzw. Jugendliche im schulischen Bereich,

insbesondere in den Fächern Mathematik und Physik haben, soll im Folgenden anhand empirischer Befunde dargestellt werden.

4.3. Der Einfluss stereotyper Erwartungen

Dass bereits Eltern genderstereotype Studienziele für ihre Kinder haben, zeigt die Studie von Dresel, Heller, Schober, und Ziegler (2001). Untersucht wurde die Erwartung der Eltern hinsichtlich geeigneter Studienfächer ihrer Kinder. 311 Eltern von Schülerinnen und Schülern der 8. Schulstufe eines Gymnasiums wurden hierzu befragt. Das Ergebnis spiegelte die konventionell geschlechterstereotypen Unterschiede wieder. Eltern schätzten für Burschen Studienfächer wie Maschinenbau oder Mathematik als geeignet ein, Mädchen hielten sie vor allem für Sprachstudien oder das Grundschullehramt prädestiniert. Ebenso wurden Leistungsunterschiede von Mädchen und Buben durch geschlechterstereotypes Denken der Eltern beobachtet (Dresel, Schober, & Ziegler, 2007).

Auch geschlechtsspezifische Erwartungen bzw. Überzeugungen der Lehrkräfte wirken sich auf die Leistungsfähigkeit, das Selbstvertrauen und die Auffassung über die eigene Begabung von Schülerinnen und Schülern aus und sind das Resultat der klassischen Geschlechtsunterschiede in Bezug auf Studien- und Berufswahl. Derart vermittelte geschlechterstereotype Erwartungen durch relevante Bezugspersonen werden internalisiert und prägen folglich das Lernen sowie den Bildungserfolg von Erwachsenen (Ludwig et al., 2006).

4.3.1. Self-fulfilling prophecy oder der Pygmalion-Effekt

Zur Erklärung schulischer Geschlechterdifferenzen aufgrund von Erwartungen in einem theoretischen Rahmen wird die „*self-fulfilling prophecy*“ (SFP; „sich selbst erfüllende Prophezeiung“) bzw. der „*Pygmalion-Effekt*“ genannt. Es handelt sich hierbei um Antizipationen bzw. Erwartungen, die zukünftige Ereignisse, wie etwa die Studienleistung, selbst hervorrufen (Ludwig, 1991, 1999). Dementsprechend werden aufgrund geschlechtsspezifischer Leistungserwartungen von relevanten Bezugspersonen, wie etwa Lehrkräften, Eltern oder Schülern, die Leistungsunterschiede zwischen Mädchen und Burschen (mit-)bedingt. Derartige SFPs werden u.a. durch stereotype Auffassungen, wie Frauen und Männer „typischerweise“ sind und sich verhalten, verursacht. Demnach haben Lehrkräfte und Eltern spezifische Erwartungen an das

Verhalten der Schülerinnen und Schüler, weshalb sich diese entsprechend differenziert verhalten (Ludwig, et al., 2006).

Studienergebnisse belegen, dass Eltern in den Domänenbereichen Mathematik und Naturwissenschaften ein geschlechtsspezifisches Denken hinsichtlich Begabung und Leistungsfähigkeit aufweisen. So zeigte sich, dass vor allem Burschen (mehr als 50%) für Mathematik und Naturwissenschaften als begabt erachtet werden und sie aufgrund der konventionellen Geschlechterstereotype auch für derartige Studienfächer als geeignet gelten. Vor allem für Mädchen, deren Eltern ausgeprägt konservative Geschlechterrollenstereotype vertreten, zeigen sich ungünstige Einstellungen hinsichtlich Leistung und Motivation in mathematisch-naturwissenschaftlichen Fächern, ganz im Gegenteil zu den Burschen, die hier Vorteile erlangen (Dresel & Schober, 2004; Ludwig, et al., 2006). Es wird hier von einem „(negativen) Golem-Effekt bei Mädchen und einem (positiven) Pygmalion-Effekt bei Jungen“ (Ludwig, et al., 2006, S. 224) gesprochen.

4.3.2. Stereotype – Threat

Das Hervorrufen negativer Stereotype bzw. stereotyper Erwartungen in Leistungssituationen kann zu reduzierten Leistungen führen (Steele, 1997). Personen, die sich in einer Leistungssituation befinden, haben demnach die Befürchtung von ihrer Gruppe aufgrund negativer Stereotype beurteilt zu werden oder eine nicht beabsichtigte Bestätigung des Stereotypen aufgrund ihres Verhaltens hervorzurufen, was sie als Bedrohung erleben (Keller, 2008). Fühlt sich eine Person aufgrund negativer Stereotype bedroht, so ist sie bestrebt, diese nicht zu bestätigen (Keller, 2008; Steele, Spencer, & Aronson, 2002). Solche Bestrebungen haben oftmals einen Abfall in Leistungssituationen zur Folge, welche man angesichts des tatsächlichen Leistungspotenzials nicht erwarten würde (Keller, 2008; Schmader, Johns, & Forbes, 2008). Dieses Phänomen wird als *Stereotype Threat* bzw. *Stereotype Bedrohung* bezeichnet. Es handelt sich dabei um einen Prozess, der sowohl die Aktivierung negativer (für eine Person relevante und leistungsbezogene) Stereotype als auch den daraus resultierenden Leistungsabfall einbezieht (Steele, 1997).

Empirische Untersuchungen konnten die negative Wirkung von Geschlechterstereotypen auf die Leistung von Frauen nachweisen. Ihme und Mauch (2007) etwa zeigten in ihrer Studie, dass Frauen nur dann eine schlechtere Performanz in leistungsbezogenen Situationen erzielen als Männer, wenn sie zuvor an den sogenannten Weiblichkeitsstereotyp erinnert wurden. Wird

dieser typische Frauenstereotyp nicht aktiviert, so weisen Frauen und Männer gleiche Leistungen auf. Analoge Ergebnisse wurden bei Untersuchungen in Leistungssituationen an Universitäten beobachtet (z.B. Spencer, Steele, & Quinn, 1999). In der Studie von Cadinu, Maass, Rosabianca, und Kiesner (2005) wurde Psychologiestudentinnen ein anspruchsvoller Mathematiktest vorgelegt. Die Teilnehmerinnen wurden einer *Stereotype Threat* und einer *Nicht-Stereotype Threat* Bedingung zugeteilt. Die Studentinnen der *Stereotype Threat* Bedingung zeigten im Vergleich zur Kontrollgruppe deutlich schlechtere Testleistungen. Ein interessanter Befund zeigte sich auch dahingehend, dass der Leistungsabfall durch ein erhöht negativ domänenspezifisches Denken in der *Stereotype Threat* Bedingung vermittelt wurde. Studentinnen in der *Stereotype Threat* Bedingung berichteten vermehrt über negative Gedanken hinsichtlich ihres Unvermögens den Test durchführen zu können bzw. deren Mangel an mathematischer Kompetenz.

Danaher und Crandall (2008) zeigten etwa, dass bereits die Abfrage des Geschlechts einer Person vor einem Test einen negativen Geschlechterstereotyp aktivieren kann und folglich zu Leistungsbeeinträchtigungen führt. Wurden die demographischen Daten jedoch nach der Testung erhoben, konnten die Leistungsunterschiede zwischen Frauen und Männern reduziert werden.

5. Einfluss von Note und Fähigkeitsselbstkonzept auf die Leistung im schlussfolgernden Denken und geschlechtsspezifische Differenzen

5.1. Zusammenhang von Note und Leistung

Noten dienen der Leistungsbeurteilung in unterschiedlichen Domänen (Unterrichts- bzw. Studienfächern) und werden oftmals als Indikatorvariable herangezogen um Schul- bzw. Studienleistung zu messen. Ihr Anspruch ist es, die reine Schulleistung bzw. den Lernerfolg möglichst valide wiederzugeben (Rost, 2010).

Zahlreiche empirische Untersuchungen konnten die gute prognostische Validität von Schulnoten auf den künftigen Studienerfolg dokumentieren (z.B. Gold & Souvignier, 2005; Rindermann & Oubaid, 1999; Steyer, Yousfi, & Würfel, 2005). In der Studie von Schmidt-Atzert (2005) etwa wurde bei der Auswahl von Studienbewerbern für das Fach Psychologie unter anderem die Schulnote als Prädiktor für den Studienerfolg herangezogen. Es konnte gezeigt werden, dass die

Abiturnote ein guter Prädiktor der Vordiplomnote, welche sich aus dem Mittelwert von sieben Einzelnoten zusammensetzte, ist.

Korrelationen von Intelligenztests mit Schulnoten weisen darauf hin, dass etwa der Zusammenhang von schlussfolgerndem Denken mit der Mathematiknote $-.45$ beträgt (Schmidt-Atzert & Amelang, 2012). Spiel, et al. (2004) erhoben in ihrer Studie die Mathematiknote als Indikator für die Leistung in Zusammenhang mit schlussfolgerndem Denken bei Gymnasiasten der 7. bis 12. Schulstufe. Im Sinne der Stufentheorie von Piaget konnten zwischen dem Stadium der konkreten Operationen und jenem der formalen Operationen zwei Übergangsstadien identifiziert werden. Es konnte demonstriert werden, dass erst im zweiten Übergangsstadium der Zusammenhang zwischen SDV-Testscore (Syllogismustest SDV, „Schlussfolgerndes Denken Verbal“) und Mathematiknote signifikant wurde und somit anzunehmen ist, dass es erst in diesem Stadium zur Anwendung schlussfolgernden Denkens kommt.

Einschränkend ist hinzuzufügen, dass die Studien- bzw. Testleistung nicht zur Gänze durch Noten konstituiert wird, da sie zusätzlich zu der objektiven Leistung auch weitere Informationen enthält, wie etwa den Zusammenhang von Note und Intelligenz oder Note und fachspezifischem Selbstkonzept (Rost, 2010). Im Folgenden soll auf den Zusammenhang des fachspezifischen Selbstkonzepts und der Leistung eingegangen werden. Zum besseren Verständnis wird zuvor das Fähigkeitsselbstkonzept definiert.

5.2. Fähigkeitsselbstkonzept

Unter Selbstkonzept versteht man die persönliche Wahrnehmung und das Wissen über die eigenen Fähigkeiten und Eigenschaften, und es ist zudem ein relativ stabiles Persönlichkeitsmerkmal (Helmke, 1992). Als Fähigkeitsselbstkonzept (oder auch „akademisches Selbstkonzept“, „schulisches Leistungs-Selbstkonzept“ oder „Selbstkonzept schulischer Fähigkeiten“) wird jener Teil des Selbstkonzeptes bezeichnet, der sich auf die Bewertung bzw. Einschätzung der eigenen Leistungsfähigkeit bezieht (Rost, 2010). Das Selbstkonzept kann auf unterschiedliche Bereiche wirken.

Das hierarchisch multidimensionale Selbstkonzeptmodell von Shavelson (Shavelson & Bolus, 1981) etwa geht von einem allgemeinen bzw. globalen Selbstkonzept an der Spitze der Hierarchie aus. Die Ebene darunter teilt sich in ein akademisches und ein nicht-akademisches Selbstkonzept. In einer weiteren Subebene wird das akademische Selbstkonzept unterteilt in

Englisch, Geschichte, Mathematik und Naturwissenschaften. Das nicht-akademische Selbstkonzept wird in die Bereiche soziales Selbstkonzept (mit einer weiteren Unterebene: Peers und „*significant others*“), emotionales Selbstkonzept (mit einer weiteren darunterliegenden Ebene: bestimmter emotionaler Zustand) und physisches Selbstkonzept (mit einer weiteren Subebene: physikalische Fähigkeit und physikalisches Erscheinungsbild) unterteilt. Die unterste Basisebene bildet die Verhaltensevaluation in spezifischen Situationen. In der revidierten Version des Modells von Shavelson wurde das akademische Selbstkonzept separat in ein mathematisches und verbales Selbstkonzept geteilt, da diese beiden Faktoren empirischen Befunden zufolge nicht miteinander korrelieren (Marsh, 1986; Marsh, Byrne, & Shavelson, 1988).

Untersuchungen zum Fähigkeitsselbstkonzept werden vor allem in Zusammenhang mit der schulischen Leistung erfasst, worauf im Folgenden näher eingegangen wird.

5.2.1. Fähigkeitsselbstkonzept und Leistung

Ob das Fähigkeitsselbstkonzept einen Einfluss auf die Leistung hat oder vice versa, ist Gegenstand zahlreicher Untersuchungen. Gemäß dem *Skill-Development-Ansatz* trägt eine verbesserte Leistung zu einem erhöhten Fähigkeitsselbstkonzept bei (Calsyn & Kenny, 1977). Im Konzept des *Self-Enhancement* kommt es hingegen aufgrund einer Verstärkung des Fähigkeitsselbstopkonzeptes zu einer verbesserten Leistung (Dickhäuser, 2006; Rost, 2010). Weiterführende Untersuchungen sprechen für ein reziprokes Effekt-Modell. In diesem Modell wirkt sich sowohl die Leistung auf das Fähigkeitsselbstkonzept als auch das Fähigkeitsselbstkonzept auf die Leistung aus. Die Auswirkung von Leistung auf das Fähigkeitskonzept ist dabei im Allgemeinen stärker als umgekehrt (Marsh & O'Mara, 2008).

In der Längsschnittstudie von Kammermeyer und Martschinke (2006) beobachtete man eine Richtungsänderung hinsichtlich der Wirkung zwischen dem Fähigkeitsselbstkonzept und der Leistung dahingehend, dass zu Schulbeginn und in den ersten beiden Jahrgangsklassen einer Grundschule das Konzept des *Self-Enhancement* wirksam ist, sich jedoch in den letzten beiden Schuljahrgängen in Richtung *Skill-Development-Ansatz* umkehrt. Begründet wurde dies mit den über die Zeit angesammelten Erfahrungen der Kinder in Leistungssituationen und den daraus resultierenden Kompetenzrückmeldungen, die sie erhielten. Der Einfluss der Leistung auf das Fähigkeitsselbstkonzept konnte bei Schülerinnen und Schülern der 3. Klasse beobachtet werden.

Wie über die eigenen Fähigkeiten geurteilt wird beruht demnach auf Kompetenzerfahrungen und darauf, wie diese interpretiert werden. Rückmeldungen die von Lehrpersonen, Eltern und Mitschülern übermittelt werden, sind dabei bedeutend. Auch Kausalattributionen, im Sinne von Schlussfolgerungen, die aus Erfolg bzw. Misserfolg gezogen werden, nehmen zur Erklärung eigener Leistungen eine bedeutsame Rolle ein (Lohrmann, Götz, & Haag, 2010). Soziale und dimensionale Vergleichsprozesse sind beteiligt wenn es gilt, fachspezifische Leistungen einzuschätzen. Das *Internal/External frame of reference* Modell (*I/E-frame of reference model*) versucht diese Vergleichsprozesse zwischen Leistung und Fähigkeitsselbstkonzept zu beschreiben (Marsh, 1986). Bei den externalen bzw. sozialen Vergleichen werden eigene Leistungen in einem Schulfach mit Leistungen von anderen Mitschülerinnen und –schülern verglichen. Bei internalen bzw. dimensional Vergleichen werden die eigenen Leistungen zwischen unterschiedlichen Fächern beurteilt (Lohrmann, et al., 2010; Marsh, 1986).

Vor allem Noten stellen hier einen wichtigen Bezugspunkt dar. Aken, Helmke, und Schneider (1997) demonstrierten in ihrer Untersuchung, dass Noten das Fähigkeitsselbstkonzept besser als standardisierte Testscores vorhersagen. Es ist anzunehmen, dass für Schülerinnen und Schüler Noten einen plastischeren Anhaltspunkt über eine hohe bzw. niedrige Fähigkeit geben als Testleistungen.

Im Rahmen external-sozialer Vergleiche resultierte die eigene Leistung, die im Vergleich mit der Leistung von Mitschülern besser war, in einem höheren Fähigkeitsselbstkonzept. Eine schlechtere Eigenleistung als die der Mitschüler führte hingegen zu einem niedrigeren Fähigkeitsselbstkonzept (Dickhäuser & Galfe, 2004). In der Studie von Möller und Köller (2004) zu internal-dimensionalen Vergleichen konnte gezeigt werden, dass Schüler mit einer durchschnittlichen mathematischen Leistung und besserer Leistung in Deutsch ein niedrigeres mathematisches Fähigkeitsselbstkonzept aufwiesen als jene, die in Deutsch eine schlechtere Leistung erlangten und ebenfalls durchschnittliche Leistungen in Mathematik erzielten.

5.2.2. Geschlechtsspezifische Leistungsunterschiede und Fähigkeitsselbstkonzept

Die PISA-Ergebnisse 2012 wiesen darauf hin, dass selbst wenn Mädchen und Burschen gleiche Ergebnisse in ihrer mathematischen Leistung erzielen, Mädchen gegenüber dem Fach Mathematik negativer eingestellt sind und weniger Lernmotivation zeigen. Mädchen schrieben sich zudem geringere mathematische Fähigkeiten zu und berichteten häufiger als Burschen,

Angst vor Mathematik zu haben. Um jedoch hohe Leistungen zu erzielen, ist das Vertrauen in die eigene Fähigkeit von Bedeutung (OECD, 2014).

Die Einstellung von Eltern allein zeigt bereits einen Einfluss auf das Fähigkeitsselbstkonzept ihrer Kinder. In der Studie von Dresel, et al. (2001) wiesen Mädchen (Gymnasiastinnen der 8. Schulstufe) von genderbezogen konservativ eingestellten Eltern ein geringeres mathematisches Fähigkeitsselbstkonzept auf. Auch Tiedemann (2000) konnte in seiner Untersuchung den Einfluss von geschlechterstereotyper Elterneinstellung auf das mathematische Fähigkeitskonzept von Grundschulern nachweisen. Eltern, die genderspezifische Ansichten vertraten, schrieben Töchtern eine geringere mathematische Fähigkeit zu als ihren Söhnen.

Niklas und Schneider (2012) zeigten in ihrer Längsschnittstudie, dass bereits nach dem ersten Grundschuljahr Mädchen bessere Leistungen im Bereich Schriftsprache (Lesen, Rechtschreiben) erzielen und Burschen höhere Kompetenzen im Bereich Mathematik aufweisen. Derart signifikante Unterschiede konnten im Vorschuljahr noch nicht gefunden werden. Es zeigte sich der Hinweis, dass das Selbstkonzept dabei einen Einfluss hat. Dieses wird jedoch erst während der Grundschulzeit ausgebildet und stabilisiert.

In einer Studie von Stipek und Gralinski (1991) beispielsweise wurde Drittklässlern einer Mittelschule ein Fragebogen zur Messung der eigenen Leistungsüberzeugung vor und nach einem angesetzten Mathematiktest vorgelegt. Es zeigte sich, dass Mädchen, im Gegensatz zu Buben, ihre Fähigkeiten, im Test gut abzuschneiden, geringer einschätzten. Sie erwarteten zudem, dass sie im Mathematiktest schlechter abschneiden würden als ihre männlichen Mitschüler und attribuierten Erfolg mit einer geringeren Wahrscheinlichkeit auf eine höhere Fähigkeit. Zu einem ähnlichen Ergebnis kamen auch Tiedemann und Faber (1995), die Grundschüler vor und nach einem Mathematiktest über ihre Einschätzungen hinsichtlich des Selbstkonzeptes befragten. Mädchen schrieben sich geringere mathematische Fähigkeiten zu und erwarteten schlechtere Leistungsergebnisse nach dem Test als ihre männlichen Mitschüler. Auch wenn das Niveau der Mathematikleistungen zwischen den Geschlechtern im späten Grundschulalter noch ausgeglichen ist, zeigen sich bereits nachweislich motivationspsychologische Veränderungen (Tiedemann & Faber, 1995). Dass sich das Fähigkeitsselbstkonzept von Schülerinnen und Schülern trotz eines annähernd gleichen Leistungsniveaus in naturwissenschaftlichen Fächern unterscheidet, zeigten Schilling, Sparfeldt, und Rost (2006) in ihrer Studie. Mädchen und Burschen erreichten in den Fächern Mathematik, Biologie, Geschichte und Physik ähnlich gute Noten. Dennoch schrieben sich Mädchen in diesen Fächern geringere Fähigkeiten zu als ihre männlichen Mitschüler. In Deutsch und Englisch

zeigten jedoch Mädchen ein höheres Fähigkeitsselbstkonzept, was durch die besseren Noten der Mädchen in diesen Fächern begründet wurde.

Wie sich aus bisherigen Befunden zusammenfassend folgern lässt, werden Geschlechtsunterschiede vor allem im spezifischen Fähigkeitsselbstkonzept beobachtet. Bei globalen Fähigkeitsselbsteinschätzungen waren derart geschlechtsspezifische Differenzen mehrfach nicht nachweisbar (Hattie, 1992; Rost, 2010).

6. Zielsetzung vorliegender Arbeit und Forschungsfragen

Ziel dieser Arbeit ist, Studierende des zweiten Studienabschnittes bzw. Master-Studierende durch Vorgabe von Syllogismen hinsichtlich ihrer Kompetenz im schlussfolgernden Denken zu untersuchen. Der Schwerpunkt liegt auf dem Zusammenhang des Geschlechterstereotyps im schlussfolgernden Denken und darin, wie sich das Fähigkeitsselbstkonzept auf die Testleistung im SDV auswirkt. Diese Untersuchung stellt eine Weiterführung der Arbeiten von Eder (2012) und Klopff (2011) dar, welche Studienanfängerinnen und Studienanfänger derselben Studienrichtungen testeten und sich auf die Stufentheorie der kognitiven Entwicklung von Piaget stützten.

