



universität
wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

Unterschiede in der Bewertung der Leistungen von
Schülerinnen und Schülern durch zukünftige
Physiklehrkräfte

verfasst von / submitted by

Karin Herta Nigischer, BSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Master of Science (MSc)

Wien, 2017 / Vienna 2017

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

A 066 840

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

Masterstudium Psychologie

Betreut von / Supervisor:

Univ.-Prof. Mag. DDr. Christiane Spiel

Inhaltsverzeichnis

Zusammenfassung.....	3
Abstract.....	4
I Einleitung	5
II Theoretischer Teil	7
1 Geschlechterunterschiede	7
2 Ansätze zur Erklärung von Geschlechterunterschieden	8
2.1. Biologische Theorien.....	8
2.2. Soziale Lerntheorie	9
2.3. Kognitive Entwicklungstheorie	10
2.4. Geschlechtsschematheorie	10
3 Geschlechterstereotype	11
4 Einfluss der Eltern	14
5 Einfluss der Lehrkräfte.....	16
6 Auswirkungen	17
6.1. Fähigkeitsselbstkonzept.....	17
6.2. Stereotype Threat	20
7 Benotung/Beurteilung	22
8 Die Rolle von Erfahrung/Expertise.....	23
9 Need for Cognitive Closure.....	25
10 Fragestellungen und Hypothesen	27
III Empirischer Teil	30
11 Methode.....	30
11.1 Untersuchungsdesign und Vorgehen	30
11.2 Erhebungsinstrument	31
11.3 Stichprobenbeschreibung	33

12 Ergebnisse	35
12.1 stereotype Beurteilung.....	35
12.2 Zusammenhang mit dem Geschlecht der „Lehrkraft“	36
12.3 Zusammenhang mit dem gewählten Zweifach	37
12.4 Zusammenhang mit Schwerpunktsetzung der Schülerin/des Schülers	38
12.5 Need for cognitive closure	38
12.6 Mögliche weitere Einflussfaktoren	39
IV Diskussion.....	41
13 Zusammenfassung und Diskussion der Ergebnisse	41
13.1 Unterschiedliche Beurteilung von Schülerinnen und Schülern	41
13.2 Weitere Zusammenhänge	43
13.3 Bedürfnis nach kognitiver Geschlossenheit.....	45
14 Limitationen.....	45
15 Conclusio	48
Abbildungsverzeichnis.....	50
Tabellenverzeichnis.....	50
Literaturverzeichnis	51
Anhang	61

Zusammenfassung

Ziel dieser Arbeit war es zu untersuchen, inwieweit sich bei Lehramtsstudierenden des Unterrichtsfaches Physik die Beurteilung einer fiktiven Antwort von SchülerInnen der Sekundarstufe je nach SchülerInnengeschlecht unterscheidet. Es handelt sich um die Replikation einer Studie von Hofer (2015) mit veränderter Stichprobe. Dazu wurde eine Onlineerhebung mit Physik - Lehramtsstudierenden österreichischer Universitäten durchgeführt, wobei der Rücklauf von 32 (12 weibliche, 20 männliche Teilnehmer) geringer als erwartet war. Vorgegeben waren eine Prüfungsfrage aus dem Bereich Newton'sche Mechanik und die fiktive Antwort einer Schülerin beziehungsweise eines Schülers, die benotet werden musste. Im Anschluss mussten noch 16 Items eines Kurzfragebogens zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit beantwortet werden, um zu überprüfen, ob ein Zusammenhang zwischen dieser Persönlichkeitseigenschaft und Unterschieden bei der Beurteilung besteht. Die Ergebnisse zeigten, dass die Lehramtsstudierenden Schüler signifikant schlechter beurteilten als Schülerinnen. Dies war ein unerwartetes Ergebnis, da es den Ergebnissen von Hofer (2015), bei der Physiklehrkräfte mit wenig Unterrichtserfahrung Schülerinnen signifikant schlechter beurteilten, genau entgegengesetzt ist. Es fanden sich in der vorliegenden Studie keine Hinweise auf einen Zusammenhang der Beurteilung mit dem Geschlecht der Studierenden, deren gewähltem Zweitfach, der Schwerpunktsetzung der SchülerInnen und dem Bedürfnis nach kognitiver Geschlossenheit.

Mögliche Gründe für diese Ergebnisse und Limitationen der Studie werden diskutiert.

Schlüsselwörter: Geschlechterstereotype, Lehramtsstudierende, Lehrerfahrung, Physikunterricht, Bedürfnis nach kognitiver Geschlossenheit

Abstract

The aim of this study was to investigate to which extent the gender of students of secondary school had influence to the rating when preservice teachers were doing the assessment of a students fictive response in the field of physics.