Als Basis dient die Studie von Spiel, et al. (2001, 2004), welche hier erweitert werden soll. Spiel, et al. (2001, 2004) untersuchten in ihrer Studie Schülerinnen und Schüler zwischen der 7. und 12. Schulstufe mittels SDV. Durch Anwendung des Mixed Rasch-Modells (MRMs), welches im Methodenteil näher beschrieben wird, wurde eine 3-Klassen-Lösung gefunden. Umgelegt auf die Stufentheorie der kognitiven Entwicklung von Piaget konnten ein konkret-operatorisches Stadium, ein Übergangsstadium 1 und ein Übergangsstadium 2 identifiziert werden. Das höchste Stadium der formalen Operationen konnte nicht beobachtet werden.

Klopff (2011) und Eder (2012), welche Studienanfängerinnen und Studienanfänger mit einem Durchschnittsalter von 20.95 Jahren ($SD = 5$) untersuchten, konnten in ihren Arbeiten vier Klassen extrahieren. Die erste Klasse konnte dem konkret-operationalen Stadium im Sinne von Piaget zugeordnet werden. Die drei weiteren Klassen wurden als Übergangsstadium 1, Übergangsstadium 2 und Übergangsstadium 3 identifiziert. Auch in diesen Untersuchungen konnte ein Stadium der formalen Operationen nicht ermittelt werden.

In vorliegender Arbeit soll herausgefunden werden, ob ein Stadium der formalen Operationen bei Studierenden im zweiten Abschnitt identifiziert werden kann und ob die Studienrichtung und das Geschlecht einen Unterschied in der SDV-Testleistung zeigen. Zudem ist von Interesse, den Zusammenhang zwischen der Studienleistung, welche mittels Abfrage der Noten relevanter Studienfächer im Diplomprüfungszeugnis des ersten Studienabschnitts operationalisiert wurde, und der SDV-Testleistung zu überprüfen.

Des Weiteren wurde der SDV für vorliegende Arbeit um drei spezifische Items, sog. Genderitems, erweitert. Die Genderitems beinhalten eine latente Genderinformation, welche zur Erfassung von Geschlechterstereotypen eigens formuliert wurden. Ob dermaßen transportierte latente Geschlechterstereotypen schließlich unterschiedliche Lösungswahrscheinlichkeiten bei weiblichen und männlichen Studierenden im SDV bewirken, soll geprüft werden. Ebenso wird untersucht, ob das Fähigkeitsselbstkonzept der Studierenden mit der SDV-Testleistung in einem Zusammenhang steht.

Hierzu wurden folgende Fragestellungen – unter Berücksichtigung von Studierenden im zweiten Abschnitt der Studienrichtungen Mathematik, Physik, Chemie und Psychologie - formuliert.

6.1. Zusammenhang der SDV-Testleistung mit Studienrichtung, Geschlecht und Studienleistung

Um ein naturwissenschaftliches Studium absolvieren zu können, ist die Kompetenz des schlussfolgernden Denkens eine notwendige Voraussetzung (OECD, 2014; Spiel & Glück, 2008). Wie bisherige Untersuchungsergebnisse gezeigt haben, unterscheiden sich Frauen und Männer hinsichtlich der Fähigkeit im schlussfolgernden Denken dahingehend, dass Männer höhere Fähigkeiten darin aufweisen als Frauen (z.B. Hyde, 1990; Kiper & Kattmann, 2003). Zudem entscheiden sich deutlich mehr Männer für ein Studium in Mathematik, Physik und Chemie als Frauen. Klopff (2011) untersuchte in ihrer Arbeit unter anderem den Zusammenhang zwischen der Kompetenz des schlussfolgernden Denkens und der Studienwahl. Die Hypothese, dass Studienanfänger der naturwissenschaftlichen Studienrichtungen, welche mehr männliche Studierende inkludierten, höhere Kompetenzen im schlussfolgernden Denken aufweisen als Studierende der Psychologie, welche sich aus mehr weiblichen Studienanfängern zusammensetzten, konnte bestätigt werden. Dass die Mathematiknote mit dem SDV-Testscore signifikant im Zusammenhang steht, konnte Spiel, et al. (2004) demonstrieren.

Aus diesen Befunden ergeben sich folgende Forschungsfragen:

Forschungsfrage 1a:

Unterscheidet sich die Fähigkeit zum schlussfolgernden Denken bezüglich der Studienrichtung?

Forschungsfrage 1b:

Unterscheidet sich die Fähigkeit zum schlussfolgernden Denken bezüglich des Geschlechts?

Forschungsfrage 1c:

Gibt es einen Zusammenhang zwischen der Studienleistung und der Fähigkeit zum schlussfolgernden Denken?

Es wird vermutet, dass Studierende der Studienrichtungen Mathematik, Physik und Chemie eine höhere SDV-Testleistung als Studierende der Psychologie aufweisen und dass Männer besser abschneiden als Frauen. Es wird zudem angenommen, dass, je erfolgreicher die Studienleistung von Studierenden ist, desto besser auch das Abschneiden im Test ausfällt.

6.2. Kompetenz im schlussfolgernden Denken bei Studierenden im zweiten Studienabschnitt bzw. Masterstudiengang

Zentrales Merkmal im formal-operationalen Stadium ist jenes des hypothetisch-deduktiven Denkens (Piaget, 1976b). Bei der Beantwortung von Syllogismen bedeutet dies eine hohe Lösungswahrscheinlichkeit über alle Argument- und Inhaltsformen hinweg (Spiel, et al., 2004). Weder bei Spiel, et al. (2001, 2004), welche Schülerinnen und Schüler zwischen der 7. und 12. Schulstufe mittels SDV untersuchten, noch bei Klopff (2011) und Eder (2012), welche deduktives Denken bei Studienanfängerinnen und Studienanfänger der Studienrichtungen Chemie, Physik, Mathematik und Psychologie ebenfalls mittels SDV erhoben, konnte ein Stadium der formalen Operationen identifiziert werden. Vorliegende Arbeit untersucht, ob bei Studierenden im zweiten Studienabschnitt ein Stadium der formalen Operationen erreicht werden kann, woraus sich folgende Fragestellung ergab:

Forschungsfrage 2:

Kann bei im Studium fortgeschrittenen Studentinnen und Studenten eine Klasse des formal-operationalen Denkens identifiziert werden?

Diese Forschungsfrage beinhaltet die Suche nach verschiedenen, hierarchisch angeordneten Stufen des schlussfolgernden Denkens, sowie deren Anzahl und Benennung. Da in vorliegender Untersuchung Studierende der Mathematik, Physik, Chemie und Psychologie im zweiten Studienabschnitt getestet werden, wird vermutet, dass ein formal-operationales Stadium als höchste Stufe identifiziert werden kann. Anzumerken ist, dass zu dieser Forschungsfrage 2 aufgrund ihres explorativen, heuristischen Charakters keine entsprechenden Hypothesenprüfungen vorgenommen werden.

6.3. Zusammenhang von Genderinformation mit Lösungswahrscheinlichkeit

In dieser Untersuchung wurde ein Item vorgegeben, welches eine latente Genderinformation impliziert. Anhand dieses Genderitems soll ermittelt werden, ob diese zusätzliche Information mit der Lösungswahrscheinlichkeit einhergeht. Das Genderitem lautet wie folgt:

„Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.“

Die sich daraus ergebende **Fragestellung 3a** lautet:

Weisen die Genderitems Unterschiede in der Lösungswahrscheinlichkeit bezüglich des Geschlechts auf?

Es kann vermutet werden, dass die Information „Maria, die gute Laune hat“ als Widerspruch zu „dass sie in Physik eine gute Note schreibt“ aufgefasst wird und daher als Item mit kontrafaktischem Inhalt gewertet wird. Entgegen der geschlechterstereotypen Auffassung, der zufolge eine Frau in einer „frauenuntypischen“ Studienrichtung wie Physik eine gute Note schreibt, wird angenommen, dass dieses Genderitem von den Testpersonen und hier vor allem von Männern weniger häufig gelöst wird.

In diesem Zusammenhang ist ebenso von Interesse, ob die Gesamtestleistung einen Einfluss auf die Lösungswahrscheinlichkeit der Genderitems hat.

Daraus hat sich folgende **Fragestellung 3b** abgeleitet:

Steht die SDV-Testleistung mit der Lösbarkeit der Genderitems in Zusammenhang?

Es wird vermutet, dass aufgrund von vorhandenen Geschlechterstereotypen die latente Information im Genderitem ein paradoxes Lösungsverhalten bewirkt und somit auch Personen, die eine gute Testleistung zeigen, die Genderitems weniger häufig lösen können.

6.4. Zusammenhang von Fähigkeitsselbstkonzept mit der Testleistung

Um zu untersuchen, ob sich Frauen und Männer hinsichtlich des Fähigkeitsselbstkonzepts unterscheiden, sich die Selbsteinschätzung im Fähigkeitsselbstkonzept durch die Bearbeitung von Syllogismen nach der Testung verändert, und schließlich das Fähigkeitsselbstkonzept im Zusammenhang mit der Testleistung steht, ergaben sich:

Fragestellung 4a:

Wird das Fähigkeitsselbstkonzept durch die Bearbeitung von Syllogismen verändert?

Fragestellung 4b:

Gibt es Unterschiede im Fähigkeitsselbstkonzept zwischen weiblichen und männlichen Studierenden?

Fragestellung 4c:

Steht das Fähigkeitsselbstkonzept mit der SDV-Testleistung in einem Zusammenhang?

Gemäß den bisherigen Befunden hat das Fähigkeitsselbstkonzept eine Auswirkung auf die Leistungsfähigkeit. Frauen schätzen sich demgemäß im Allgemeinen weniger fähig ein als Männer.

6.5. Übersicht der Forschungshypothesen

Zusammenfassend werden aus den interessierenden Fragestellungen die entsprechend gerichteten Alternativhypothesen abgeleitet und angeführt.

Hypothesen zu den Faktoren Studienrichtung und Geschlecht:

- $H_1^{(1a)}$: Die Fähigkeit zum schlussfolgernden Denken unterscheidet sich zwischen den Studierenden der einzelnen Studienrichtungen dahingehend, dass Studierende der Naturwissenschaften eine höhere Lösungswahrscheinlichkeit als Studierende der Psychologie zeigen.
- $H_1^{(1b)}$: Die Fähigkeit zum schlussfolgernden Denken unterscheidet sich dahingehend, dass Männer eine höhere Lösungswahrscheinlichkeit als Frauen zeigen.
- $H_1^{(1c)}$: Die Studienleistung in Form von Noten weist einen negativen Zusammenhang mit der Fähigkeit zum schlussfolgernden Denken auf.

Entsprechende Hypothesen ($H^{(2)}$) für die Identifikation von verschiedenen Lösungsklassen im Sinne von Entwicklungsstadien beim deduktiven Schlussfolgern sind aufgrund der explorativen Art der Fragestellung nicht vorgesehen.

Hypothesen zu den Genderitems:

- $H_1^{(3a)}$: Die Genderitems, mit ihrer latent enthaltenen Information, weisen Geschlechtsunterschiede dahingehend auf, dass Männer kontrafaktische Genderitems weniger häufig lösen als Frauen.
- $H_1^{(3b)}$: Die SDV-Testleistung steht mit der Lösbarkeit der Genderitems in einem positiven Zusammenhang.

Hypothesen zum Fähigkeitsselbstkonzept:

- $H_1^{(4a)}$: Das Fähigkeitsselbstkonzept nimmt nach Bearbeitung der Syllogismen ab.
- $H_1^{(4b)}$: Männliche Studierende weisen eine höhere Einschätzung zum Fähigkeitsselbstkonzept gegenüber weiblichen Studierenden auf.
- $H_1^{(4c)}$: Das Fähigkeitsselbstkonzept steht mit der SDV-Testleistung in einem positiven Zusammenhang: eine höhere Selbsteinschätzung geht mit einer hohen Lösungswahrscheinlichkeit einher.

III. EMPIRIE

7. Methode und Verfahren

7.1. Untersuchungsdesign und Stichprobe

Da es sich um eine Erweiterung der Untersuchungen von Klopff (2011) und Eder (2012) handelt, sollte sich die Stichprobe dieser Arbeit ebenfalls aus Studentinnen und Studenten der Studienrichtungen Mathematik, Physik, Chemie und Psychologie zusammensetzen. Im Unterschied zu den Arbeiten von Klopff und Eder, die die Ausprägung im schlussfolgernden Denken an Studierenden zu Studienbeginn testeten, wurden in vorliegender Arbeit Studierende im zweiten Abschnitt bzw. im Masterstudium untersucht. Die Zugehörigkeit zu einem höheren Studienabschnitt war somit ein Einschlusskriterium.

Um diese Stichprobe zu erreichen, wurde das Lehrveranstaltungsverzeichnis in den jeweiligen Studienrichtungen an der Universität Wien und an der Technischen Universität Wien dahingehend durchsucht, um all jene Lehrveranstaltungen und Vorlesungen herauszufinden, welche in den jeweiligen Studienfächern im zweiten Abschnitt bzw. im Masterstudium angeboten werden. Dabei wurden die E-Mail-Adressen der Lehrveranstaltungsleiterinnen und Lehrveranstaltungsleiter notiert, um diese in weiterer Folge zwecks eines Erhebungstermins zu kontaktieren. Den Vortragenden wurde in einer E-Mail das Thema der Diplomarbeit, das Erhebungsinstrument und die Dauer der Bearbeitung (etwa 20 Minuten) in wenigen Worten dargestellt. In dieser E-Mail wurden zudem die Lehrveranstaltungsleiterinnen und -leiter ersucht, zu Beginn bzw. am Ende ihrer Lehreinheit entsprechend Zeit zur Verfügung zu stellen, um die Testung durchführen zu können. Die Erhebung wurde im Sommersemester 2011 und Wintersemester 2011/2012 durchgeführt.

Als unabhängige Variablen (UVs) wurden die beiden zweistufigen Zwischensubjektfaktoren Geschlecht (weiblich, männlich) und Studienrichtung (Naturwissenschaften, Psychologie) festgelegt. Als abhängige Variable (AV) wurde die im nachfolgenden Abschnitt beschriebene Fähigkeit zum schlussfolgernden Denken (*reasoning*) herangezogen.

7.2. Erhebungsinstrument

7.2.1. SDV

Zur Operationalisierung der Untersuchung wurde der „Leistungsprofiltest Schlussfolgerndes Denken-Verbal“ (SDV) von Spiel, et al. (2003) in einer erweiterten Form als Erhebungsinstrument herangezogen. Etwaige Erweiterungen werden in diesem Kapitel weiter unten beschrieben. Zuvor soll der SDV jedoch in seiner ursprünglichen Version erläutert werden.

Mit dem SDV ist es möglich, Leistungshöhe und Leistungsprofil im schlussfolgernden Denken bei Jugendlichen und Erwachsenen zu erheben (Spiel, et al., 2004). Grundlage dieses Tests bildet dabei Piagets Entwicklungstheorie. Hierbei werden konditionale Syllogismen vorgegeben und systematisch variiert. Neben den vier Argumentformen (MP, NA, AK, MT) werden drei mögliche Inhaltsformen (konkret, abstrakt, kontrafaktisch) und die Möglichkeit der Negation (negiert, unnegiert) modifiziert. Somit ergeben sich $4 \times 3 \times 2 = 24$ Variationsmöglichkeiten an Syllogismen (Göbler, 2001).

Diese 24 Items werden aus sechs konditionalen Syllogismen und jeweils vier möglichen Argumentformen gebildet (Spiel, et al., 2004). In den Untersuchungen von Spiel, et al. (2001, 2004) wurde der SDV mit 24 Items vorgegeben. Die Negationen werden dabei immer nur in der Hauptprämisse im Antezedenten verwirklicht, um die Beschränkung auf 24 Items nicht zu überschreiten. Abbildung 1 verdeutlicht die sechs konditionalen Syllogismen in ihrer ursprünglichen Form.

	konkret	abstrakt	kontrafaktisch
Ohne Negation des Antezedenten	Wenn die Sonne scheint, trägt Tina einen roten Rock.	Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.	Wenn es Abend ist, geht die Sonne auf.
Mit Negation des Antezedenten	Wenn die Sonne nicht scheint, trägt Peter eine blaue Hose.	Wenn X nicht zur Gruppe B gehört, dann hat X die Eigenschaft c.	Wenn es nicht Abend ist, geht die Sonne unter.

Abbildung 1. Sechs Beispielimis konditionaler Syllogismen aus dem ursprünglichen SDV.

Um eine robustere Schätzung auf Einzelebene zu erzielen, wurden dem SDV weitere 8 Items mit hoher Trennschärfe hinzugefügt. Diese erweiterte Version des SDV beinhaltet insgesamt 32 Items. Die Hauptprämissen der neuen Items sind in Abbildung 2 exemplarisch angeführt.

	konkret	abstrakt	kontrafaktisch
Ohne Negation des Antezedenten	Wenn Herr Müller gute Laune hat, fährt er nach Berlin	Wenn A auf B trifft, dann entscheidet sich A für c.	Wenn es regnet, ist die Straße trocken.
Mit Negation des Antezedenten	Wenn Frau Schuster keine gute Laune hat, geht sie spazieren.	Wenn C nicht auf D trifft, entscheidet sich C für m.	Wenn es nicht regnet, ist die Straße nass.

Abbildung 2. Neue hinzugefügte SDV-Items.

Klopf (2011) und Eder (2012) gaben in ihren Arbeiten die erweiterte Version des SDV mit 32 Items vor.

Nachstehend soll ein Beispielitem aus dem in dieser Untersuchung verwendeten SDV veranschaulicht werden.

Wenn es Abend ist, geht die Sonne auf.

Es ist Abend.

☐ Die Sonne geht auf.

☐ Die Sonne geht nicht auf.

☐ Vielleicht geht die Sonne auf, vielleicht aber auch nicht.

☐ Ich kann diese Aufgabe nicht lösen.

Dieses unnegierte Item ist in der Argumentform MP formuliert und hat einen kontrafaktischen Inhalt. Die Items im SDV werden jeweils als Hauptprämisse („wenn, dann“) mit jeweiliger Unterprämisse vorgeben. Negationen werden dabei ausschließlich im Antezedenten der Hauptprämisse gebildet. Die Items unterscheiden sich hinsichtlich ihrer Argumentform MP, MT, AK oder NA und ihres Aufgabeninhalts konkret, abstrakt oder kontrafaktisch. Insgesamt stehen zur Beantwortung der Items vier Antwortmöglichkeiten zur Verfügung, wobei die Option „Ich kann diese Aufgabe nicht lösen“ deshalb formuliert wurde, um ein eventuelles Raten der

Testpersonen zu vermeiden. Dem SDV liegt eine ausführliche Instruktion bei, welche den Testpersonen eine konkrete Anweisung und Information zur Beantwortung der Items gibt.

7.2.2. Erweiterung um ein Genderitem

In vorliegender Untersuchung wurde der SDV mit 32 Items um ein zusätzliches Item (formuliert in drei Argumentformen), welches eine implizite Genderinformation enthält, erweitert. Das Genderitem wurde von Frau Univ.-Prof. DDr. Mag. Christiane Spiel formuliert und lautet wie folgt:

„Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.“

Die darin enthaltene latente Genderinformation „Maria“, als prototypisch weibliche Person, welche auch „gute Laune“ hat, schreibt in einem eher „frauenuntypischen Fach wie Physik“ eine „gute Note“ soll darauf schließen lassen, dass dieses Item von den Testpersonen aufgrund vorhandener Geschlechterstereotype als Item mit kontrafaktischem Inhalt angesehen wird und entsprechend schwer gelöst wird. Dieses Genderitem wurde in den drei Argumentformen, MP, NA und AK als jeweils kontrafaktisch und unnegiert vorgegeben.

Um mögliche Reihenfolgeeffekte zu überprüfen, wurde der SDV in zwei Versionen A und B vorgegeben. Das Genderitem wurde dabei in seinen drei Ausführungen (MP, NA und AK, kontrafaktisch, unnegiert) in der Version A in der ersten Hälfte des Tests vorgegeben (Präsentation der Items auf Position 1, 11 und 21) und in Version B in der zweiten Hälfte des Tests (Präsentation der Items auf Position 15, 25 und 35). Somit ergab sich eine Testverlängerung von 32 Items auf 35 Items. Die Testversionen A und B sind im Anhang beigelegt.

7.3. Untersuchungsdurchführung

Die Untersuchung wurde im Juni 2011, im Wintersemester 2011/2012 sowie ein weiteres Mal zu Beginn des Sommersemesters 2012 anhand einer Stichprobe aus der Grundgesamtheit der interessierenden Studienrichtungen durchgeführt.

Vor der Testung wurde zunächst eine kurze Instruktion gegeben und auf die Freiwilligkeit der Teilnahme und die anonymisierte Datenanalyse hingewiesen. Es wurde darauf aufmerksam gemacht, die Anleitung durchzulesen, den Fragebogen vollständig auszufüllen und genau, aber zügig durchzuarbeiten. Sodann wurde der SDV randomisiert in den Paper-Pencil Versionen A und B an die Studierenden verteilt.

Die Berechnungen im Rahmen der deskriptiv- und inferenzstatistischen Analysen erfolgten mittels Statistiksoftware IBM SPSS® 17.0 für Windows. Im Rahmen der Hypothesenprüfung wurde vorab ein Signifikanzniveau von $\alpha = 5\%$ festgelegt. Ergebnisse der Hypothesenprüfungen, die diesen Wert erreichen oder unterschreiten, werden als signifikant bezeichnet. Zur Berechnung des Mixed Rasch-Modells wurde die Programmsoftware WINMIRA® (Version 1.741) herangezogen.