This study is a replication of a study by Hofer (2015) with a modified sample. An online survey with physics preservice teachers from Austrian universities was conducted, where the response of 32 (12 female, 20 male participants) was lower than expected. All participants had to grade the same fictional answer of a female or male secondary school student to an examination question from the field of Newtonian mechanics. In addition 16 items of a short questionnaire had to be answered to assess the need for cognitive closure of the participants. This was to check, whether there is a relationship between this personality trait and differences in assessment. Results showed, that male school students were significantly graded worse by preservice teachers than female school students. This result was unexpected and not in line with the results from the study by Hofer (2015), which had shown a gender bias in grading against female students by physics teachers with less experience. In the present study, there was no evidence for a relationship between differences in assessment and participants gender, their chosen second subject of study, the focus of the school students and the need for cognitive closure.

Various reasons for the results and the limitations of this study are discussed.

Keywords: gender stereotypes, preservice teachers, teaching experience, physics instruction, need for cognitive closure

I Einleitung

Kinder werden schon früh in ihrer Entwicklung mit Ansichten und Erwartungen konfrontiert, wie Mädchen oder Buben sein sollen oder wo ihre geschlechtsspezifischen besonderen Fähigkeiten liegen. Diese Einstellungen werden von den Kindern gelernt und haben Einfluss auf die Entwicklung ihres Selbstkonzeptes. Bereits in der ersten Klasse Volksschule gibt es signifikante Unterschiede in der Leistung und im Selbstkonzept von Mädchen und Buben. Während Buben sich mehr in Mathematik zutrauen, schätzen Mädchen ihr Können in Lesen und Rechtschreiben höher ein (Niklas & Schneider, 2012). Diese Unterschiede ziehen sich über die gesamte Schullaufbahn weiter und beeinflussen die Ausbildungswahl der Kinder. Nur 33% der AbsolventInnen von naturwissenschaftlichen Fächern sind Frauen und sogar nur 21% in Studien des Ingenieurs-, Fertigungs- und Bauwesens (OECD, 2015). Auch im Gehalt gibt es noch immer große Unterschiede zwischen den Geschlechtern. „Unabhängig von der Bildungsstufe haben Frauen niedrigere Einkommen als Männer. In Österreich verdient eine 35-44-jährige Frau mit Tertiärabschluss 71% dessen, was ein Mann mit gleichem Alter und Bildungsabschluss verdient. ...“ (OECD, 2015).

Wodurch entstehen diese Unterschiede, die die gesamte Bildungskarriere und Lebensspanne betreffen? Ein Erklärungsansatz ist das Vorhandensein von stereotypen Einstellungen zu Frauen und Männern in der Gesellschaft.

Die vorliegende Arbeit beschäftigt sich mit Geschlechterstereotypen im mathematisch-naturwissenschaftlichen Bereich, deren Entstehung, Einflussfaktoren und Auswirkungen. Der Fokus liegt im Schulbereich und bei den Lehrkräften und ihrem Einfluss auf die Leistung und Motivation der SchülerInnen. Die zugrundeliegende Studie stellt eine Replikation der Studie von Hofer (2015) dar, die Unterschiede in der Bewertung von Leistungen von Schülerinnen und Schülern durch Physiklehrkräfte fand, wobei diesmal die Stichprobe aus Physik-Lehramtsstudierenden bestand.

Dazu werden im Theorieteil in Kapitel 1 zunächst bestehende Geschlechterunterschiede dargestellt. Danach widmet sich Kapitel 2 der Erklärung von Geschlechterunterschieden in der Literatur und Kapitel 3 explizit den Geschlechterstereotypen als einen wesentlichen Erklärungsansatz. Die folgenden Kapitel beleuchten den Einfluss wesentlicher SozialisationsagentInnen, nämlich Eltern (Kapitel 4) und Lehrkräfte (Kapitel 5). Mögliche Auswirkungen von Geschlechterstereotypen werden in Kapitel 6 behandelt. Kapitel 7 gibt einen kurzen Einblick über die Benotung im (österreichischen) Schulsystem. In den letzten beiden theoretischen Kapiteln werden die Rolle von Erfahrung der Lehrkräfte bei der Beurteilung ihrer Schülerinnen (Kapitel 8) und das Konstrukt der Persönlichkeitseigenschaft *need for cognitive closure* (Kapitel 9) näher beleuchtet. Anschließend werden die Zielsetzungen und die aufgestellten Fragestellungen und Hypothesen dargestellt (Kapitel 10).

Im empirischen Teil werden zuerst die Methode (Kapitel 11) und dann die Ergebnisse der Untersuchung (Kapitel 12) berichtet. Zum Abschluss werden die Ergebnisse noch einmal zusammengefasst, in Zusammenhang mit dem theoretischen Hintergrund und weiteren Forschungsergebnissen gebracht (Kapitel 13) und mögliche Einschränkungen der Studie diskutiert (Kapitel 14).