7.4. Spezifisches Auswertungsverfahren

Als sinnvolles Instrument zur Identifikation von kognitiven Entwicklungsstadien bietet sich das Mixed Rasch-Modell (MRM) an (Spiel, et al., 1997; Spiel & Glück, 2008; Spiel, et al., 2001, 2004) und wird dementsprechend in vorliegender Arbeit herangezogen.

Das Mixed Rasch-Modell (MRM; Rost, 1990; Rost & von Davier, 1995) ist eine Verbindung aus dem Rasch-Modell und der Latent Class Analyse. Es baut im Wesentlichen auf dem dichotom logistischen Modell von Rasch (Rasch-Modell, RM) auf, welches prüft, inwieweit Items dieselbe zugrundeliegende, latente Eigenschaft erfassen (Rost, 2004). Das RM nimmt an, dass für alle Personen dieselben Itemparameter gelten, wobei die Itemschwierigkeit über alle Personen-gruppen gleich bleibt. Die Latent Class Analyse wiederum fordert, dass alle Personen derselben Klasse für alle Items auch dieselben Lösungswahrscheinlichkeiten aufweisen.

Das Mixed Rasch-Modell als Verbindung dieser beiden Ansätze hingegen lässt für unterschiedliche Itemschwierigkeiten unterschiedliche Personengruppen zu. Das bedeutet, dass Personen unterschiedlichen latenten Klassen zugeteilt werden können, innerhalb derer das Rasch-Modell gilt (Rost, 2004).

Da im MRM alle Antwortmuster einbezogen werden, ist es möglich, latente Subklassen von Personen zu identifizieren. Nachdem das MRM innerhalb dieser latenten Klassen auch quantitative Personenfähigkeitsunterschiede zulässt, können neben quantitativen auch qualitative Leistungen beurteilt werden (Spiel & Glück, 2008; Spiel, et al., 2004). In den Untersuchungen

von Spiel, et al. (2004), Klopff (2011) und Eder (2012) wurden durch Anwendung des MRM sogenannte Übergangsstadien zwischen dem konkret-operationalen und dem formalen Stadium im Sinne Piagets identifiziert. Das MRM stellt somit, als adäquates Instrument zur Erforschung kognitiver Entwicklungsstufen im schlussfolgernden Denken, das Verfahren der Wahl dar.

Als informationstheoretisches Maß zur Überprüfung der Modellgeltung und der optimalen Klassenanzahl dient das „Best Information Criterion“ (BIC), wobei die Modellpassung umso besser ausfällt, je kleiner dieser Kennwert ist. Rost (2004) bezeichnet den BIC als geeignete Kennzahl bei großen Stichproben. Aufgrund des Stichprobenumfangs in vorliegender Untersuchung wurde der BIC als geeignetes Maß herangezogen.

Für weitere und detailliertere Informationen zum Mixed Rasch-Modell und dessen Modellgeltung sei auf Rost (1990, 2004) und Rost und von Davier (1995) verwiesen.

7.5. Stichprobe

Die Gesamtstichprobe umfasste 395 Studierende, welche sich im zweiten Abschnitt bzw. im Masterstudium ihrer jeweiligen Studienrichtung befanden. Das Durchschnittsalter lag bei 24.05 ($SD = 4.78$) Jahren. Insgesamt wurden 184 weibliche und 211 männliche Studierende getestet.

205 Studierende entfielen dabei auf die Studienrichtung Psychologie. Der Altersdurchschnitt lag hier bei 25.41 ($SD = 5.33$) Jahren. 190 Studierende entfielen insgesamt auf die Studienrichtungen Physik ($n = 111$, Altersdurchschnitt = 22.35 ($SD = 3.68$) Jahre), Chemie ($n = 6$, Altersdurchschnitt 22.83 ($SD = 4.49$) Jahre) und Mathematik ($n = 73$, Altersdurchschnitt 22.93 ($SD = 3.41$) Jahre). Aufgrund der Vergleichbarkeit der naturwissenschaftlichen Studienrichtungen Chemie, Mathematik und Physik können diese Untersuchungsteilnehmer zusammengefasst werden. Der Altersmittelwert dieser Gruppe lag bei 22.95 ($SD = 3.60$).

Tabelle 2

Vierfeldertafel zu Geschlecht und Studienrichtung

			Geschlecht		Gesamt
			weiblich	männlich	
Studien- richtung	Psychologie	Anzahl	143	62	205
		Anteilswert	69,8%	30,2%	100,0%
	Naturwissen- schaften	Anzahl	41	149	190
		Anteilswert	21,6%	78,4%	100,0%
Gesamt		Anzahl	184	211	395
		Anteilswert	46,6%	53,4%	100,0%

Insgesamt ist das Geschlechterverhältnis in der vorliegenden Untersuchung annähernd gleichverteilt (siehe Tabelle 2), während innerhalb der Studienrichtungen die Anteilswerte der weiblichen und männlichen Studierenden deutliche Unterschiede aufweisen.

Die Berechnung der Prüfgröße fiel mit $\chi^2(1) = 91.98$, $p < .001$ signifikant aus. Es kann angenommen werden, dass die Verteilung der Geschlechter auf die beiden Studienrichtungen unterschiedlich ist. Darüber hinaus wurde geprüft, ob diese Unterschiedlichkeit auch für die jeweiligen Studienrichtungen repräsentativ ist. Die Anteilswerte in den Naturwissenschaften der Universität Wien und TU Wien im Studienjahr 2011/2012 betrug für männliche Studierende 71,41% und für weibliche Studierende 28,59%; für die Studienrichtung Psychologie der Universität Wien für männliche Studierende 23,39% und für weibliche Studierende 76,61%. Mittels Chi-Quadrat-Anpassungstest wurden diese theoretisch zu erwartenden gegen die beobachteten Verteilungen verglichen; mit $\chi^2(1) = 4.58$, $p = .032$ für die Naturwissenschaften und $\chi^2(1) = 5.43$, $p = .020$ für Psychologie wurden jeweils ein etwas erhöhter Männeranteil in den Stichproben untersucht. Die Daten für die zu erwartenden Verteilungen stammen von der Technischen Universität Wien – TISS (2011) sowie vom Bundesministerium für Wissenschaft, Forschung und Wirtschaft (2011) aus dem Wintersemester 2011 und beziehen sich auf ordentlich Studierende der Masterstudien und Diplomstudien der jeweiligen Studienrichtung.

36 (9,1% der) Untersuchungsteilnehmerinnen und Untersuchungsteilnehmer haben eine andere Muttersprache als Deutsch. Es können keine Verteilungsunterschiede in Abhängigkeit von den beiden Studienrichtungen mit $\chi^2(1) = 2.69$, $p = .101$ beobachtet werden.

8. Ergebnisse

In vorliegender Untersuchung wurden, im Gegensatz zu den Arbeiten von Klopff (2011) und Eder (2012), Teilnehmerinnen und Teilnehmer mit nicht-deutscher Erstsprache nicht ausgeschlossen. Es war daher mittels t-Test für unabhängige Stichproben zu prüfen, ob die SDV-Lösungswahrscheinlichkeit einen Unterschied in Abhängigkeit von der Erstsprache aufweist (siehe Tabelle 3).

Tabelle 3

Lösungswahrscheinlichkeit im SDV in Abhängigkeit der Erstsprache

	Erstsprache	n	<i>M</i>	<i>SD</i>
SDV	Deutsch	359 (90,9%)	.641	.214
	Nicht Deutsch	36 (9,1%)	.625	.221

Die Berechnung der entsprechenden Prüfgröße (Homogenität der Varianzen konnte mit $p = .701$ angenommen werden) fiel mit $t(393) = 0.43$, $p = .664$ ($d = 0.08$) nicht signifikant aus. Die beiden Gruppen unterscheiden sich nicht hinsichtlich der SDV-Lösungswahrscheinlichkeit.

8.1. Unterschiede im schlussfolgernden Denken bezüglich Studienrichtung und Geschlecht

Mittels zweifaktorieller univariater ANOVA wurden die Unterschiedlichkeiten in der SDV-Leistung in Abhängigkeit von der Studienrichtung und dem Geschlecht überprüft. Die abhängige Variable (AV) bildet die mittlere Lösungswahrscheinlichkeit im SDV; die Studienrichtung und das Geschlecht repräsentieren die beiden unabhängigen Variablen (UV). Tabelle 4 zeigt die Lösungswahrscheinlichkeit der SDV-Leistung in Abhängigkeit von Studienrichtung und Geschlecht.

Tabelle 4

Kennwerte SDV Leistung in Abhängigkeit von Studienrichtung und Geschlecht

Studienrichtung	Geschlecht	<i>M</i>	<i>SD</i>	n
Psychologen	weiblich	.558	.159	143
	männlich	.598	.151	62
	Gesamt	.570	.158	205
Naturwissenschaftler	weiblich	.705	.236	41
	männlich	.717	.243	149
	Gesamt	.715	.241	190
Gesamt	weiblich	.591	.189	184
	männlich	.682	.227	211
	Gesamt	.640	.214	395

Das Ergebnis zeigt einen signifikanten Effekt mit $F(1, 391) = 31.84, p < .001$ ($\eta^2_p = .075$) und weist darauf hin, dass die mittlere Lösungswahrscheinlichkeit einen Unterschied bezüglich der Studienrichtung aufweist. Studierende aus den naturwissenschaftlichen Studienrichtungen haben eine höhere Leistung als jene der Fachrichtung Psychologie im SDV. Somit kann die Alternativhypothese $H_1^{(1a)}$ angenommen werden. Des Weiteren zeigt sich, dass das Geschlecht in keinem signifikanten Zusammenhang mit der mittleren Lösungswahrscheinlichkeit steht, die Prüfgröße fiel mit $F(1, 391) = 1.24, p = .267$ nicht signifikant aus. Es kann kein Niveauunterschied zwischen den Geschlechtern angenommen werden. Mit $F(1, 391) = 0.32, p = .571$ sind darüber hinaus keine Wechselwirkungen aus Studienrichtung x Geschlecht anzunehmen. Die Alternativhypothese $H_1^{(1b)}$ wird somit zurückgewiesen.

8.2. Zusammenhang der Studienleistung mit der Lösungswahrscheinlichkeit im SDV

Die durchschnittliche Studienleistung wurde aus den Noten dreier prüfungsrelevanter Fächer gebildet, um eine möglichst adäquate Charakteristik des Studienerfolgs zu erreichen. Mittels Produkt-Moment Korrelation wurde der Zusammenhang zwischen Studienleistung und der SDV Lösungswahrscheinlichkeit untersucht. Hohe Werte (schlechte Noten) korrelieren mit einer niedrigen Leistung (Lösungswahrscheinlichkeit) im SDV zu $r = -.210$ ($p < .001$, einseitig, $n = 360$). Der Korrelationskoeffizient nach Pearson weist auf einen niedrigen bis knapp mittelhohen negativen Zusammenhang hin. Die $H_1^{(1c)}$ kann angenommen werden.

Die Abbildung 3 zeigt das entsprechende bivariate Streudiagramm für die Studienleistung und die Lösungswahrscheinlichkeit im SDV unter Berücksichtigung der beiden Studienrichtungen.

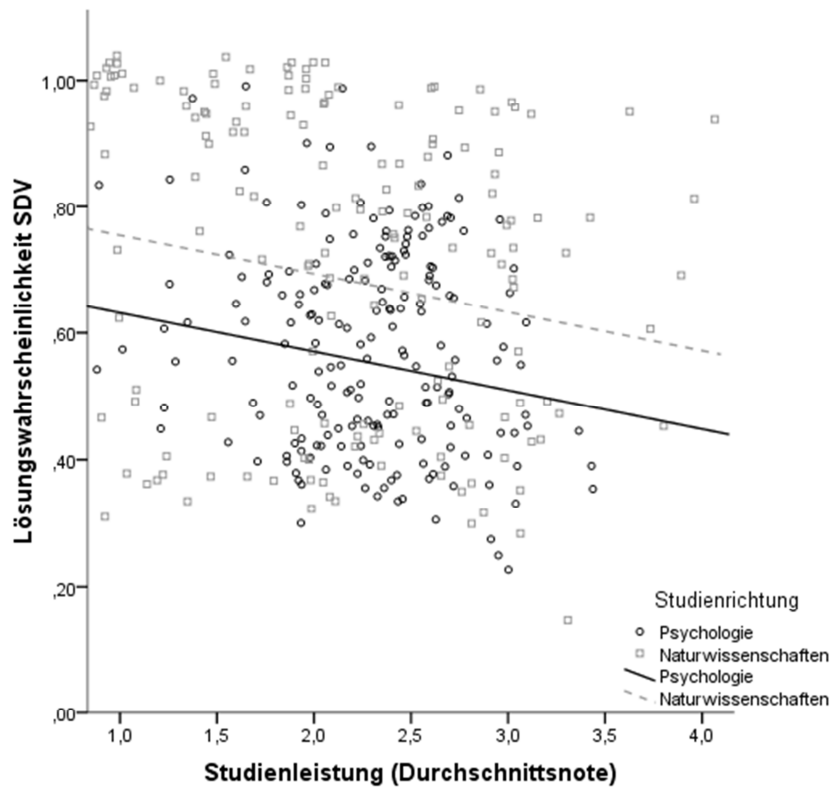


Abbildung 3. Bivariates Streudiagramm für Studienleistung und Lösungswahrscheinlichkeit im SDV unter Berücksichtigung der beiden Studienrichtungen.

8.3. Extraktion der Lösungsklassen

Wie erwähnt dient der BIC als informationstheoretisches Maß zur Überprüfung der Modellgeltung. Je kleiner der BIC-Wert, desto höher das Ausmaß der Modellpassung. In vorliegender Arbeit wurden mittels WINMIRA[®] die BIC-Werte sowohl unter Einbeziehung der Genderitems (in Summe 35 Items) als auch ohne Genderitems (in Summe 32 Items) berechnet.

Tabelle 5

BIC-Werte der Mixed Rasch-Modelle mit und ohne Genderitems mit unterschiedlicher Klassenanzahl

Klassenanzahl	BIC-Werte mit Genderitems (35 Items)	BIC-Werte ohne Genderitems (32 Items)
1	14454.00	13433.20
2	13248.21	12260.90
3	12621.07	11734.70
4	12449.05	11572.31
5	12347.91	11490.80
6	12372.74	11512.33
7	12397.19	11543.55
8	12483.59	11615.54
9	12695.20	11769.45
10	12740.61	11836.33

Tabelle 5 veranschaulicht, dass der niedrigste BIC-Wert in beiden Berechnungsmethoden bei der Klassenanzahl 5 liegt. Somit sollte eine Fünf-Klassenlösung das Modell am besten beschreiben. Die Zuordnungswahrscheinlichkeit zu den einzelnen Lösungsklassen ist in Tabelle 6 dargestellt. Da Berechnungen unter WINMIRA® nur auf vollständigen Datensätzen beruhen können, waren vier Testpersonen, bei welchen die SDV-Testung unvollständig war, auszuschließen. Die Anzahl der gültigen Fälle verringerte sich von 395 auf 391.

Tabelle 6

Mittlere Zuordnungswahrscheinlichkeit gereiht nach Klasse mit Genderitems

Klasse (mit Genderitems)	N	M (Sicherheit)	SD
1	97	.9818	.0684
2	30	.9983	.0058
3	15	.9998	.0004
4	147	.9063	.1536
5	102	.8921	.1676
Insgesamt	391	.9320	.1380

Die Zuordnungssicherheit zu den einzelnen Klassen erreicht auf Grundlage der WINMIRA® - Prozedur unter Berücksichtigung aller 35 Items im Durchschnitt 93,20%.

8.3.1. Lösungswahrscheinlichkeit in Abhängigkeit der WINMIRA – Klassen

In diesem Abschnitt (Tabelle 7 bis Tabelle 11) sind die mittleren Lösungswahrscheinlichkeiten je Klasse, unter Berücksichtigung der Argumentform (MP, MT, NA und AK) und des Aufgabeninhaltes (konkret, abstrakt und kontrafaktisch), dargestellt.

Tabelle 7

Mittlere Lösungswahrscheinlichkeiten Klasse 1

Klasse 1 (n = 97)	MP	MT	NA	AK
konkret	.8299	.6117	.3093	.2500
abstrakt	.9381	.7251	.1753	.1005
kontrafaktisch	.8591	.7010	.2715	.1924

Tabelle 8

Mittlere Lösungswahrscheinlichkeiten Klasse 2

Klasse 2 (n = 30)	MP	MT	NA	AK
konkret	.0167	.2000	.8000	.3667
abstrakt	.4833	.3556	.2167	.7833
kontrafaktisch	.2333	.0667	.6111	.3556

Tabelle 9

Mittlere Lösungswahrscheinlichkeiten Klasse 3

Klasse 3 (n = 15)	MP	MT	NA	AK
konkret	.1333	.0667	.8222	.9167
abstrakt	.8000	.7778	.0333	.1000
kontrafaktisch	.2667	.3111	.5778	.6667

Tabelle 10

Mittlere Lösungswahrscheinlichkeiten Klasse 4

Klasse 4 (n = 147)	MP	MT	NA	AK
konkret	.9694	.4807	.8912	.6361
abstrakt	.9456	.5011	.7687	.4643
kontrafaktisch	.9388	.4286	.8231	.5488

Tabelle 11

Mittlere Lösungswahrscheinlichkeiten Klasse 5

Klasse 5 (n = 102)	MP	MT	NA	AK
konkret	1.00	.5392	.9052	.9730
abstrakt	.9853	.5392	.9020	.9436
kontrafaktisch	1.00	.4902	.8922	.9739

Abbildung 4 stellt zur besseren Übersichtlichkeit die mittleren Lösungswahrscheinlichkeiten aller fünf Lösungsklassen, getrennt nach Argumentform MP, MT, NA und AK sowie nach Aufgabeninhalt konkret, abstrakt und kontrafaktisch, dar.

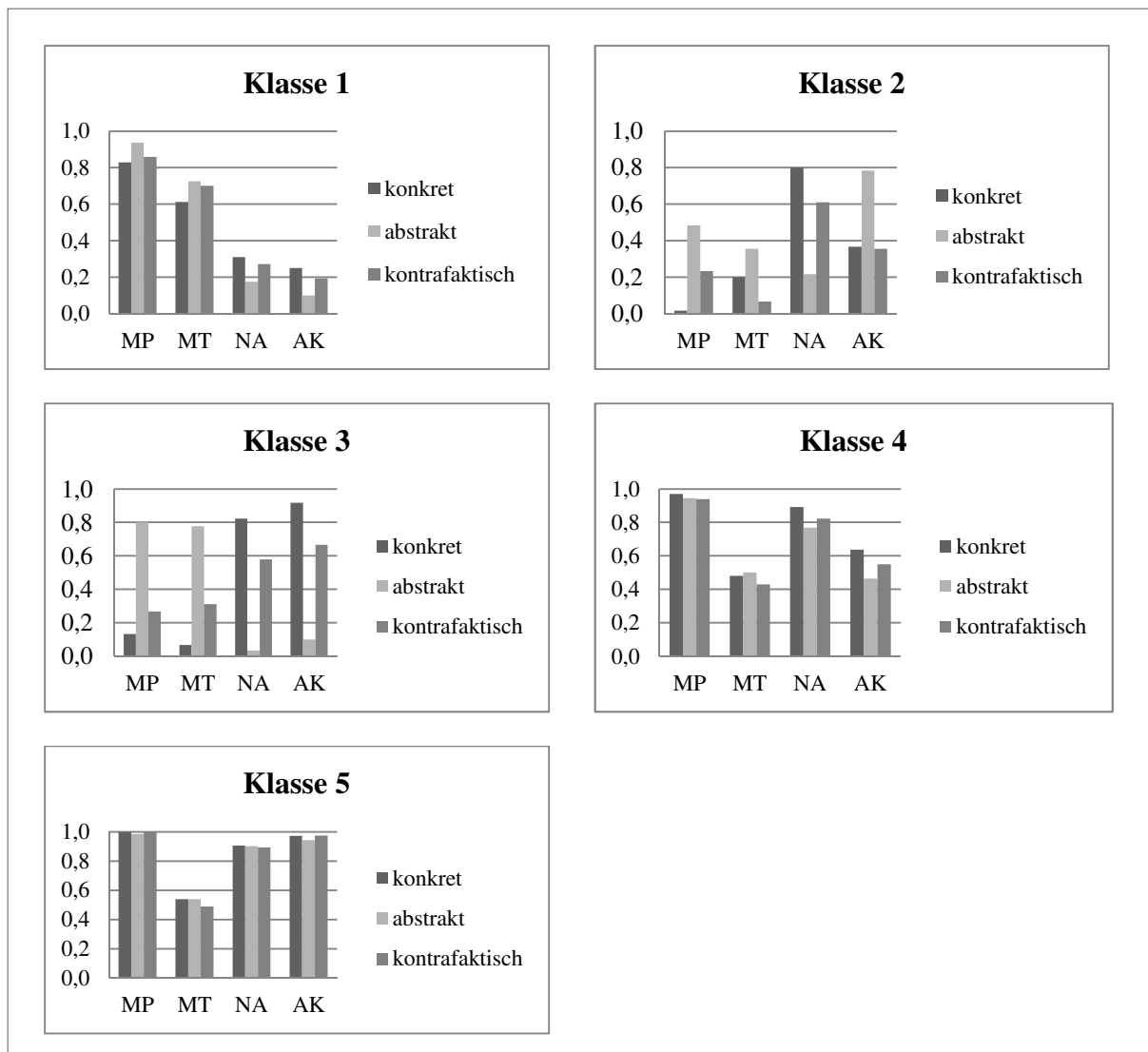


Abbildung 4. Mittlere Lösungswahrscheinlichkeiten der 35 Items getrennt über die vier Argument- und drei Inhaltsformen pro Klasse.

Aufgrund der Lösungswahrscheinlichkeiten von 391 gültigen Fällen in den Argument- und Inhaltsformen können die ermittelten Klassen in einer aufsteigenden Hierarchie wie folgt charakterisiert werden. Die zugehörigen Lösungswahrscheinlichkeiten in den einzelnen Bedingungen und je Klasse können der Tabelle 12 entnommen werden.

Klasse 1

Diese Klasse (97 Teilnehmerinnen und Teilnehmer; 24,8%) weist ein bikonditionales Antwortmuster auf. Dabei werden, unabhängig von Inhalt und Negation, die invaliden Aufgaben (NA und AK) seltener gelöst (.214, $SD = .173$), die validen Aufgaben (MP und MT) hingegen häufiger (.764; $SD = .163$); $t(96) = 23.87, p < .001$. Im Detail werden abstrakte Aufgabeninhalte bei den validen Bedingungen häufiger (.810, $SD = .190$) als bei den invaliden Bedingungen (.125, $SD = .195$) gelöst. Ebenso werden konkrete, valide Aufgaben häufiger (.699, $SD = .282$) als konkrete, invalide (.275, $SD = .271$) gelöst und kontrafaktische, valide Aufgabeninhalte häufiger (.780, $SD = .178$) als kontrafaktische, invalide (.232, $SD = .205$).

Klasse 2

Mit 30 beobachteten Teilnehmerinnen und Teilnehmern (7,7%) werden hier vor allem bei den NA-Aufgaben konkrete aber auch kontrafaktische Inhalte häufiger gelöst; Aufgaben mit abstraktem Inhalt jedoch selten. Bei den AK-Aufgaben zeigt sich das genaue Gegenteil. Abstrakte Aufgabeninhalte werden häufig gelöst, konkrete und kontrafaktische jedoch selten. Bei den validen Aufgaben hingegen zeigen sich geringe Lösungswahrscheinlichkeiten (.223, $SD = .117$), wobei bei der Aufgabenform MP Aufgaben mit konkretem Inhalt fast gar nicht gelöst werden können; bei den invaliden Aufgaben zeigt sich eine höhere Lösungswahrscheinlichkeit (.544, $SD = .118$); $t(29) = -13.33, p < .001$.

Klasse 3

Mit 15 beobachteten Teilnehmerinnen und Teilnehmern ist diese Lösungskategorie im Vergleich die kleinste (3,8%). Auffällig in dieser Klasse ist, dass bei validen Aufgaben (.383, $SD = .115$) vor allem abstrakte Aufgabeninhalte häufig (.787, $SD = .177$) gelöst werden können. Konkrete und kontrafaktische Aufgabeninhalte kaum bis sehr selten. Die invaliden Aufgaben (.544, $SD = .125$) zeichnen sich durch eine häufigere Lösungswahrscheinlichkeit, vor allem der konkreten, jedoch auch kontrafaktischen Aufgabeninhalte aus; $t(14) = -2.78, p = .015$. Abstrakte Aufgabeninhalte können hier kaum bis nur selten gelöst

werden ($.078$, $SD = .124$). Konkrete, valide Aufgaben werden seltener ($.093$, $SD = .128$) als konkrete, invalide Aufgaben ($.876$, $SD = .161$) gelöst.

Klasse 4

In dieser im Vergleich größten Klasse (147 Teilnehmerinnen und Teilnehmer; 37,6%) konnten MP-Aufgaben häufig gelöst werden, knapp gefolgt von NA und AK Aufgaben. MT-Aufgaben werden hier am wenigsten oft gelöst. Bei den validen Aufgaben zeigt sich eine mittlere bis bereits höhere Lösungswahrscheinlichkeit ($.680$, $SD = .186$) und bei den invaliden Aufgaben eine ähnliche Lösungswahrscheinlichkeit ($.670$, $SD = .190$); die Berechnung der entsprechenden Prüfgröße des t-Tests für verbundene Stichproben fällt mit $t(146) = 0.70$, $p = .488$ nicht signifikant aus.

Klasse 5

Personen in dieser Lösungsklasse (102 Teilnehmerinnen und Teilnehmer; 26,1%) können MP-Aufgaben fast zur Gänze lösen. Auch bei den Aufgaben des Typs NA und AK weisen sie hohe Lösungswahrscheinlichkeiten auf. MT-Aufgaben löst diese Klasse hingegen am wenigsten häufig. Bei den validen Aufgaben zeigen sich mittlere bis bereits höhere Lösungswahrscheinlichkeiten ($.730$, $SD = .242$) und bei den invaliden Aufgaben eine relativ hohe Lösungswahrscheinlichkeit ($.936$, $SD = .128$); $t(101) = -9.99$, $p < .001$. Abstrakte, invalide Aufgabeninhalte konnten in dieser Lösungsklasse besser ($.930$, $SD = .136$) als abstrakte, valide Aufgabeninhalte ($.718$, $SD = .272$) gelöst werden.

Tabelle 12

Lösungswahrscheinlichkeiten (M, SD) in den Argument- und Inhaltsbedingungen je Klasse

Klasse (Häufigkeit)		val	inval	konk val	konk inval	abs val	abs inval	kontra val	kontra inval
1 (97)	<i>M</i>	.764	.214	.699	.275	.810	.125	.780	.232
	<i>SD</i>	.163	.173	.282	.271	.190	.195	.178	.205
2 (30)	<i>M</i>	.223	.544	.127	.552	.407	.594	.150	.483
	<i>SD</i>	.117	.117	.207	.149	.083	.143	.160	.148
3 (15)	<i>M</i>	.383	.544	.093	.876	.787	.078	.289	.622
	<i>SD</i>	.115	.125	.128	.161	.177	.124	.305	.292
4 (147)	<i>M</i>	.680	.670	.676	.745	.679	.566	.684	.686
	<i>SD</i>	.186	.190	.224	.216	.219	.278	.220	.222
5 (102)	<i>M</i>	.730	.936	.724	.944	.718	.930	.745	.933
	<i>SD</i>	.242	.128	.267	.141	.272	.136	.226	.160
Gesamt (391)	<i>M</i>	.667	.612	.630	.671	.705	.535	.668	.620
	<i>SD</i>	.240	.309	.309	.326	.240	.369	.274	.326

Anmerkung. Argumentform: val = valide, inval = invalide; Inhaltsform: konk = konkret, abs = abstrakt, kontra = kontrafaktisch.

Darüber hinaus wurde die mittlere Lösungswahrscheinlichkeit in Abhängigkeit der Argumentformen MP, MT, NA und AK ermittelt.

Tabelle 13

Lösungswahrscheinlichkeit der vier Argumentformen (n = 391)

	MP	MT	NA	AK
<i>M (SD)</i>	.867 (.242)	.512 (.332)	.679 (.326)	.563 (.344)

Tabelle 13 gibt die mittleren Lösungswahrscheinlichkeiten in Abhängigkeit der vier Argumentformen wieder. Die Prüfgröße der einfaktoriellen ANOVA mit Messwiederholung (Sphärizität, geprüft mittels Mauchly-Test, konnte mit $p < .001$ nicht angenommen werden, wodurch die Korrektur nach Greenhouse-Geisser mit $\varepsilon = .736$ erforderlich war) fiel mit $F(2.21, 861.37) = 131.38, p < .001 (\eta^2_p = .252)$ signifikant aus. Die Berechnung von einfachen a-priori Kontrasten für den Vergleich der nachfolgenden Kategorie gegen die erste (MP) ergab

jeweils mit $p < .001$ signifikante Unterschiede. MP ist demgemäß mit einer mittleren Lösungswahrscheinlichkeit von 86,7% im Vergleich die am leichtesten zu lösende Argumentform.

8.3.2. Lösungswahrscheinlichkeit SDV unter Berücksichtigung der Klassenzugehörigkeit und Negation

Es wurde geprüft, ob und inwieweit sich die Lösungswahrscheinlichkeit der Syllogismen zwischen den beiden Negationsformen *unnegiert* und *negiert* in Abhängigkeit der Lösungsklassen unterscheidet (siehe Tabelle 14).

Tabelle 14

Deskriptivstatistische Kennwerte zu den mittleren Lösungswahrscheinlichkeiten

Moderator	Lösungsklasse	<i>M</i>	<i>SD</i>	N
unnegiert	1	.5019	.1384	97
	2	.3684	.1114	30
	3	.5298	.0674	15
	4	.7533	.1625	147
	5	.8684	.1468	102
	Gesamt	.6829	.2209	391
Moderator	Lösungsklasse	<i>M</i>	<i>SD</i>	N
negiert	1	.4227	.1406	97
	2	.4313	.1117	30
	3	.4000	.0660	15
	4	.5808	.1960	147
	5	.8100	.1794	102
	Gesamt	.5830	.2272	391

Die Beurteilung der Wechselwirkung aus Klassenzugehörigkeit x Negation fiel mit $F(4, 386) = 35.47$, $p < .001$ ($\eta^2_p = .269$) signifikant aus; es kann eine Interaktion der Klassenzugehörigkeit in Verbindung mit der Negation auf die Lösungswahrscheinlichkeit SDV angenommen werden. Daher ist die Interpretation des signifikanten Haupteffekts des Moderators Negation mit $F(1, 386) = 88.30$, $p < .001$ ($\eta^2_p = .186$) bezüglich einer Abnahme der Lösungs-

wahrscheinlichkeit von unnegiert ($M = .6829$, $SD = .2209$) zu negiert ($M = .5830$, $SD = .2272$) zu relativieren. Die Berechnung von nachfolgenden t-Tests für verbundene Stichproben ergab, dass die Lösungsklassen 1, 3, 4 und 5 eine signifikante Abnahme der Lösungswahrscheinlichkeiten aufweisen (p 's $< .001$), während für die Lösungsklasse 2 mit $p = .004$ eine Zunahme der Lösungswahrscheinlichkeit bei Vorliegen einer Negation angenommen werden kann. Die Tabelle 15 zeigt die Kennwerte und Prüfgrößen der post-hoc t-Testungen.

Tabelle 15

Kennwerte der Differenz der Lösungswahrscheinlichkeit von unnegiert zu negiert mit Prüfgrößen

Lösungsklasse	Differenz (M)	SD	t	df	p
1	.079	.125	6.22	96	<.001
2	-.063	.108	-3.18	29	.004
3	.130	.095	5.28	14	<.001
4	.173	.127	16.46	146	<.001
5	.058	.075	7.89	101	<.001

Der Zwischensubjektfaktor Lösungsklasse fiel mit $F(4, 386) = 106.45$, $p < .001$ ($\eta^2_p = .525$) signifikant aus; die einzelnen Lösungsklassen zeigen unterschiedliche Niveaus in der Lösungswahrscheinlichkeit SDV. Die Abbildung 5 zeigt die Lösungswahrscheinlichkeiten SDV der 5 Lösungsklassen unter Berücksichtigung des Moderators Negation.

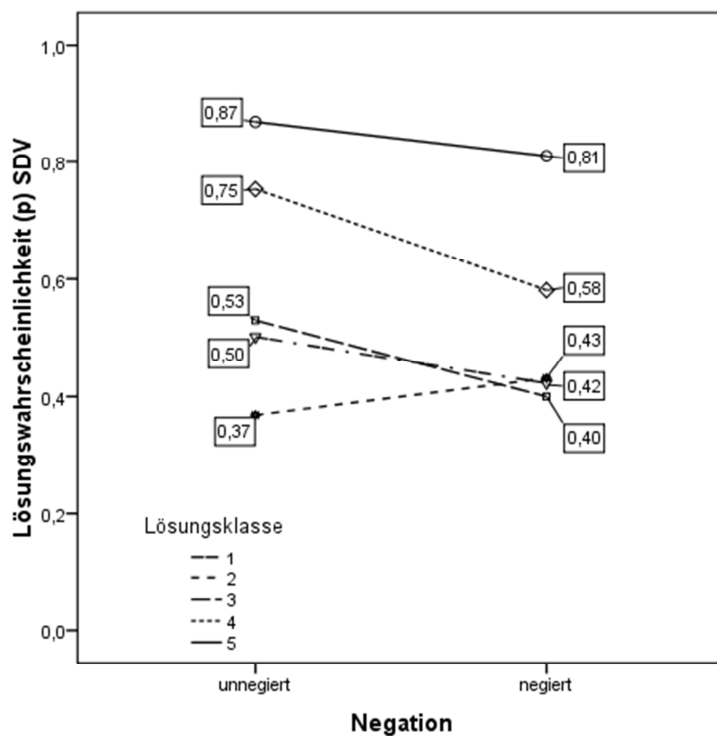


Abbildung 5. Lösungswahrscheinlichkeiten SDV in Abhängigkeit von Lösungsklasse und Negation.

Lösungsklasse und Studienrichtung

Von Interesse war, ob die beiden Studienrichtungen ein unterschiedliches Niveau in der Lösungsklasse aufweisen bzw. erreichen. In der Tabelle 16 finden sich die entsprechenden Häufigkeiten und Anteilswerte je Lösungsklasse.

Tabelle 16

Kontingenztafel mit Häufigkeiten und Anteilswerten der Lösungsklasse je Studienrichtung

Studienrichtung		Lösungsklasse					Gesamt
		1	2	3	4	5	
Psychologie	Anzahl	68	2	12	79	40	201
	% innerhalb	33,8%	1,0%	6,0%	39,3%	19,9%	100,0%
	Stand. Residuum	2.6	-3.4	1.5	0.4	-1.7	
Naturwissen- schaften	Anzahl	29	28	3	68	62	190
	% innerhalb	15,3%	14,7%	1,6%	35,8%	32,6%	100,0%
	Stand. Residuum	-2.6	3.5	-1.6	-0.4	1.8	
Gesamt	Anzahl	97	30	15	147	102	391
	% innerhalb	24,8%	7,7%	3,8%	37,6%	26,1%	100,0%

Die Prüfgröße fiel mit $\chi^2(4) = 48.91$, $p < .001$ signifikant aus; es kann ein Verteilungsunterschied bezüglich der Lösungsklassen bei den Studienrichtungen angenommen werden. Studierende der Psychologie weisen unter Berücksichtigung der standardisierten Residuen häufiger die Lösungsklasse 1 auf, während wiederum Studierende der Naturwissenschaften häufiger die Lösungsklasse 2 erreichen. In der Lösungsklasse 3 sind Studierende der Psychologie tendenziell häufiger vertreten, während die Lösungsklasse 5 von Studierenden der Naturwissenschaften tendenziell häufiger erreicht wird. In der Lösungsklasse 4 unterscheiden sich die Anteilswerte der beiden Studienrichtungen nicht.

Darüber hinaus wurden die erreichten Lösungsklassen auch getrennt für die Geschlechter untersucht. In der Tabelle 17 finden sich die entsprechenden Häufigkeiten und Anteilswerte je Lösungsklasse, unter Berücksichtigung des Geschlechts und der Studienrichtung.

Tabelle 17

Kontingenztafel mit Häufigkeiten und Anteilswerten der Lösungsklasse je Studienrichtung und Geschlecht

Geschlecht Studienrichtung			Lösungsklasse					Gesamt
			1	2	3	4	5	
weiblich	Psychologie	Anzahl	51	2	9	51	27	140
		% innerhalb	36,4%	1,4%	6,4%	36,4%	19,3%	100,0%
	Naturwissenschaften	Anzahl	7	6	0	13	15	41
		% innerhalb	17,1%	14,6%	,0%	31,7%	36,6%	100,0%
	Gesamt	Anzahl	58	8	9	64	42	181
		% innerhalb	32,0%	4,4%	5,0%	35,4%	23,2%	100,0%
männlich	Psychologie	Anzahl	17	0	3	28	13	61
		% innerhalb	27,9%	,0%	4,9%	45,9%	21,3%	100,0%
	Naturwissenschaften	Anzahl	22	22	3	55	47	149
		% innerhalb	14,8%	14,8%	2,0%	36,9%	31,5%	100,0%
	Gesamt	Anzahl	39	22	6	83	60	210
		% innerhalb	18,6%	10,5%	2,9%	39,5%	28,6%	100,0%

Die Prüfgrößen fielen sowohl für weibliche Studierende mit $\chi^2(4) = 23.15$, $p < .001$ als auch für männliche Studierende mit $\chi^2(4) = 16.76$, $p = .002$ signifikant aus; es kann jeweils dieser

Verteilungsunterschied bezüglich der Lösungsklassen auch bei den Geschlechtern angenommen werden.

8.4. Zusammenhang der Genderinformation im Genderitem mit der Testleistung

Zur Operationalisierung der im SDV vorgegebenen Syllogismen wurden diese vorab nach ihrem semantischen Aufgabeninhalt, Argumentform und Negation, wie in Tabelle 18 ersichtlich, kodiert.

Tabelle 18

Kodierung der Items

Semantischer Aufgabeninhalt	Argumentform	Negationen
konkret = 1	MP = 1	unnegiert = 1
abstrakt = 2	MT = 2	negiert = 2
kontrafaktisch = 3	NA = 3	
	AK = 4	

Zum Beispiel repräsentiert die Kodierung 311 ein Item mit kontrafaktischem Aufgabeninhalt, Argumentform MP und keiner Negation (unnegiert). Als Itemformulierung für diese spezielle Kombination von Moderatorvariablen sei das Item 1 „Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note. Maria hat gute Laune.“ angeführt.

Es wurde zunächst eine Reliabilitätsanalyse durchgeführt, um zu untersuchen, ob die hinzugefügten Genderitems sich ebenfalls konsistent verhalten und das Gleiche messen. Der Reliabilitätskoeffizient für die dichotomen Items ergab mit $\alpha_{20} = .895$ ($n = 391$) nach Kuder Richardson (KR-20) den Hinweis auf eine ausreichend hohe Messgenauigkeit. Die Tabelle 23 im Anhang 1 zeigt die korrigierten Trennschärfen und Koeffizienten nach Cronbach's Alpha unter Auslassung des entsprechenden Items. Für die Items 5, 15 und 22 wurden korrigierte Trennschärfen $< .20$ beobachtet, womit jedoch für die Gesamtskala keine Verminderung der Konsistenz anzunehmen ist. Die drei Genderitems (1, 11 und 21) zeigten Trennschärfen $\geq .436$. Die 35 Items bilden eine ausreichend konsistente und zuverlässige Skala, die homogen misst (vergleiche Bortz & Döring, 2006). Die Genderitems verhalten sich, unter Berücksichtigung der Ratewahrscheinlichkeit [0.33], im Vergleich zu den anderen Items mittelschwierig.

Im Folgenden wird der Effekt der Genderinformation in spezifischen Items auf die Lösungswahrscheinlichkeit unter Berücksichtigung des Geschlechts der Studierenden analysiert. Zur Berechnung wurden für die drei Genderitems (Item 1, Item 11 und Item 21) jeweils eine Vierfeldertafel erstellt und zwar mit dem Lösungsverhalten (gelöst/nicht gelöst) als abhängige Variable und Geschlecht (m/w) als unabhängige Variable.

Item 1 (Kontrafaktisch/MP/unnegiert)

Tabelle 19

Item 1, Genderitem nach Geschlecht

Geschlecht		Item 1. 311		Gesamt
		nicht gelöst	gelöst	
weiblich	Anzahl	38	146	184
	Erwartete Anzahl	37.7	146.3	184.0
	% innerhalb von Geschlecht	20,7%	79,3%	100,0%
	Standardisierte Residuen	.0	.0	
männlich	Anzahl	43	168	211
	Erwartete Anzahl	43.3	167.7	211.0
	% innerhalb von Geschlecht	20,4%	79,6%	100,0%
	Standardisierte Residuen	.0	.0	
Gesamt	Anzahl	81	314	395
	Erwartete Anzahl	81.0	314.0	395.0
	% innerhalb von Geschlecht	20,5%	79,5%	100,0%

In Tabelle 19 ist ersichtlich, dass 146 (79,3%) von 184 Frauen dieses Item (1) lösen konnten und von den insgesamt 211 männlichen Studierenden 168 (79,6%). Es kann hier kein Verteilungsunterschied in Bezug auf die Lösungshäufigkeit in Abhängigkeit vom Geschlecht beobachtet werden, die Prüfgröße fiel mit $\chi^2(1) = 0.004$, $p = .947$ nicht signifikant aus. Für beide Geschlechter war dieses Item gleich leicht lösbar.

Item 11 (Kontrafaktisch/NA/unnegiert)

Tabelle 20

Item 11, Genderitem nach Geschlecht

Geschlecht		Item 11. 331		Gesamt
		nicht gelöst	gelöst	
weiblich	Anzahl	45	139	184
	Erwartete Anzahl	41.0	143.0	184.0
	% innerhalb von Geschlecht	24,5%	75,5%	100,0%
	Standardisierte Residuen	.6	-.3	
männlich	Anzahl	43	168	211
	Erwartete Anzahl	47.0	164.0	211.0
	% innerhalb von Geschlecht	20,4%	79,6%	100,0%
	Standardisierte Residuen	-.6	.3	
Gesamt	Anzahl	88	307	395
	Erwartete Anzahl	88.0	307.0	395.0
	% innerhalb von Geschlecht	22,3%	77,7%	100,0%

Dieses Item (11) konnte von 139 (75,5%) der insgesamt 184 weiblichen Studierenden gelöst werden und von 168 (79,6%) der insgesamt 211 männlichen Studierenden (siehe Tabelle 20). Auch hier kann kein signifikanter Verteilungsunterschied in der Lösungshäufigkeit in Abhängigkeit vom Geschlecht angenommen werden, $\chi^2(1) = 0.944$, $p = .331$. Dieses Item war ebenso sowohl für Frauen als auch Männer gleich leicht zu lösen.

Item 21 (Kontrafaktisch/AK/unnegiert)

Tabelle 21

Item 21, Genderitem nach Geschlecht

Geschlecht		Item 21. 341		Gesamt
		nicht gelöst	gelöst	
weiblich	Anzahl	70	111	181
	Erwartete Anzahl	59.3	121.7	181.0
	% innerhalb von Geschlecht	38,7%	61,3%	100,0%
	Standardisierte Residuen	1.4	-1.0	
männlich	Anzahl	58	152	210
	Erwartete Anzahl	68.7	141.3	210.0
	% innerhalb von Geschlecht	27,6%	72,4%	100,0%
	Standardisierte Residuen	-1.3	.9	
Gesamt	Anzahl	128	263	391
	Erwartete Anzahl	128.0	263.0	391.0
	% innerhalb von Geschlecht	32,7%	67,3%	100,0%

Dieses Item (21) weist Unterschiede in der Lösungshäufigkeit auf; die Prüfgröße fiel mit $\chi^2(1) = 5.395$, $p = .020$ signifikant aus. 61,3% der weiblichen Studierenden und 72,4% der männlichen Studierenden lösten dieses Item. Männliche Studierende lösen dieses Item somit häufiger (siehe Tabelle 21). Somit muss die gerichtete Alternativhypothese $H_1^{(3a)}$, dass Frauen dieses Item häufiger lösen, verworfen werden.

Um dieses Ergebnis interpretieren zu können, ist prinzipiell die Unterschiedlichkeit des Lösungsverhaltens bei AK-Items zwischen Männern und Frauen von Interesse. Die Berechnung eines t-Test für unabhängige Stichproben ergab mit $t(389) = -3.99$, $p < .001$ ($d = 0.40$) ein signifikantes Ergebnis, womit Männer ($M = .6264$, $SD = .3338$) die AK-Argumentform häufiger lösen als Frauen ($M = .4897$, $SD = .3424$).

Die Lösungswahrscheinlichkeit aller drei Genderitems steht in einem positiven Zusammenhang mit der SDV-Testleistung. Die Berechnung des Korrelationskoeffizienten der Produkt-Moment-Korrelation nach Pearson fiel mit $r = .649$ ($p < .001$, einseitig, $n = 391$) signifikant aus. Die Hypothese $H_1^{(3b)}$ kann demgemäß angenommen werden.

Reihenfolgeeffekt

Um mögliche Reihenfolgeeffekte zu überprüfen, wurden für die beiden Vorgabenversionen A und B sowohl für die Variante mit Genderitems (35 Items) und ohne Genderitems (32 Items) jeweils t-Tests für unabhängige Stichproben berechnet (siehe Tabelle 22). Für die Berechnung bildeten die gemittelten SDV Summenscores (*mean scores*) die abhängigen Variablen und die beiden Vorgabeversionen die Stufen der unabhängigen Variablen. Es sollte geprüft werden, ob die beiden Testformen A und B bezüglich ihres Syllogismenscores unterschiedlich ausfallen.

Tabelle 22

Mittelwerte für beide Testversionen A und B mit 35 und 32 Items

	Vorgabereihenfolge	N	M	SD
SDV Meanscore (35 Items)	A	222	.6413	.2216
	B	173	.6376	.2056
SDV Meanscore (32 Items)	A	222	.6336	.2182
	B	173	.6241	.2081

Die entsprechenden Prüfgrößen fielen sowohl für den SDV-Meanscore (35 Items) mit $t(393) = 0.17$, $p = .865$ als auch für den SDV-Meanscore (32 Items) mit $t(393) = 0.44$, $p = .662$ jeweils nicht signifikant aus. Die beiden Vorgabeversionen A und B unterscheiden sich nicht bezüglich der Lösungswahrscheinlichkeit, die unterschiedliche Positionierung der Genderitems wirkte sich nicht auf die Testleistung in den beiden Versionen aus.

8.5. Einfluss des Fähigkeitsselbstkonzepts weiblicher Studierender auf das Lösen der Genderitems

Im Rahmen der Erhebung wurde die Frage „Wie gut glauben Sie, solche Aufgaben lösen zu können?“ gestellt. Nach der Testbearbeitung war dies erneut einzuschätzen, wobei die Formulierung nunmehr lautete: „Wie gut konnten Sie diese Aufgaben lösen?“ Das Antwortformat lag auf einer vierstufigen Selbsteinschätzungsskala vor, wobei (1) *gut* bis (4) *schlecht* darstellt. Die Prüfung auf Veränderung mittels t-Test für abhängige Stichproben ergibt mit $t(388) = -7.78$, $p < .001$ ($d = 0.71$) ein signifikantes Resultat. Es kann eine niedrigere Selbsteinschätzung nach der Testbearbeitung angenommen werden. Die Hypothese $H_1^{(4a)}$ kann angenommen werden.

Die entsprechende Korrelation für die Stabilität der Selbsteinschätzung vor und nach der Testvorgabe fiel mit $r = .533$ ($p < .001$, $n = 389$) signifikant aus und zeigt einen mäßig hohen Zusammenhang in der Selbsteinschätzung. Getrennt nach Geschlecht fallen die Korrelationskoeffizienten für Frauen mit $r = .355$ und Männer mit $r = .594$ jeweils signifikant ($p < .001$) aus.

Die Prüfung der Hypothese $H_1^{(4b)}$ wurde mittels einer zweifaktoriellen Varianzanalyse mit Messwiederholung, unter der Berücksichtigung des Zwischensubjektfaktors Geschlecht vorgenommen. Geprüft wurde der Unterschied in der Selbsteinschätzung in Abhängigkeit vom Geschlecht und die Wechselwirkung aus der Faktorstufenkombination Geschlecht x Erhebungszeitpunkt. Die Homogenität der Varianzen konnte zu den beiden Erhebungszeitpunkten jeweils angenommen werden ($p > .05$).

Die Berechnung der entsprechenden Prüfgröße für den Zwischensubjektfaktor Geschlecht fiel mit $F(1, 387) = 46.57$, $p < .001$ ($\eta^2_p = .107$) signifikant aus. Die Prüfgröße für die Wechselwirkung aus Geschlecht x Erhebungszeitpunkt fällt mit $F(1, 387) = 2.38$, $p = .124$ ($\eta^2 = .006$) nicht signifikant aus. Abbildung 6 zeigt die Selbsteinschätzung in Abhängigkeit von Erhebungszeitpunkt und Geschlecht.

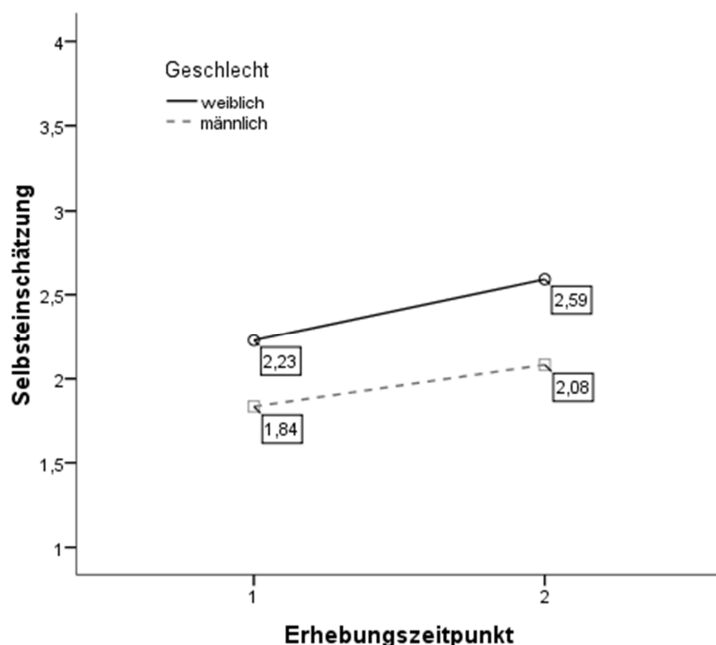


Abbildung 6. Selbsteinschätzung in Abhängigkeit von Erhebungszeitpunkt und Geschlecht.

Die Hypothese $H_1^{(4b)}$ kann angenommen werden. Es ist festzuhalten, dass zwischen den Geschlechtern ein Unterschied dahingehend anzunehmen ist, dass sich Männer zu beiden Zeitpunkten besser einschätzen.

Die Selbsteinschätzung wurde vor und nach der Testung mit den vier Stufen (1) *gut* bis (4) *schlecht* erhoben; der Zusammenhang der Einschätzung vor der Testleistung fiel mit $r = -.305$ ($p < .001$, einseitig, $n = 391$) und danach mit $r = -.365$ ($p < .001$, einseitig, $n = 389$) jeweils signifikant aus. Es kann sowohl vor als auch nach der Testung mit den Syllogismen ein mäßiger Zusammenhang mit der Selbsteinschätzung beobachtet werden. Personen mit einer ausgeprägteren Selbsteinschätzung weisen eine höhere Testleistung auf. Die Hypothese $H_1^{(4c)}$ kann angenommen werden.

9. Diskussion

Hauptziel dieser Studie war, unter dem Aspekt des Geschlechterstereotyps, die Kompetenz schlussfolgernden Denkens von Studierenden im zweiten Studienabschnitt der Studienrichtungen Mathematik, Physik, Chemie und Psychologie zu untersuchen. Es handelte sich dabei um eine Erweiterung der Arbeiten von Klopff (2011) und Eder (2012), wobei die Studien von Spiel, et al. (2001, 2004) die Basis bildeten. Der Schwerpunkt lag dabei auf den im SDV eigens hinzugefügten Genderitems sowie der Extrahierung und Zuordnung von Lösungsklassen im Sinne der Stufentheorie von Piaget (1976b, 1984).

Die Vorgabe des SDV erfolgte an der Universität Wien sowie der Technischen Universität Wien bei 395 Studierenden mit einem Durchschnittsalter von 24.05 ($SD = 4.78$) Jahren. Erwartungsgemäß fanden sich hauptsächlich weibliche Studierende im Studienfach Psychologie (69,8%) wieder, und die übrigen 21,6% verteilten sich auf die Studienrichtungen Mathematik, Physik und Chemie (die aufgrund ihrer inhaltlichen Nähe als naturwissenschaftliche Studienrichtungen zusammengefasst wurden). Die Verteilung ergab bei den männlichen Studierenden 30,2% im Studienfach Psychologie und 78,4% in den klassischen naturwissenschaftlichen Studienrichtungen. Somit entsprachen diese Anteilswerte in der Stichprobe nur annähernd der Verteilung an inskribierten Studentinnen und Studenten, wie sie von der Technischen Universität Wien und Universität Wien angegeben wurden. Da die Kenntnis der deutschen Sprache notwendig ist, um die Instruktion des SDV zu verstehen, wurde die Erstsprache miterhoben. 9,1% der Testpersonen hatten eine andere Erstsprache als Deutsch.

Es zeigte sich, dass die Erstsprache keinen Einfluss auf die Lösungswahrscheinlichkeit im SDV hatte. Somit konnte der Faktor Sprache als konfundierende Variable ausgeschlossen werden.

Festzuhalten ist, dass die SDV-Itemschwierigkeiten im Sinne der klassischen Testtheorie über die Lösungswahrscheinlichkeit definiert wurden und daher eine Stichprobenabhängigkeit anzunehmen ist.

Zur SDV-Testleistung bezüglich der Studienrichtung konnte gezeigt werden, dass die Lösungswahrscheinlichkeit im SDV Unterschiede aufweist. Wie angenommen lösten, wie bereits Klopff (2011) bei Studienanfängerinnen und Studienanfängern zeigte, Studierende der naturwissenschaftlichen Studienfächer (Mathematik, Physik und Chemie) im zweiten Studienabschnitt bzw. Masterstudium die Items im SDV häufiger als Studierende der Studienrichtung Psychologie. Darüber hinaus konnte belegt werden, dass keine Unterschiede in der Lösungswahrscheinlichkeit in Abhängigkeit vom Geschlecht auftraten. Diese Interpretation der Ergebnisse zu den Haupteffekten konnte ohne Einschränkung vorgenommen werden, da keine Interaktionseffekte zwischen Studienrichtung und Geschlecht auftraten.

Um zu überprüfen, ob und inwieweit die Studienleistung in einem Zusammenhang mit der Lösungswahrscheinlichkeit der Items steht, wurde der Durchschnitt dreier relevanter Noten aus dem ersten Diplomprüfungszeugnis der jeweiligen Studienrichtung herangezogen. Dieser Indikator für die Studienleistung geht mit der Lösungswahrscheinlichkeit im SDV mit $r = -.210$ einher. Es konnte ein schwacher Zusammenhang belegt werden, dass bessere Noten mit einer höheren Lösungswahrscheinlichkeit einhergehen.

Auf Grundlage der WINMIRA[®]-Prozedur konnten fünf Lösungsklassen, sowohl unter Berücksichtigung von 35 Items als auch 32 Items (ohne Genderitems), extrahiert werden. Die Kategorisierung der Fälle zu einer 5-Klassenlösung auf Basis der WINMIRA[®]-Prozedur kann aufgrund der hohen Zuordnungssicherheit (durchschnittlich 93,20% bei 35 Items und 95,98% bei 32 Items) als sinnvoll bezeichnet werden.

Die in der vorliegenden Untersuchung erreichte Klassenanzahl korrespondiert weitestgehend mit den Untersuchungen von Spiel, et al. (2001, 2004), welche neben einem konkret-operationalen Stadium ein Übergangsstadium 1 und 2 identifizieren konnten sowie Klopff (2011) und Eder (2012), welche noch ein zusätzliches Übergangsstadium 3 identifizierten. Darüber hinaus kann auf Grundlage dieser Arbeit ein weiteres Übergangsstadium 4 angenommen werden.

Im Folgenden werden diese Lösungsklassen näher beschrieben.

Lösungsklasse 1 (Konkret-operationales Stadium)

Dieses Stadium wurde sowohl bei Spiel, et al. (2001, 2004) als auch bei Klopff (2011) und Eder (2012) dem konkret-operationalen Stadium zugeordnet. Auch in vorliegender Arbeit zeichnet sich die Lösungsklasse 1 durch ein bikonditionales Antwortmuster aus und kann ebenso dem konkret-operationalen Stadium zugeordnet werden. Dabei werden unabhängig von Negation und Inhalt valide Aufgaben (MP und MT) häufiger gelöst als invalide Aufgaben (NA und AK).

Lösungsklasse 2 und 3 (Übergangsstadium 1 und 2)

Lösungsklasse 2 und 3 zeigen in gewissen Kombinationen von Aufgabeninhalten ein Lösungsverhalten, welches sich von den anderen Lösungsklassen (1, 4 und 5) deutlich unterscheidet. Diese beiden Klassen können nicht direkt mit dem Übergangsstadium 1 von Klopff (2011) und Eder (2012) verglichen werden. Dies führt dazu, dass diese beiden Klassen differenziert zu beschreiben sind, da sie über die Inhaltsformen hinweg heterogene Lösungswahrscheinlichkeiten zeigen.

Studierende in Lösungsklasse 2 erzielen bei den validen Aufgaben schlechte Leistungen. Bei den NA-Aufgaben können vor allem konkrete aber auch kontrafaktische Items häufiger gelöst werden als abstrakte. Bei den AK-Aufgaben werden abstrakte Items häufiger gelöst als konkrete und kontrafaktische.

Lösungsklasse 3 bildet vergleichsweise die kleinste Klasse. Bei den validen Aufgaben (MP und MT), und hier vor allem bei abstrakten Inhalten, können gute Leistungen erzielt werden; konkrete und kontrafaktische Aufgaben werden hingegen selten gelöst. Das genaue Gegenteil repräsentiert sich bei den invaliden Aufgaben, bei denen sowohl die konkreten als auch die kontrafaktischen Items häufig gelöst werden, die abstrakten hingegen kaum.

Aufgrund der häufigeren Lösungswahrscheinlichkeit der invaliden Aufgaben zeigen Personen in diesen beiden Klassen erste Kompetenzen im deduktiven Schlussfolgern und sollen somit nicht mehr dem konkret-operationalen Stadium zugeordnet werden. Sie bilden folglich die ersten beiden Übergangsstadien, 1 und 2, zum formal-operationalen Stadium.

Lösungsklasse 4 (Übergangsstadium 3)

Verglichen mit den Lösungsklassen 2 und 3 werden hier die Aufgaben über die Argument- und Inhaltsformen hinweg wieder gleichmäßiger gelöst. Hohe Lösungswahrscheinlichkeiten ergeben sich bei MP-Aufgaben. Auch die Lösung der invaliden Aufgaben ist bereits häufiger möglich. Am wenigsten häufig werden MT-Aufgaben gelöst. Dieser Leistungseinbruch bei der Argumentform MT sowie das unabhängigere Lösungsverhalten über die Inhalte hinweg deuten auf einen Entwicklungsfortschritt im deduktiven Schlussfolgern hin. Diese Beobachtung geht auch mit der Literatur einher, wonach ein Leistungsabfall bei der Argumentform MT als Entwicklungsfortschritt interpretiert wird (Byrnes & Overton, 1986; Janveau-Brennan & Markovits, 1999; Kuhn, 1977; Spiel, et al., 2004; Wildman & Fletcher, 1977); ebenso wie die Nivellierung der Lösungswahrscheinlichkeiten über den Inhalt hinweg (Markovits & Vachon, 1990). Somit werden Personen, die sich in dieser Lösungsklasse befinden, dem Übergangsstadium 3 zugeordnet.

Lösungsklasse 5 (Übergangsstadium 4)

Lösungsklasse 5 ist gekennzeichnet durch eine hohe bis sehr hohe Lösungswahrscheinlichkeit bei MP, NA und AK Aufgaben. Bei MT-Aufgaben werden, wie auch in Lösungsklasse 4, schlechtere Leistungen erzielt. Das Antwortmuster dieser Klasse weist im Vergleich das jeweils höchste Lösungsverhalten auf. Auch speziell für negierte Bedingungen weist diese Klasse die höchsten Werte auf.

Verglichen mit der Lösungsklasse 4 von Klopff (2011) und Eder (2012), bei welchen bei den invaliden Aufgaben und MP eine durchschnittliche mittlere Lösungswahrscheinlichkeit von .88 ($SD = .078$) erreicht wurde, wurde in Lösungsklasse 5 vorliegender Untersuchung mit einer durchschnittlichen Lösungswahrscheinlichkeit von .9521 ($SD = .0944$) ein höherer Wert beobachtet, während die Argumentform MT hier mit .5229 ($SD = .4284$) durchschnittlich ausfiel. Darüber hinaus haben 33 Personen (32,4%) in dieser höchsten Lösungsklasse alle 35 Aufgaben lösen können. In den darunter liegenden Übergangsstadien konnte keine Person alle Aufgaben lösen. In der vorliegenden Untersuchung zeigte sich selbst in der höchsten Lösungsklasse 5 zwar mit .9521 ($SD = .0944$) ein hoher Wert für MP-, NA- und AK-Aufgaben, während jedoch die Argumentform MT nur .5229 ($SD = .4284$) erreichte. Dies kann erklären, dass das Stadium der formalen Operationen, welches durch eine sehr hohe Lösungswahrscheinlichkeit über alle Argument- und Inhaltsformen hinweg gekennzeichnet wäre, wider Erwarten, wie schon bei Klopff (2011) und Eder (2012), nicht identifiziert werden konnte. Somit werden Personen der

Lösungsklasse 5 dem Übergangsstadium 4 zugeordnet. Insgesamt, wie auch bei Spiel, et al. (2001, 2004) sowie Klopff (2011) und Eder (2012) konnte die Argumentform MP mit einer mittleren Lösungswahrscheinlichkeit von $M = .867$ ($SD = .242$) am häufigsten gelöst werden; gefolgt von den Argumentformen NA mit $M = .679$ ($SD = .326$), AK mit $M = .563$ ($SD = .344$) und MT mit $M = .512$ ($SD = .332$).

Eine interessante Beobachtung ergibt sich bei kontrafaktisch invaliden Bedingungen, dass die entsprechende Lösungswahrscheinlichkeit nur in dieser Aufgabenart mit der Klassenzugehörigkeit aufsteigend einhergeht und dass mit dieser Korrespondenz eine schnelle und ökonomische Zuordnungsmöglichkeit zu den Klassen vorgeschlagen werden kann. Je höher die Lösungswahrscheinlichkeit bei kontrafaktisch invaliden Bedingungen ausfällt, desto höher ist auch das erreichte Stadium. Diese Beobachtung stimmt auch mit jener von Schröder (1989) überein, dass es zu einem kontinuierlichen Anstieg der Performanz beim Lösen kontrafaktischer Aufgaben in den Stadien kommt.

Es wurde zudem beobachtet, dass das Erreichen der Lösungsklassen mit den Studienrichtungen zusammenhängt. Studierende der Naturwissenschaften erreichen gegenüber Studierenden der Psychologie häufiger ein höheres Stadium; dieser Verteilungsunterschied ist unabhängig vom Geschlecht. Es ist naheliegend, dass die beobachteten Verteilungsunterschiede zwischen den Studienrichtungen nicht auf das Geschlecht zurückzuführen sind und insofern den in der Gesellschaft vertretenen Geschlechterstereotyp nicht bestätigt.

Im SDV wurden neben unnegierten Syllogismen auch negierte Items vorgegeben. Bei Negationen handelt es sich um eine Moderatorvariable welche die Aufgabenschwierigkeit erhöht und bereits im konkret-operationalen Stadium gelöst werden sollte. Die Ergebnisse weisen einerseits darauf hin, dass nur für Personen im Übergangsstadium 1 (Lösungsklasse 2) negierte Aufgaben leichter lösbar sind, während Personen in allen anderen Stadien die unnegierten Aufgaben leichter fallen und andererseits, dass im konkret-operationalem Stadium die Lösungswahrscheinlichkeit von negierten Items mit 42,27% durchschnittlich hoch ausfällt. Dieses Ergebnis stimmt überwiegend mit jenen von Moshman (1977), Wildman und Fletcher (1977) und Roberge und Mason (1978) überein, bei welchen negierte Aufgaben schwieriger lösbar waren als unnegierte. Eine Ausnahme in vorliegender Untersuchung bildet, wie bereits erwähnt, Lösungsklasse 2.

Dem SDV wurden eigens drei Genderitems hinzugefügt und somit der SDV von 32 auf 35 Items erweitert. Bei diesen drei Genderitems handelt es sich um die Formen MP, NA und AK, jeweils

kontrafaktisch und unnegiert. Im Vergleich zu den anderen Items sind die Genderitems reliabel und erweisen sich als mittelschwierig. Bei der Interpretation des Zusammenhanges zwischen Trennschärfe und Lösungswahrscheinlichkeit der Genderitems können diese als unauffällig im Vergleich mit den übrigen Items beschrieben werden. Entgegen der Annahme waren beim Genderitem 1 (MP) und 11 (NA) keine Unterschiede zwischen den Geschlechtern festzustellen. Es kann festgehalten werden, dass diese beiden Items für beide Geschlechter gleich schwierig waren. Hingegen wurde das Genderitem 21 (AK) von 72,4% der Männer und von 61,3% der Frauen gelöst. Dieser Unterschied stellte sich als signifikant heraus und weist darauf hin, dass Männer dieses Item häufiger lösten als Frauen. Somit kann einerseits angenommen werden, dass die inkludierte Genderinformation keine Auswirkung auf die Lösungswahrscheinlichkeit hat und dass andererseits die höhere Lösungswahrscheinlichkeit des Genderitems 21 (AK) durch Männer dadurch interpretiert werden kann, dass Männer AK-Aufgaben generell etwas häufiger lösen als Frauen. Der standardisierte Effekt weist mit $d = 0.40$ auf einen kleinen Unterschied zwischen den Geschlechtern hin. Es konnte zudem ein positiver Zusammenhang der Lösungswahrscheinlichkeit aller drei Genderitems mit der SDV-Testleistung dahingehend beobachtet werden, dass Studierende, die eine gute SDV-Testleistung zeigten auch die Genderitems häufiger lösen konnten.

Der SDV wurde randomisiert in zwei Versionen, A und B, vorgegeben. Es konnten keine Reihenfolgeeffekte aufgedeckt werden, sodass sich hieraus keine Auswirkungen auf die Lösungswahrscheinlichkeit ergaben. Die beiden Versionen unterschieden sich durch die verschiedene Anordnung der Genderitems im Test. In Version A waren die Genderitems im ersten Teil des Tests verteilt (Item 1, 11 und 21) und in Version B in der zweiten Hälfte des SDV (Item 15, 25 und 35). Die unterschiedliche Positionierung der Genderitems im SDV hatte keine Auswirkung auf deren Lösungswahrscheinlichkeit.

Das Fähigkeitsselbstkonzept betreffend schätzten sich die Männer zu beiden Erhebungszeitpunkten höher ein als ihre weiblichen Kommilitoninnen. Diese Beobachtung korrespondiert mit Befunden aus der Literatur (Schilling, et al., 2006; Stipek & Gralinski, 1991; Tiedemann & Faber, 1995). Der Zusammenhang der SDV-Testleistung mit dem Fähigkeitsselbstkonzept fiel sowohl vor ($r = -.305$) als auch nach der Testung ($r = -.365$) mäßig hoch aus, womit eine mittlere Fähigkeitseinschätzung anzunehmen ist. Weiters schätzten sich beide Geschlechter zum zweiten Erhebungszeitpunkt, nach der Testung, selbstkritischer ein als davor.

10. Zusammenfassung

Vorliegende Untersuchung behandelt unter Berücksichtigung des Geschlechterstereotyps den Übergang vom konkreten zum formalen Stadium im Sinne der Stadientheorie von Piaget (1972b) und darüber hinaus den Einfluss einer latenten Genderinformation in spezifischen Items (sog. Genderitems) auf die Lösungswahrscheinlichkeit, welche dem SDV hinzugefügt wurden. Diese Arbeit stellt eine Erweiterung der Studien von Klopff (2011) und Eder (2012) dar, wobei die Untersuchung von Spiel, et al. (2001, 2004) die Basis bildet.

Die Erhebung wurde an fortgeschrittenen Studierenden im zweiten Abschnitt beziehungsweise im Masterstudium der Studienrichtungen Physik, Chemie, Mathematik und Psychologie durchgeführt. Als Erhebungsinstrument wurde der Test zum Schlussfolgernden Denken Verbal (SDV), der ursprünglich aus 32 konditionalen Syllogismen besteht und für diese Untersuchung um drei Genderitems erweitert wurde, in den Versionen A und B vorgelegt. In Version A waren die Genderitems in der ersten Hälfte des Tests positioniert, in Version B in der zweiten Hälfte. Durch Anwendung des Mixed Rasch-Modells konnten fünf Lösungsklassen identifiziert werden. Lösungsklasse 1 wurde aufgrund des bikonditionalen Antwortmusters dem konkret-operationalen Stadium zugeordnet. Zudem konnten jeweils ein Übergangsstadium 1 und 2 ermittelt werden, deren Lösungswahrscheinlichkeiten deutlich inhaltsabhängig waren. Studierende im Übergangsstadium 3 zeigten bereits ein homogeneres Antwortmuster und wiesen, bis auf die Argumentform MT, durchwegs hohe Lösungswahrscheinlichkeiten auf. Das Übergangsstadium 4 zeigte über alle Argument- und Inhaltsformen hinweg die höchsten Lösungswahrscheinlichkeiten, jedoch war auch hier die Argumentform MT die am wenigsten häufig gelöste. Wie bereits bei Klopff (2011) und Eder (2012) konnte auch in vorliegender Untersuchung keine Lösungsklasse, die das formale Stadium repräsentiert, identifiziert werden.

Im Allgemeinen konnte kein Unterschied der Lösungswahrscheinlichkeit bei Syllogismen zwischen Männern und Frauen festgestellt werden. Studierende der klassischen naturwissenschaftlichen Studienrichtungen (Mathematik, Physik und Chemie) konnten Syllogismen häufiger lösen als Studierende der Psychologie, wobei die entsprechende Effektgröße auf einen mäßigen Unterschied hinweist. Die Studienleistung selbst, operationalisiert anhand von Noten, weist einen schwachen Zusammenhang mit der SDV-Testleistung auf.

Die dem SDV hinzugefügten Genderitems der Argumentformen MP, NA und AK, jeweils kontrafaktisch und unnegiert, erwiesen sich im Vergleich zu den anderen Items als trennscharf

und mittelschwierig. Während bei den Formen MP und NA keine Geschlechtsunterschiede in den Lösungswahrscheinlichkeiten beobachtet werden konnten, wurde das Genderitem der Form AK von Männern häufiger gelöst als von Frauen. Die Genderinformation in den Genderitems dürfte somit keinen besonderen Einfluss auf die Lösungswahrscheinlichkeit haben. Die häufigere Lösbarkeit der Argumentform AK durch Männer kann dadurch erklärt werden, dass Männer AK-Aufgaben generell häufiger lösten als Frauen.

Hinsichtlich des Fähigkeitsselbstkonzeptes schätzen sich Männer zu beiden Erhebungszeitpunkten höher ein als Frauen. Gleichzeitig fiel die Einschätzung zum Fähigkeitsselbstkonzept nach der Testung für beide Geschlechter niedriger aus.

Zusammenfassend kann belegt werden, wie dies bereits in den Untersuchungen von Spiel, et al. (2001, 2004), Klopff (2011) und Eder (2012) gezeigt wurde, dass die kognitive Entwicklung zwischen dem konkret-operationalen und dem formalen Stadium in Übergangsphasen verläuft. Ein formales Stadium konnte jedoch auch bei fortgeschrittenen Studierenden nicht identifiziert werden, sofern die erreichte Klassifikation durch das MRM als Grundlage dieser Interpretation dienen kann. Die Annahme, dass aufgrund von vorhandenen Geschlechterstereotypen die im Genderitem beinhaltete latente Information zu einer niedrigeren Lösungswahrscheinlichkeit der Genderitems bei männlichen Studierenden führt, konnte im Rahmen dieser Arbeit nicht bestätigt werden. Die Genderinformation hat offensichtlich keinen Einfluss auf die Lösungswahrscheinlichkeit der Genderitems.

IV. LITERATURVERZEICHNIS

- Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, 93(5), 751-763.
- Aken, M. A. G. v., Helmke, A., & Schneider, W. (1997). Selbstkonzept und Leistung - Dynamik ihres Zusammenspiels: Ergebnisse aus dem SCHOLASTIK-Projekt. In F. Weinert & A. Helmke (Eds.), *Entwicklung im Grundschulalter*. Weinheim: Beltz.
- Bara, B. G., Bucciarelli, M., & Lombardo, V. (2001). Model theory of deduction: A unified computational approach. *Cognitive Science*, 25(6), 839-901.
- Beckmann, J. F., & Guthke, J. (1999). Psychodiagnostik des schlußfolgernden Denkens: Handbuch zur Adaptiven Computergestützten Intelligenz-Lerntestbatterie für Schlußfolgerndes Denken (ACIL). Göttingen: Hogrefe.
- Beller, S., & Kuhn Münch, G. (2007). What causal conditional reasoning tells us about people's understanding of causality. *Thinking & Reasoning*, 13(4), 426 - 460.
- Beller, S., & Spada, H. (1999). Inhaltseffekte beim propositionalen Schließen: Wie interagiert konzeptuelles Wissen mit syntaktischer Struktur? (Forschungsbericht Nr. 138). Freiburg: Psychologisches Institut der Albert-Ludwigs-Universität Freiburg i. Br.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. ed.). Heidelberg: Springer.
- Buggle, F. (1997). *Die Entwicklungspsychologie Jean Piagets*. Stuttgart [u.a.]: Kohlhammer.
- Bundesministerium für Wissenschaft Forschung und Wirtschaft (2011). *uni:data - Zahlen und Fakten zum Hochschulbereich*. Wien: BM:WFW. Abgerufen 11.05.2013 von <http://wissenschaft.bmwf.at/bmwf/wissenschaft-hochschulen/universitaeten/statistiken/>.
- Byrnes, J. P., & Overton, W. F. (1986). Reasoning about certainty and uncertainty in concrete, causal, and propositional contexts. *Developmental Psychology*, 22(6), 793 - 799.
- Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? *Psychological Science*, 16(7), 572-578.
- Calsyn, R., & Kenny, D. (1977). Self-concept of ability and perceived evaluations by others: Cause or effect of academic achievement. *Journal of Educational Psychology*, 69, 136-145.
- Caplan, P. J., Crawford, M., Hyde, J. S., & Richardson, J. T. E. (1997). *Gender differences in human cognition*. New York, Oxford: Oxford University Press.

- Danaher, K., & Crandall, C. S. (2008). Stereotype threat in applied settings re-examined. *Journal of Applied Social Psychology*, 38(6), 1639.
- Dickhäuser, O. (2006). Fähigkeitsselbstkonzepte Entstehung, Auswirkung, Förderung. *Zeitschrift für Pädagogische Psychologie*, 20(1/2), 5-8.
- Dickhäuser, O., & Galfe, E. (2004). Besser als ..., schlechter als ... - Leistungsbezogene Vergleichsprozesse in der Grundschule. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 36(1), 1-9.
- Dresel, M., Heller, K. A., Schober, B., & Ziegler, A. (2001). Geschlechtsunterschiede im mathematisch-naturwissenschaftlichen Bereich: Motivations- und selbstwertschädliche Einflüsse der Eltern auf Ursachenerklärungen ihrer Kinder in Leistungskontexten. In C. Finkbeiner & G. W. Schnaitmann (Eds.), *Lehren und Lernen im Kontext empirischer Forschung und Fachdidaktik* (pp. 270-288). Donauwörth: Auer.
- Dresel, M., & Schober, B. (2004). *Golem und Pygmalion: Scheitert die Chancengleichheit von Mädchen in den Köpfen von Müttern und Vätern?* Paper presented at the Vortrag auf dem Kongress der Deutschen Gesellschaft für Erziehungswissenschaften, 21. - 24. März 2004, Zürich.
- Dresel, M., Schober, B., & Ziegler, A. (2007). Golem und Pygmalion. Scheitert die Chancengleichheit von Mädchen im mathematisch-naturwissenschaftlich-technischen Bereich am geschlechtsstereotypen Denken der Eltern. In P. H. Ludwig & H. Ludwig (Eds.), *Erwartungen in himmelblau und rosarot. Effekte, Determinanten und Konsequenzen von Geschlechterdifferenzen in der Schule* (pp. 61-81). Weinheim/München: Juventa.
- Eagly, A. H. (1987). *Sex differences in social behavior*. NJ: Erlbaum.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573-598.
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Trautner (Eds.), *The developmental social psychology of gender*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Eckes, T. (2002). Paternalistic and envious gender stereotypes: testing predictions from the stereotype content model. *Sex Roles*, 47(3/4), 99-114.
- Eckes, T. (2004). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen. In R. Becker & B. Kortendiek (Eds.), *Handbuch Frauen- und Geschlechterforschung: Theorie, Methoden, Empirie*. Wiesbaden: VS Verlag für Sozialwissenschaften.

- Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen. In R. Becker, & Kortendiek, B. (Ed.), *Handbuch Frauen- und Geschlechterforschung* (pp. 171-182). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Eder, U. (2012). *Identifikation von Entwicklungsstadien im deduktiven Schlussfolgern auf Basis der Stadientheorie von Piaget bei StudienanfängerInnen naturwissenschaftlicher Fächer an der Universität Wien*. Unveröffentlichte Diplomarbeit, Universität Wien.
- Evans, J. S. B. T., Clibbens, J., & Rood, B. (1995). Bias in conditional inference: Implications for mental models and mental logic. *The Quarterly Journal of Experimental Psychology*, 48A(3), 644 - 670.
- Evans, J. S. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human Reasoning: The psychology of deduction*: Hove: Lawrence Erlbaum.
- Eysenck, M. W., & Keane, M. T. (2000). *Cognitive Psychology: A student's handbook* (4th ed.). Hove, UK: Psychology Press.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902.
- Flammer, A. (2009). *Entwicklungstheorien: Psychologische Theorien der menschlichen Entwicklung*. Bern: Huber.
- Gärtner, K., Schwabe, M., & Sommer-Binder, G. (2010). Frauenbericht 2010, Teil 1: Statistische Analyse zur Entwicklung der Situation von Frauen in Österreich, Kapitel 2: Bildung. Abgerufen 11.07.2013, von <http://www.bka.gv.at/site/7207/default.aspx>
- Gauffroy, C., & Barrouillet, P. (2011). The primacy of thinking about possibilities in the development of reasoning. *Developmental Psychology*, 47(4), 1000-1011.
- Gold, A., & Souvignier, E. (2005). Prognose der Studierfähigkeit - Ergebnisse aus Längsschnittanalysen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 37, 214-222.
- Gößler, H. (2001). *Untersuchung zur Entwicklung schlussfolgernden Denkens bei konditionalen Syllogismen aus einer Piagetschen Entwicklungsperspektive*. Unveröffentlichte Diplomarbeit, Universität Graz.
- Hattie, A. (1992). *Self-concept*. Hillsdale, NJ: Erlbaum.
- Helmke, A. (1992). *Selbstvertrauen und schulische Leistungen*. Göttingen: Hogrefe.
- Hyde, J. S. (1990). Meta-analysis and the psychology of gender differences. *Signs*, 16, 55-73.
- Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, 107(2), 139-155.

- Ihme, T. A., & Mauch, M. (2007). Werbung als implizite Aktivierungsquelle von Geschlechterstereotypen und ihr Einfluss auf Mathematikleistungen sowie auf das Computerwissen bei Mädchen und Jungen. *Empirische Pädagogik*, 21(3), 291-305.
- Janveau-Brennan, G., & Markovits, H. (1999). The development of reasoning with causal conditionals. *Developmental Psychology*, 35(4), 904 - 911.
- Johnson-Laird, P. N. (1983). *Mental Models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, Massachusetts: Harvard University Press.
- Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual Review of Psychology*, 50, 109-135.
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. London: Psychology Press.
- Kammermeyer, G., & Martschinke, S. (2006). Selbstkonzept- und Leistungsentwicklung in der Grundschule - Ergebnisse der KILIA-Studie. *Empirische Pädagogik*, 20, 245-259.
- Kanka, M., Wagner, P., Schober, B., & Spiel, C. (2011). Gender-stereotyped attitudes and behavior in kindergarten students. *The International Journal of Learning*, 18(2), 291-304.
- Kanka, M., Wagner, P., Schober, B., & Spiel, C. (2013). Gender-stereotyped attitudes in kindergarten students: A multicausal analysis. *The European Journal of Social & Behavioural Sciences*, 8(5), 1294-1299.
- Keller, J. (2008). Stereotype als Bedrohung. In L. P. Ericson & B. Six (Eds.), *Stereotype, Vorurteile und soziale Diskriminierung. Theorien, Befunde und Interventionen*. Weinheim: Beltz.
- Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 138(3), 427 - 457.
- Kiper, H., & Kattmann, U. (2003). Basiskompetenz im Vergleich - Übersicht über Ergebnisse der PISA-Studie 2000. In B. Moschner, H. Kiper & U. Kattmann (Eds.), *PISA 2000 als Herausforderung. Perspektiven für Lehren und Lernen*. Baltmannsweiler: Schneider Verlag.
- Klopf, M. (2011). *Die Kompetenz im schlussfolgernden Denken bei Studentinnen und Studenten zu Studienbeginn. Ein Vergleich zwischen Psychologie- und Naturwissenschaftsstudierenden*. Unveröffentlichte Diplomarbeit, Universität Wien.
- Knight, G. P., Fabes, G. A., & Higgins, D. A. (1996). Concerns about drawing causal inferences from meta-analyses: an example in the study of gender differences in aggression. *Psychological Bulletin*, 119(3), 410-421.
- Kodroff, J. K., & Roberge, J. J. (1975). Development analysis of the conditional reasoning abilities of primary-grade children. *Developmental Psychology*, 11(1), 21 - 28.

- Kubinger, K. D. (2006). *Psychologische Diagnostik. Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Kuhn, D. (1977). Conditional reasoning in children. *Developmental Psychology*, 13(4), 342 - 353.
- Leighton, J. P. (2004). Defining and describing reason. In R. J. S. J. P. Leighton (Ed.), *The nature of reasoning*. Cambridge: Cambridge University Press.
- Lohrmann, K., Götz, T., & Haag, L. (2010). Zusammenhänge von fachspezifischen Leistungen und Fähigkeitsselbstkonzepten im Grundschulalter. In B. Schwarz (Ed.), *Erziehungswissenschaftliche Forschung - nachhaltige Bildung: Beiträge zur 5. DGfE-Sektionstagung "Empirische Bildungsforschung" AEPF-KBBB im Frühjahr 2009* (pp. 296-303). Landau: Verlag Empirische Pädagogik e.V., 2010.
- Ludwig, P. H. (1991). *Sich selbst erfüllende Prophezeiungen im Alltagsleben. Theorie und empirische Basis von Erwartungseffekten und Konsequenzen für die Pädagogik*. Stuttgart: Hogrefe.
- Ludwig, P. H. (1999). *Imagination. Sich selbst erfüllende Vorstellungen zur Förderung von Lernprozessen*. Opladen: Leske + Budrich.
- Ludwig, P. H., Dresel, M., Finsterwald, M., Fischer, N., Holder, M. C., Rustemeyer, R., . . . Ziegler, A. (2006). Erwartungen in himmelblau und rosarot: Erklärung für Geschlechterdifferenzen im lebenslangen Lernen. In R. Fatke & H. Merckens (Eds.), *Bildung über die Lebenszeit*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Markovits, H., Fleury, M.-L., Quinn, S., & Venet, M. (1998). The development of conditional reasoning and the structure of semantic memory. *Child Development*, 69(3), 742 - 755.
- Markovits, H., & Vachon, R. (1990). Conditional reasoning, representation, and level of abstraction. *Developmental Psychology*, 26, 942 - 951.
- Markovits, H., Venet, M., Janveau-Brennan, G., Malfait, N., Pion, N., & Vadeboncoeur, I. (1996). Reasoning in young children: Fantasy and information retrieval. *Child Development*, 67, 2857 - 2872.
- Marsh, H. W. (1986). Verbal and math self-concepts: an internal/external frame of reference model. *American Educational Research Journal*, 23, 129 - 149.
- Marsh, H. W., Byrne, B. M., & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, 80, 366 - 380.
- Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and

- multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin*, 34(4), 542-552.
- Metz-Göckel, S. (1988). Geschlechterverhältnisse, Geschlechtersozialisation und Geschlechtsidentität. *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie*, 8(2), 85-97.
- Möller, J., & Köller, O. (2004). Die Genese akademischer Selbstkonzepte: Effekte dimensionaler und sozialer Vergleiche. *Psychologische Rundschau*, 55(1), 19-27.
- Moshman, D. (1977). Consolidation and stage formation in the emergence of formal operations. *Developmental Psychology*, 13(2), 95-100.
- Müller, U., Overton, W. F., & Reese, K. (2001). Development of conditional reasoning: A longitudinal study. *Journal of Cognition and Development*, 2(1), 27 - 49.
- Newstead, S. E., Handley, S. J., Harley, C., Wright, H., & Farrelly, D. (2004). Individual differences in deductive reasoning. *The Quarterly Journal of Experimental Psychology*, 57A(1), 33 - 60.
- Niklas, F., & Schneider, W. (2012). Die Anfänge geschlechts-spezifischer Leistungsunterschiede in mathematischen und schriftsprachlichen Kompetenzen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(3), 123-138.
- OECD. (2014). *PISA 2012 Ergebnisse: Was Schülerinnen und Schüler wissen und können (Band I, Überarbeitete Ausgabe, Februar 2014): Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften*. W. Bertelsmann Verlag: Germany.
- Örter, R., & Montada, L. E. (2008). *Entwicklungspsychologie*. Weinheim: Beltz.
- Overton, W. F., Ward, S. L., Black, J., Noveck, I. A., & O'Brien, D. P. (1987). Form and content in the development of deductive reasoning. *Developmental Psychology*, 23(1), 22 - 30.
- Piaget, J. (1972a). Intellectual evolution from adolescence to adulthood. *Human Development*, 15, 1-12.
- Piaget, J. (1972b). *Theorien und Methoden der modernen Erziehung*. Wien: Molden.
- Piaget, J. (1975a). *Das Erwachen der Intelligenz beim Kinde* Stuttgart: Ernst Klett.
- Piaget, J. (1975b). *Der Aufbau der Wirklichkeit beim Kinde* (Vol. Gesammelte Werke, Bd. 2, Studienausgabe). Stuttgart: Ernst Klett.
- Piaget, J. (1976a). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Piaget, J. (1976b). *Psychologie der Intelligenz*. München: Kindler.
- Piaget, J. (1984). Die intellektuelle Entwicklung im Jugend- und im Erwachsenenalter. In D. G. T. Schöfthaler (Ed.), *Soziale Struktur und Vernunft. Jean Piagets Modell entwickelten*

- Denkens in der Diskussion kulturvergleichender Forschung* (pp. 47-60). Frankfurt am Main: Suhrkamp.
- Piaget, J. (2003). *Meine Theorie der geistigen Entwicklung*. New York: John Wiley & Sons, Inc.
- Piaget, J., & Inhelder, B. (1977). *Von der Logik des Kindes zur Logik des Heranwachsenden*. Olten: Walger-Verlag.
- Reverberi, C., Cherubini, P., Frackowiak, R. S. J., Caltagirone, C., Paulesu, E., & Macaluso, E. (2010). Conditional and syllogistic deductive tasks dissociate functionally during premise integration. *Human Brain Mapping, 31*, 1430-1445.
- Rindermann, H., & Oubaid, V. (1999). Auswahl von Studienanfängern durch Universitäten - Kriterien, Verfahren und Prognostizierbarkeit des Studienerfolgs. *Zeitschrift für Differentielle und Diagnostische Psychologie, 20*, 172-191.
- Roberge, J. J., & Mason, E. J. (1978). Effects of negation on adolescents' class and conditional reasoning abilities. *The Journal of General Psychology, 98*(2), 187-195.
- Roberge, J. J., & Paulus, D. H. (1971). Developmental patterns for children's class and conditional reasoning abilities. *Developmental Psychology, 4*(2), 191 - 200.
- Rost, D. H. (2010). *Handwörterbuch Pädagogische Psychologie*. Weinheim, Babel: Beltz.
- Rost, J. (1990). Rasch models in latent classes: An integration of two approaches in item analysis. *Applied Psychological Measurement, 14*, 271-282.
- Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion*. Bern: Verlag Hans Huber.
- Rost, J., & von Davier, M. (1995). Mixture distribution Rasch models. In G.H. Fischer & I. W. Molenaar (Eds.), *Rasch models: Foundations, recent developments, and applications* (pp. 257-268). New York: Springer.
- Rustemeyer, R. (1988). Geschlechtsstereotype und ihre Auswirkungen auf das Sozial- und Leistungsverhalten. *Zeitschrift für Sozialisationsforschung und Erziehungssoziologie, 8*(2), 115-129.
- Scharlau, I. (1996). *Jean Piaget zur Einführung*. Hamburg: Junius.
- Schilling, S. R., Sparfeldt, J. R., & Rost, D. H. (2006). Facetten schulischen Selbstkonzepts: Welchen Unterschied macht das Geschlecht? *Zeitschrift für Pädagogische Psychologie, 20*(1/2), 9-18.
- Schmader, T., Johns, M., & Forbes, C. E. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review, 115*(2), 330-356.
- Schmidt-Atzert, L. (2005). Prädiktion von Studienerfolg bei Psychologiestudenten. *Psychologische Rundschau, 56*(2), 131-133.

- Schmidt-Atzert, L., & Amelang, M. (2012). *Psychologische Diagnostik*. Berlin Heidelberg: Springer-Verlag.
- Schröder, E. (1989). *Vom konkreten zum formalen Denken. Individuelle Entwicklungsverläufe von der Kindheit bis zum Jugendalter*. Bern: Hans Huber.
- Shavelson, R. J., & Bolus, R. (1981). Self-concept: The interplay of theory and methods. *Journal of Educational Psychology*, 74(1), 3-17.
- Sodian, B. (2007). Entwicklung des Denkens. In M. Hasselhorn & W. Schneider (Eds.), *Handbuch der Entwicklungspsychologie* (pp. 244-254). Göttingen: Hogrefe.
- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4-28.
- Spiel, C., Gittler, G., Sirsch, U., & Glück, J. (1997). Application of the Rasch model for testing Piaget's theory of cognitive development. In R. Langeheine, & Rost, J. (Ed.), *Applications of latent trait and latent class models in the social sciences* (pp. 111-117). Münster: Waxmann.
- Spiel, C., & Glück, J. (2008). A modal-based test of competence profile and competence level in deductive reasoning. In J. Hartig, E. Klieme, & D. Leutner (Ed.), *Assessment of competencies in educational contexts* (pp. 41 - 60). Göttingen: Hogrefe & Huber.
- Spiel, C., Glück, J., & Gössler, H. (2001). Stability and change of unidimensionality: The sample case of deductive reasoning. *Journal of Adolescent Research*, 16(2), 150 - 168.
- Spiel, C., Glück, J., & Gössler, H. (2003). SDV. Leistungsprofiltest Schlussfolgerndes Denken - Verbal. *Leistungsprofiltest Schlussfolgerndes Denken - Verbal*. Wien: Universität Wien, Arbeitsbereich Bildungspsychologie und Evaluation.
- Spiel, C., Glück, J., & Gössler, H. (2004). Messung von Leistungsprofil und Leistungshöhe im schlussfolgernden Denken im SDV - Die Interpretation von Piagets Entwicklungskonzept und Item-Response Modellen. *Diagnostica* 50(3), 145-152.
- Staudenmeyer, H., & Bourne, L. E. (1977). Learning to interpret conditional sentences: A developmental study. *Developmental Psychology*, 13(6), 616 - 623.
- Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *The American Psychologist*, 52(6), 613-629.
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, 34, 379-440.

- Steins, G. (2010). Einführung in "Psychologie und Geschlechterforschung". In G. Steins (Ed.), *Handbuch Psychologie und Geschlechterforschung* (pp. 11-23). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Steyer, R., Yousfi, S., & Würfel, K. (2005). Prädiktion von Studienerfolg: Der Zusammenhang zwischen Schul- und Studiennoten im Diplomstudiengang Psychologie. *Psychologische Rundschau*, 56, 129-131.
- Stipek, D. J., & Gralinski, J. H. (1991). Gender differences in children's achievement-related beliefs and emotional responses to success and failure in mathematics. *Journal of Educational Psychology*, 83(3), 361-371.
- Stöger, H., Ziegler, A., & David, H. (2004). What is a specialist? Effects of the male concept of a successful academic person on the performance in a thinking task. *Psychology Science*, 46(4), 514-530.
- Technische Universität Wien (2011). *Öffentliche Statistiken zur Lehre der TU Wien*. Wien: TU Wien. Abgerufen 09.05.2012 von https://tiss.tuwien.ac.at/statistik/public_lehre.
- Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, 22(6), 742 - 758.
- Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of childrens' concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92(1), 144-151.
- Tiedemann, J., & Faber, G. (1995). Mädchen im Mathematikunterricht: Selbstkonzept und Kausalattributionen im Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27, 61-71.
- Waldmann, M. R., & Weinert, F. E. (1990). *Intelligenz und Denken. Perspektiven der Hochbegabungsforschung*. Göttingen: Hogrefe.
- Wildman, T. M., & Fletcher, H. J. (1977). Developmental increases and decreases in solutions of conditional syllogism problems. *Developmental Psychology*, 13(6), 630 - 636.

V. ANHANG

Anhang 1 Tabelle 23	91
Anhang 2 SDV Testversionen A und B	92
Anhang 3 Kurzfassung	117
Anhang 3 Abstract	118
Anhang 4 Curriculum Vitae	120

Anhang 1Tabelle 23: *SDV Lösungswahrscheinlichkeiten, korrigierte Itemtrennschärfen und Cronbach's Alpha bei Auslassung des entsprechenden Items ($k = 35$, $n = 391$)*

Itemnummer (Struktur)	Lösungs- wahrscheinlichkeit (p)	Korrigierte Trennschärfe (r_{it})	Cronbachs α , wenn Item weggelassen
1. 311	.793	.436	.892
2. 121	.583	.325	.894
3. 131	.724	.401	.893
4. 232	.527	.650	.888
5. 211	.962	.018	.896
6. 111	.862	.391	.893
7. 342	.471	.406	.892
8. 242	.476	.472	.891
9. 341	.540	.644	.888
10. 311	.862	.392	.893
11. 331	.775	.511	.891
12. 231	.645	.576	.889
13. 212	.864	.401	.893
14. 322	.414	.273	.895
15. 221	.619	.166	.897
16. 141	.662	.594	.889
17. 122	.468	.334	.894
18. 241	.591	.445	.892
19. 142	.552	.597	.889
20. 312	.910	.266	.894
21. 341	.673	.550	.890
22. 222	.506	.166	.897
23. 321	.619	.238	.895
24. 112	.813	.484	.891
25. 331	.624	.464	.891
26. 132	.777	.439	.892
27. 332	.637	.436	.892
28. 131	.721	.437	.892
29. 242	.440	.603	.889
30. 322	.407	.334	.894
31. 221	.573	.286	.895
32. 141	.616	.586	.889
33. 122	.422	.366	.893
34. 241	.532	.465	.891
35. 142	.642	.412	.892

Anmerkung. Legende der Itemstruktur: 1. Stelle: Semantischer Aufgabeninhalt (1 = konkret, 2 = abstrakt, 3 = kontrafaktisch); 2. Stelle: Syllogismen (1 = MP, 2 = MT, 3 = NA, 4 = AK); 3. Stelle: Negation (1 = unnegiert, 2 = negiert). Die drei Genderitems (1, 11 und 21) sind hervorgehoben.

Anhang 2

Nachstehend ist der in vorliegender Untersuchung verwendete SDV Test in den Version A und B hinzugefügt. Bei Version A finden sich die in den verschiedenen Studienrichtungen verwendete erste Seite des Fragebogens wieder.

A/Psy = Testversion A / Psychologie

A/P = Testversion A / Physik

A/P_TU = Testversion A / Physik_Technische Universität Wien

A/M_TU = Testversion A / Mathematik_Technische Universität Wien

A/C = Testversion A / Chemie

B/Psy = Testversion B / Psychologie

SDV Testversion A

FRAGEBOGEN

Angaben zur Person

Geschlecht: ☐ weiblich
☐ männlich

Alter: _____

Muttersprache: _____

Höchste abgeschlossene Schulbildung der Mutter:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Höchste abgeschlossene Schulbildung des Vaters:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Angaben zu Ihren bisherigen Noten im Psychologiestudium

Noten im 1. Diplomprüfungszeugnis (Gesamtnoten)

1

Allgemeine Psychologie	
Psychologische Methodenlehre	
Biologische Psychologie	

FRAGEBOGEN

Angaben zur Person

Geschlecht: ☐ weiblich
☐ männlich

Alter: _____

Muttersprache: _____

Höchste abgeschlossene Schulbildung der Mutter:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Höchste abgeschlossene Schulbildung des Vaters:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Angaben zu Ihren bisherigen Noten im Physikstudium

Noten im 1. Diplomprüfungszeugnis (Gesamtnoten)	
Einführung in die Physik I	
Einführung in die Physik II	
Analysis für die Physik I	
Analysis für die Physik II	

Wie gut schätzen Sie Ihre Studienleistung generell ein?

☐ gut ☐ eher gut ☐ eher schlecht ☐ schlecht

FRAGEBOGEN

Angaben zur Person

Geschlecht: ☐ weiblich
☐ männlich

Alter: _____

Muttersprache: _____

Höchste abgeschlossene Schulbildung der Mutter:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Höchste abgeschlossene Schulbildung des Vaters:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Angaben zu Ihren bisherigen Noten im Physikstudium

(Bitte entweder für Diplomstudium ODER Bachelorstudium angeben!)

Noten im 1. Diplomprüfungszeugnis (Gesamtnoten)	
Grundlagen der Physik A	
Elektronik	
Mathematik A	

Noten im Bachelorzeugnis (Modulnoten)	
Modul Grundlagen der Physik	
Modul Theoretische Physik	
Modul Mathematik	

Wie gut schätzen Sie Ihre Studienleistung generell ein?

☐ gut ☐ eher gut ☐ eher schlecht ☐ schlecht

FRAGEBOGEN

Angaben zur Person

Geschlecht: ☐ weiblich
☐ männlich

Alter: _____

Muttersprache: _____

Höchste abgeschlossene Schulbildung der Mutter:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Höchste abgeschlossene Schulbildung des Vaters:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Angaben zu Ihren bisherigen Noten im Mathematikstudium

Noten im 1. Diplomprüfungszeugnis (Gesamtnoten)	
Lineare Algebra	
Analysis	
Wahrscheinlichkeitstheorie und Statistik	

Wie gut schätzen Sie Ihre Studienleistung generell ein?

☐ gut ☐ eher gut ☐ eher schlecht ☐ schlecht

FRAGEBOGEN

Angaben zur Person

Geschlecht: ☐ weiblich ☐ männlich	Alter: _____
Muttersprache: _____	

Höchste abgeschlossene Schulbildung der Mutter:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Höchste abgeschlossene Schulbildung des Vaters:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Angaben zu Ihren bisherigen Noten im Chemiestudium

(Bitte entweder für Diplomstudium **ODER** Bachelorstudium angeben!)

Noten im 1. Diplomprüfungszeugnis (Gesamtnoten)	Noten im Bachelorzeugnis (Modulnoten)
Allgemeine Chemie	Modul Organische Chemie I
Organische Chemie	Modul Organische Chemie II
Physikalische Chemie	Modul Physikalische Chemie I
	Modul Physikalische Chemie II

Universität: ☐ Universität Wien
☐ Technische Universität

Wie gut schätzen Sie Ihre Studienleistung generell ein?

☐ gut ☐ eher gut ☐ eher schlecht ☐ schlecht

Wie gut schätzen Sie Ihre Studienleistung generell ein?

☐ gut ☐ eher gut ☐ eher schlecht ☐ schlecht

Instruktion

Im Folgenden bearbeiten Sie bitte einige Aufgaben zum logischen Denken.

Diese Aufgaben bestehen immer aus zwei Aussagen. Bei der Beantwortung der Aufgaben sollen Sie davon ausgehen, dass diese Aussagen wahr sind, auch wenn sie Ihrer Erfahrung nicht entsprechen.

Lassen Sie sich nicht davon irritieren, dass bei einigen Aufgaben die ersten Aussagen gleich sind. Da sich diese Aufgaben in ihren zweiten Aussagen unterscheiden, gleich keine Aufgabe einer anderen.

Den beiden Aussagen folgen immer drei Antwortmöglichkeiten, **von denen immer nur eine richtig ist**.

Kreuzen Sie bei jeder Aufgabe die Lösung an, von der Sie glauben, dass sie die richtige ist.

Die Möglichkeit „Ich kann diese Aufgabe nicht lösen“ soll nur gewählt werden, wenn Sie mit einer Aufgabe überhaupt nichts anfangen können.

Eine Aufgabe könnte zum Beispiel so aussehen:

Alle blonden Buben sind groß,
Josef ist ein großer Bub.

- ☐ Josef ist blond.
- ☐ Josef ist nicht blond.
- ☐ Vielleicht ist Josef blond, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

In dieser Aufgabe ist die dritte Antwort richtig, da man aufgrund der zwei Aussagen nicht wissen kann, ob Josef blond ist oder nicht.

In den Aufgaben, die Sie bearbeiten sollen, kommen Aussagen vor wie zum Beispiel:

Wenn Josef krank ist, dann liegt er im Bett.

Bedenken Sie, dass diese Aussage

NICHT bedeutet, dass Josef nur dann im Bett liegt, wenn er krank ist. Diese Aussage bedeutet **NICHT**: *Wenn Josef im Bett liegt, dann ist er krank.*

Es kann auch andere Gründe geben, dass Josef im Bett liegt (z.B. wenn er schläft).

Lesen Sie jede Aufgabe sorgfältig durch und **denken Sie über jede Aufgabe neu nach**.

Achten Sie auch darauf, ob in den Aussagen das Wort „nicht“ vorkommt.

Halten Sie sich nicht zu lange mit einer Aufgabe auf, da Sie **alle 35 Aufgaben** beantworten sollen.

Bedenken Sie, dass Sie die Aufgaben **nur aufgrund der jeweils vorgegebenen Aussagen** beantworten sollen, unabhängig von Ihren Erfahrungen.

Wie gut glauben Sie, solche Aufgaben lösen zu können?

☐ gut

☐ eher gut

☐ eher schlecht

☐ schlecht

VIEL SPASS UND GUTES GELINGEN!

1. Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.

Maria hat gute Laune.

- ☐ Maria schreibt in Physik eine gute Note.
- ☐ Maria schreibt in Physik keine gute Note.
- ☐ Vielleicht schreibt Maria in Physik eine gute Note, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

2. Wenn die Sonne scheint, trägt Tina einen roten Rock.

Tina trägt keinen roten Rock.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

3. Wenn die Sonne scheint, trägt Tina einen roten Rock.

Die Sonne scheint nicht.

- ☐ Tina trägt einen roten Rock.
- ☐ Tina trägt keinen roten Rock.
- ☐ Vielleicht trägt Tina einen roten Rock, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

4. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.

X gehört zur Gruppe B.

- ☐ X hat die Eigenschaft c.
- ☐ X hat nicht die Eigenschaft c.
- ☐ Vielleicht hat X die Eigenschaft c, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

5. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.

Y gehört zur Gruppe F.

- ☐ Y hat die Eigenschaft g.
- ☐ Y hat nicht die Eigenschaft g.
- ☐ Vielleicht hat Y die Eigenschaft g, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

6. Wenn die Sonne scheint, trägt Tina einen roten Rock.

Die Sonne scheint.

- ☐ Tina trägt einen roten Rock.
- ☐ Tina trägt keinen roten Rock.
- ☐ Vielleicht trägt Tina einen roten Rock, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

7. Wenn es **nicht** Abend ist, geht die Sonne unter.
 Die Sonne geht unter.
☐ Es ist Abend.
☐ Es ist nicht Abend.
☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
☐ Ich kann diese Aufgabe nicht lösen.
8. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.
 X hat die Eigenschaft c.
☐ X gehört zur Gruppe B.
☐ X gehört nicht zur Gruppe B.
☐ Vielleicht gehört X zur Gruppe B, vielleicht aber auch nicht.
☐ Ich kann diese Aufgabe nicht lösen.
9. Wenn es Abend ist, geht die Sonne auf.
 Die Sonne geht auf.
☐ Es ist Abend.
☐ Es ist nicht Abend.
☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
☐ Ich kann diese Aufgabe nicht lösen.
10. Wenn es Abend ist, geht die Sonne auf.
 Es ist Abend.
☐ Die Sonne geht auf.
☐ Die Sonne geht nicht auf.
☐ Vielleicht geht die Sonne auf, vielleicht aber auch nicht.
☐ Ich kann diese Aufgabe nicht lösen.
11. Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.
 Maria hat keine gute Laune.
☐ Maria schreibt in Physik eine gute Note.
☐ Maria schreibt in Physik keine gute Note.
☐ Vielleicht schreibt Maria in Physik eine gute Note, vielleicht aber auch nicht.
☐ Ich kann diese Aufgabe nicht lösen.
12. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.
 Y gehört nicht zur Gruppe F.
☐ Y hat die Eigenschaft g.
☐ Y hat nicht die Eigenschaft g.
☐ Vielleicht hat Y die Eigenschaft g, vielleicht aber auch nicht.
☐ Ich kann diese Aufgabe nicht lösen.

13. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.

X gehört nicht zur Gruppe B.

- ☐ X hat die Eigenschaft c.
- ☐ X hat nicht die Eigenschaft c.
- ☐ Vielleicht hat X die Eigenschaft c, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

14. Wenn es **nicht** Abend ist, geht die Sonne unter.

Die Sonne geht nicht unter.

- ☐ Es ist Abend.
- ☐ Es ist nicht Abend.
- ☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

15. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.

Y hat nicht die Eigenschaft g.

- ☐ Y gehört zur Gruppe F.
- ☐ Y gehört nicht zur Gruppe F.
- ☐ Vielleicht gehört Y zur Gruppe F, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

16. Wenn die Sonne scheint, trägt Tina einen roten Rock

Tina trägt einen roten Rock.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

17. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Peter trägt keine blaue Hose.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

18. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.

Y hat die Eigenschaft g.

- ☐ Y gehört zur Gruppe F.
- ☐ Y gehört nicht zur Gruppe F.
- ☐ Vielleicht gehört Y zur Gruppe F, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

19. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Peter trägt eine blaue Hose.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

20. Wenn es **nicht** Abend ist, geht die Sonne unter.

Es ist nicht Abend.

- ☐ Die Sonne geht unter.
- ☐ Die Sonne geht nicht unter.
- ☐ Vielleicht geht die Sonne unter, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

21. Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.

Maria schreibt in Physik eine gute Note.

- ☐ Maria hat gute Laune.
- ☐ Maria hat keine gute Laune.
- ☐ Vielleicht hat Maria gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

22. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.

X hat nicht die Eigenschaft c.

- ☐ X gehört zur Gruppe B.
- ☐ X gehört nicht zur Gruppe B.
- ☐ Vielleicht gehört X zur Gruppe B, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

23. Wenn es Abend ist, geht die Sonne auf.

Die Sonne geht nicht auf.

- ☐ Es ist Abend.
- ☐ Es ist nicht Abend.
- ☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

24. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Die Sonne scheint nicht.

- ☐ Peter trägt eine blaue Hose.
- ☐ Peter trägt keine blaue Hose.
- ☐ Vielleicht trägt Peter eine blaue Hose, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

25. Wenn es Abend ist, geht die Sonne auf.

Es ist nicht Abend.

- ☐ Die Sonne geht auf.
- ☐ Die Sonne geht nicht auf.
- ☐ Vielleicht geht die Sonne auf, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

26. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Die Sonne scheint.

- ☐ Peter trägt eine blaue Hose.
- ☐ Peter trägt keine blaue Hose.
- ☐ Vielleicht trägt Peter eine blaue Hose, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

27. Wenn es **nicht** Abend ist, geht die Sonne unter.

Es ist Abend.

- ☐ Die Sonne geht unter.
- ☐ Die Sonne geht nicht unter.
- ☐ Vielleicht geht die Sonne unter, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

28. Wenn Herr Müller gute Laune hat, fährt er nach Berlin.

Herr Müller hat keine gute Laune.

- ☐ Herr Müller fährt nach Berlin.
- ☐ Herr Müller fährt nicht nach Berlin.
- ☐ Vielleicht fährt Herr Müller nach Berlin, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

29. Wenn C **nicht** auf D trifft, dann entscheidet sich C für m.

C entscheidet sich für m.

- ☐ C trifft auf D.
- ☐ C trifft nicht auf D.
- ☐ Vielleicht trifft C auf D, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

30. Wenn es **nicht** regnet, ist die Straße nass.

Die Straße ist nicht nass.

- ☐ Es regnet.
- ☐ Es regnet nicht.
- ☐ Vielleicht regnet es, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

31. Wenn A auf B trifft, dann entscheidet sich A für c.

A entscheidet sich nicht für c.

- ☐ A trifft auf B.
- ☐ A trifft nicht auf B.
- ☐ Vielleicht trifft A auf B, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

32. Wenn Herr Müller gute Laune hat, fährt er nach Berlin.

Herr Müller fährt nach Berlin.

- ☐ Herr Müller hat gute Laune.
- ☐ Herr Müller hat keine gute Laune.
- ☐ Vielleicht hat Herr Müller gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

33. Wenn Frau Schuster **keine** gute Laune hat, geht sie spazieren.

Frau Schuster geht nicht spazieren.

- ☐ Frau Schuster hat gute Laune.
- ☐ Frau Schuster hat keine gute Laune.
- ☐ Vielleicht hat Frau Schuster gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

34. Wenn A auf B trifft, dann entscheidet sich A für c.

A entscheidet sich für c.

- ☐ A trifft auf B.
- ☐ A trifft nicht auf B.
- ☐ Vielleicht trifft A auf B, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

35. Wenn Frau Schuster **keine** gute Laune hat, geht sie spazieren.

Frau Schuster geht spazieren.

- ☐ Frau Schuster hat gute Laune.
- ☐ Frau Schuster hat keine gute Laune.
- ☐ Vielleicht hat Frau Schuster gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

Wie gut konnten Sie diese Aufgaben lösen?

- ☐ gut
 ☐ eher gut
 ☐ eher schlecht
 ☐ schlecht

Vielen Dank für Ihre Hilfe!

SDV Testversion B**FRAGEBOGEN****Angaben zur Person**

Geschlecht: ☐ weiblich
☐ männlich

Alter: _____

Muttersprache: _____

Höchste abgeschlossene Schulbildung der Mutter:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Höchste abgeschlossene Schulbildung des Vaters:

- ☐ Kein Schulabschluss
- ☐ Pflichtschule
- ☐ Lehre
- ☐ Fachschule
- ☐ Matura
- ☐ Universität / Fachhochschule

Angaben zu Ihren bisherigen Noten im Psychologiestudium

1

Noten im 1. Diplomprüfungszeugnis (Gesamtnoten)

Allgemeine Psychologie	
Psychologische Methodenlehre	
Biologische Psychologie	

Wie gut schätzen Sie Ihre Studienleistung generell ein?

☐ gut ☐ eher gut ☐ eher schlecht ☐ schlecht

Instruktion

Im Folgenden bearbeiten Sie bitte einige Aufgaben zum logischen Denken.

Diese Aufgaben bestehen immer aus zwei Aussagen. Bei der Beantwortung der Aufgaben sollen Sie davon ausgehen, dass diese Aussagen wahr sind, auch wenn sie Ihrer Erfahrung nicht entsprechen.

Lassen Sie sich nicht davon irritieren, dass bei einigen Aufgaben die ersten Aussagen gleich sind. Da sich diese Aufgaben in ihren zweiten Aussagen unterscheiden, gleich keine Aufgabe einer anderen.

Den beiden Aussagen folgen immer drei Antwortmöglichkeiten, **von denen immer nur eine richtig ist.**

Kreuzen Sie bei jeder Aufgabe die Lösung an, von der Sie glauben, dass sie die richtige ist.

Die Möglichkeit „Ich kann diese Aufgabe nicht lösen“ soll nur gewählt werden, wenn Sie mit einer Aufgabe überhaupt nichts anfangen können.

Eine Aufgabe könnte zum Beispiel so aussehen:

Alle blonden Buben sind groß,

Josef ist ein großer Bub.

☐ Josef ist blond.

☐ Josef ist nicht blond.

☐ Vielleicht ist Josef blond, vielleicht aber auch nicht.

☐ Ich kann diese Aufgabe nicht lösen.

In dieser Aufgabe ist die dritte Antwort richtig, da man aufgrund der zwei Aussagen nicht wissen kann, ob Josef blond ist oder nicht.

In den Aufgaben, die Sie bearbeiten sollen, kommen Aussagen vor wie zum Beispiel:

Wenn Josef krank ist, dann liegt er im Bett.

Bedenken Sie, dass diese Aussage

NICHT bedeutet, dass Josef nur dann im Bett liegt, wenn er krank ist. Diese Aussage bedeutet **NICHT**: *Wenn Josef im Bett liegt, dann ist er krank.*

Es kann auch andere Gründe geben, dass Josef im Bett liegt (z.B. wenn er schläft).

Lesen Sie jede Aufgabe sorgfältig durch und **denken Sie über jede Aufgabe neu nach.**

Achten Sie auch darauf, ob in den Aussagen das Wort „nicht“ vorkommt.

Halten Sie sich nicht zu lange mit einer Aufgabe auf, da Sie **alle 35 Aufgaben** beantworten sollen.

Bedenken Sie, dass Sie die Aufgaben nur aufgrund der jeweils vorgegebenen

Aussagen beantworten sollen, unabhängig von Ihren Erfahrungen.

Wie gut glauben Sie, solche Aufgaben lösen zu können?

☐ gut

☐ eher gut

☐ eher schlecht

☐ schlecht

VIEL SPASS UND GUTES GELINGEN!

1.

1. Wenn die Sonne scheint, trägt Tina einen roten Rock.

Tina trägt keinen roten Rock.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

2. Wenn die Sonne scheint, trägt Tina einen roten Rock.

Die Sonne scheint nicht.

- ☐ Tina trägt einen roten Rock.
- ☐ Tina trägt keinen roten Rock.
- ☐ Vielleicht trägt Tina einen roten Rock, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

3. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.

X gehört zur Gruppe B.

- ☐ X hat die Eigenschaft c.
- ☐ X hat nicht die Eigenschaft c.
- ☐ Vielleicht hat X die Eigenschaft c, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

4. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.

Y gehört zur Gruppe F.

- ☐ Y hat die Eigenschaft g.
- ☐ Y hat nicht die Eigenschaft g.
- ☐ Vielleicht hat Y die Eigenschaft g, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

5. Wenn die Sonne scheint, trägt Tina einen roten Rock.

Die Sonne scheint.

- ☐ Tina trägt einen roten Rock.
- ☐ Tina trägt keinen roten Rock.
- ☐ Vielleicht trägt Tina einen roten Rock, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

6. Wenn es **nicht** Abend ist, geht die Sonne unter.

Die Sonne geht unter.

- ☐ Es ist Abend.
- ☐ Es ist nicht Abend.
- ☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

7. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.

X hat die Eigenschaft c.

- ☐ X gehört zur Gruppe B.
- ☐ X gehört nicht zur Gruppe B.
- ☐ Vielleicht gehört X zur Gruppe B, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

8. Wenn es Abend ist, geht die Sonne auf.

Die Sonne geht auf.

- ☐ Es ist Abend.
- ☐ Es ist nicht Abend.
- ☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

9. Wenn es Abend ist, geht die Sonne auf.

Es ist Abend.

- ☐ Die Sonne geht auf.
- ☐ Die Sonne geht nicht auf.
- ☐ Vielleicht geht die Sonne auf, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

10. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.

Y gehört nicht zur Gruppe F.

- ☐ Y hat die Eigenschaft g.
- ☐ Y hat nicht die Eigenschaft g.
- ☐ Vielleicht hat Y die Eigenschaft g, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

11. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.

X gehört nicht zur Gruppe B.

- ☐ X hat die Eigenschaft c.
- ☐ X hat nicht die Eigenschaft c.
- ☐ Vielleicht hat X die Eigenschaft c, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

12. Wenn es **nicht** Abend ist, geht die Sonne unter.

Die Sonne geht nicht unter.

- ☐ Es ist Abend.
- ☐ Es ist nicht Abend.
- ☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

13. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.

Y hat nicht die Eigenschaft g.

- ☐ Y gehört zur Gruppe F.
- ☐ Y gehört nicht zur Gruppe F.
- ☐ Vielleicht gehört Y zur Gruppe F, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

14. Wenn die Sonne scheint, trägt Tina einen roten Rock

Tina trägt einen roten Rock.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

15. Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.

Maria hat gute Laune.

- ☐ Maria schreibt in Physik eine gute Note.
- ☐ Maria schreibt in Physik keine gute Note.
- ☐ Vielleicht schreibt Maria in Physik eine gute Note, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

16. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Peter trägt keine blaue Hose.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

17. Wenn Y zur Gruppe F gehört, dann hat Y die Eigenschaft g.

Y hat die Eigenschaft g.

- ☐ Y gehört zur Gruppe F.
- ☐ Y gehört nicht zur Gruppe F.
- ☐ Vielleicht gehört Y zur Gruppe F, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

18. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Peter trägt eine blaue Hose.

- ☐ Die Sonne scheint.
- ☐ Die Sonne scheint nicht.
- ☐ Vielleicht scheint die Sonne, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

19. Wenn es **nicht** Abend ist, geht die Sonne unter.

Es ist nicht Abend.

- ☐ Die Sonne geht unter.
- ☐ Die Sonne geht nicht unter.
- ☐ Vielleicht geht die Sonne unter, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

20. Wenn X **nicht** zur Gruppe B gehört, dann hat X die Eigenschaft c.

X hat nicht die Eigenschaft c.

- ☐ X gehört zur Gruppe B.
- ☐ X gehört nicht zur Gruppe B.
- ☐ Vielleicht gehört X zur Gruppe B, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

21. Wenn es Abend ist, geht die Sonne auf.

Die Sonne geht nicht auf.

- ☐ Es ist Abend.
- ☐ Es ist nicht Abend.
- ☐ Vielleicht ist es Abend, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

22. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Die Sonne scheint nicht.

- ☐ Peter trägt eine blaue Hose.
- ☐ Peter trägt keine blaue Hose.
- ☐ Vielleicht trägt Peter eine blaue Hose, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

23. Wenn es Abend ist, geht die Sonne auf.

Es ist nicht Abend.

- ☐ Die Sonne geht auf.
- ☐ Die Sonne geht nicht auf.
- ☐ Vielleicht geht die Sonne auf, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

24. Wenn die Sonne **nicht** scheint, trägt Peter eine blaue Hose.

Die Sonne scheint.

- ☐ Peter trägt eine blaue Hose.
- ☐ Peter trägt keine blaue Hose.
- ☐ Vielleicht trägt Peter eine blaue Hose, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

25. Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.

Maria hat keine gute Laune.

- ☐ Maria schreibt in Physik eine gute Note.
- ☐ Maria schreibt in Physik keine gute Note.
- ☐ Vielleicht schreibt Maria in Physik eine gute Note, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

26. Wenn es **nicht** Abend ist, geht die Sonne unter.

Es ist Abend.

- ☐ Die Sonne geht unter.
- ☐ Die Sonne geht nicht unter.
- ☐ Vielleicht geht die Sonne unter, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

27. Wenn Herr Müller gute Laune hat, fährt er nach Berlin.

Herr Müller hat keine gute Laune.

- ☐ Herr Müller fährt nach Berlin.
- ☐ Herr Müller fährt nicht nach Berlin.
- ☐ Vielleicht fährt Herr Müller nach Berlin, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

28. Wenn C **nicht** auf D trifft, dann entscheidet sich C für m.

C entscheidet sich für m.

- ☐ C trifft auf D.
- ☐ C trifft nicht auf D.
- ☐ Vielleicht trifft C auf D, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

29. Wenn es **nicht** regnet, ist die Straße nass.

Die Straße ist nicht nass.

- ☐ Es regnet.
- ☐ Es regnet nicht.
- ☐ Vielleicht regnet es, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

30. Wenn A auf B trifft, dann entscheidet sich A für c.

A entscheidet sich nicht für c.

- ☐ A trifft auf B.
- ☐ A trifft nicht auf B.
- ☐ Vielleicht trifft A auf B, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

31. Wenn Herr Müller gute Laune hat, fährt er nach Berlin.

Herr Müller fährt nach Berlin.

- ☐ Herr Müller hat gute Laune.
- ☐ Herr Müller hat keine gute Laune.
- ☐ Vielleicht hat Herr Müller gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

32. Wenn Frau Schuster **keine** gute Laune hat, geht sie spazieren.

Frau Schuster geht nicht spazieren.

- ☐ Frau Schuster hat gute Laune.
- ☐ Frau Schuster hat keine gute Laune.
- ☐ Vielleicht hat Frau Schuster gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

33. Wenn A auf B trifft, dann entscheidet sich A für c.

A entscheidet sich für c.

- ☐ A trifft auf B.
- ☐ A trifft nicht auf B.
- ☐ Vielleicht trifft A auf B, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

34. Wenn Frau Schuster **keine** gute Laune hat, geht sie spazieren.

Frau Schuster geht spazieren.

- ☐ Frau Schuster hat gute Laune.
- ☐ Frau Schuster hat keine gute Laune.
- ☐ Vielleicht hat Frau Schuster gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

35. Wenn Maria gute Laune hat, schreibt sie in Physik eine gute Note.

Maria schreibt in Physik eine gute Note.

- ☐ Maria hat gute Laune.
- ☐ Maria hat keine gute Laune.
- ☐ Vielleicht hat Maria gute Laune, vielleicht aber auch nicht.
- ☐ Ich kann diese Aufgabe nicht lösen.

Wie gut konnten Sie diese Aufgaben lösen?

- ☐ gut ☐ eher gut ☐ eher schlecht ☐ schlecht

Vielen Dank für Ihre Hilfe!

Anhang 3

Kurzzusammenfassung

Ziel dieser Untersuchung war, unter Berücksichtigung des Geschlechterstereotyps, die Erhebung schlussfolgernden Denkens bei fortgeschrittenen Studierenden der Studienrichtungen Mathematik, Physik, Chemie (zusammengefasst Naturwissenschaften) und Psychologie im Sinne der Stadientheorie Piagets. Mittels des Syllogismustests SDV (Schlussfolgerndes Denken Verbal) wurde an 205 Psychologiestudierenden (w = 143, m = 62) und 190 Naturwissenschaftsstudierenden (w = 41, m = 149) schlussfolgerndes Denken erhoben. Wie vermutet erreichten Studierende der Naturwissenschaften eine höhere SDV-Testleistung als Studierende der Psychologie. Im Allgemeinen wurden keine Unterschiede in der SDV-Testleistung zwischen Männern und Frauen festgestellt. Die Studienleistung, operationalisiert anhand von Noten, wies einen schwachen Zusammenhang mit der SDV-Testleistung auf. Durch Anwendung des Mixed Rasch-Modells konnten, im Sinne der Stadientheorie Piagets, neben einem konkret-operationalen Stadium noch 4 Übergangsstadien identifiziert werden, jedoch kein Stadium der formalen Operationen. Zusätzlich wurden dem SDV spezifische Items hinzugefügt, welche zur Erfassung von Geschlechterstereotypen eine latente Genderinformation inkludierten. Von Interesse war, ob derart transportierte latente Geschlechterstereotypen unterschiedliche Lösungswahrscheinlichkeiten bei weiblichen und männlichen Studierenden im SDV bewirken. Die Ergebnisse zeigten, dass die Genderinformation keinen Einfluss auf die Lösungswahrscheinlichkeit der spezifischen Items bei weiblichen und männlichen Studierenden hat. Zudem wurde das Fähigkeitsselbstkonzept der Studierenden vor und nach der Testung erhoben. Männer schätzten sich zu beiden Erhebungszeitpunkten höher ein als Frauen. Zudem fiel die Einschätzung zum Fähigkeitsselbstkonzept nach der Testung für beide Geschlechter niedriger aus.

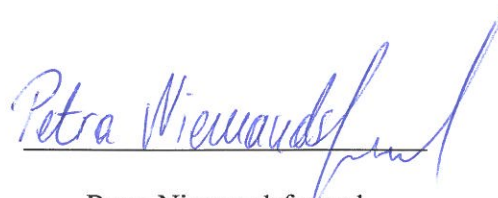
Anhang 4

Abstract

The purpose of this study was to examine reasoning in advanced students of mathematics, physics, chemistry (summarized as natural sciences) and psychology, in terms of the cognitive developmental theory of Piaget, considering gender stereotypes. Via syllogism test DRV (deductive reasoning verbal) 205 students of psychology ($f = 143$, $m = 62$) and 190 students of natural science ($f = 41$, $m = 149$) were examined to survey reasoning. As hypothesized, students of natural science achieved a higher DRV-test performance than students of psychology. Generally no gender differences in terms of solving syllogisms were found. The study performance, which was operationalized by grades, indicated a weak correlation with the DVR-test performance. Using a Mixed Rasch-Model, a concrete-operational stage according to Piaget was found. Also four further intermediate stages could be identified but no formal stage. Furthermore, specific items were added to the DRV which included latent gender information to acquire gender stereotypes. It was of interest, if gender stereotypes, transported in such a way, cause different solutions relating to syllogisms between female and male students. Obviously, the gender information has no influence on solving gender items between female and male students. Data of self-concept were collected before and after the test. It could be shown that male students have a higher self-concept than female students at either time of measurement. Moreover, the self-concept estimation was inferior in both female and male students after the test.

Hiermit bestätige ich, dass die vorliegende Arbeit in allen relevanten Teilen selbstständig durchgeführt wurde.

Wien, März 2015

A handwritten signature in blue ink, reading "Petra Niemandsfreund", written over a horizontal line.

Petra Niemandsfreund

Anhang 5

CURRICULUM VITAE

Persönliche Daten

Name: Petra Niemandsfreund

Schul- und Weiterbildung

2004- 2015 Studium an der Universität Wien, DS Psychologie

1993 - 1998 Höhere Lehranstalt für wirtschaftliche Berufe, 1030 Wien

Beruflicher Werdegang und Praktika

02/2014 – 01/2015	Bildungskarenz
02/2012 – 01/2014	Administrative Coordinator Regulatory Affairs Amgen GmbH, 1040 Wien
10/2010 – 06/2011	Projektstudium: Mitarbeit an der Evaluation des Schulentwicklungsprojekts „neue Wiener HAK“ Universität Wien im Arbeitsbereich Bildungspsychologie & Evaluation
09/2009 – 01/2012	Associate Admin Coordinator Amgen GmbH, 1040 Wien
04/2009 – 07/2009	Praktikantin Psychologie Kolpinghaus „Gemeinsam Leben“, 1100 Wien
04/2006 – 02/2009	Rezeptionistin Canon Austria GmbH, 1100 Wien
12/2005 – 04/2006	Sekretariat Merkur Treuhand Wirtschafts- und Steuerberatungsges.m.b.H, 1130 Wien
09/2004 – 09/2005	Front Desk Assistent Sandoz GmbH, 1220 Wien
01/2001 – 09/2004	Außendienstmitarbeiterin im Bereich KMU ONE GmbH, 1210 Wien
05/1999 – 12/2000	Assistentin: - Assistentin Leiter Direktvertrieb - Project Assistant of Head of Direct Sales ONE GmbH, 1210 Wien

07/1998 – 04/1999 **Sekretariat**
IG Facility Management GmbH, 1010 Wien

Besondere Kenntnisse und Fähigkeiten

EDV: MS Office: Outlook, Excel, Word, Power Point;
 SPSS, WINMIRA

Maschinschreiben: 10-Finger-System

Fremdsprachen: Englisch: in Wort und Schrift
 Spanisch: Grundkenntnisse
 Französisch: Grundkenntnisse



Wien, März 2015