II Theoretischer Teil

1 Geschlechterunterschiede

„Aktuelle Daten und Befunde weisen darauf hin, dass es nach wie vor deutliche geschlechterbezogene Segregationen im österreichischen Schul- und Bildungsbereich gibt, welche in der Folge zu diversen Benachteiligungen am Arbeitsmarkt und in den privaten Beziehungen der Menschen führen. Schule hat die Aufgabe, darauf hinzuwirken, dass Schülerinnen und Schüler ihre Kompetenzen, Interessen und Handlungsspielräume möglichst breit entwickeln können – [...]“ (Bundesministerium für Bildung, 2017).

Bei der Schulwahl zeigen sich in Österreich deutliche Geschlechterdifferenzen. Deutlich mehr als die Hälfte aller SchülerInnen der zehnten Schulstufe besuchen geschlechtstypische Schulformen (PISA, 2015). Im Schuljahr 2015/16 wurden wirtschaftsberufliche und sozialberufliche Schulen hauptsächlich von Mädchen besucht (wirtschaftsberufliche 86,7%, sozialberufliche 79,8%), während bei den technisch gewerblichen Schulen Burschen mit 74,1% deutlich überrepräsentiert waren (Stockinger & Schestak, 2015). 23% der männlichen Universitätsabsolventen beginnen in einem technisch-mathematischen Beruf zu arbeiten, aber nur 7,5% der weiblichen Absolventinnen (OECD, 2011).

Aber nicht nur die Schulwahl zeigt Geschlechterdifferenzen, sondern auch die Leistung der SchülerInnen ist unterschiedlich.

In der länderübergreifenden Schulstudie TIMSS 2011 wies Österreich im internationalen Vergleich sowohl in der Volksschule als auch in der Sekundarstufe eine größere Geschlechterdifferenz zugunsten der Buben im Bereich Mathematik auf. Auch in den Naturwissenschaften erzielten Buben signifikant bessere Leistungen als Mädchen. Nur 6% der Mädchen waren auf der höchsten Kompetenzstufe, im Vergleich zu 11% der Buben. Der größte Unterschied zeigte sich im Fach Erdkunde, gefolgt von Physik (Suchan, Wallner-Paschon, Bergmüller, & Schreiner, 2011).

Auch bei den 15-16 jährigen SchülerInnen sind die Befunde nicht anders. Im internationalen Vergleich haben in Österreich Buben den größten Vorsprung

gegenüber Mädchen hinsichtlich der Leistung in den naturwissenschaftlichen Fächern. Besonders groß ist der Unterschied in der Lernmotivation (PISA, 2015). Demgegenüber drücken Mädchen bereits in der Volksschule eine höhere Lesefreudigkeit aus als Buben. In der Schulstudie PIRLS 2011 stimmten 68% der Mädchen der Aussage „Ich lese gern“ völlig zu, während dies nur 45% der Buben taten (Bruneforth, Lassnigg, Vogenhuber, Schreiner, & Breit, 2016).

Auch bei den Abschlüssen der Sekundär- und Tertiärausbildung zeigen sich Unterschiede zwischen den Geschlechtern. „Im Jahr 2013 wurden insgesamt 58% aller Reife- und Diplomprüfungen von Frauen abgeschlossen“ (Bruneforth et al., 2016, S. 132). Ebenso liegen bei den Bachelorabschlüssen Frauen mit 60% der AbsolventInnen voran, während diese Zahl dann bei Masterabschlüssen auf 55% sinkt. Bei der Beendigung von Doktoratsprogrammen liegen Frauen in Österreich mit 44% dann aber hinter den Männern (OECD, 2015).

Ein viel diskutiertes Beispiel ist auch der Medizinaufnahmetest in Österreich. Im Jahr 2017 waren wiederum zwar rund 60% der TestteilnehmerInnen weiblich, aber nur 53% der Plätze gingen an Frauen (studium.at, 2017).

Wie kommt es nun zu diesen Geschlechterunterschieden? Der nächste Abschnitt widmet sich den verschiedenen Zugängen zur Erklärung von Geschlechterunterschieden.

2 Ansätze zur Erklärung von Geschlechterunterschieden

Verschiedene psychologische Theorien versuchen die Entstehung von Geschlechterunterschieden zu erklären. Eine wesentliche Unterscheidung besteht darin, ob davon ausgegangen wird, dass Geschlechterunterschiede bereits angeboren und damit nicht veränderbar sind oder, dass sie durch Einflüsse aus der Umwelt entstehen.

2.1. Biologische Theorien

Am einfachsten und auch sehr leicht nachvollziehbar erscheinen biologische Erklärungen. Geschlechterunterschiede kommen demnach durch unterschiedliche angeborene Anlagen zustande. Sie werden daher als unveränderbar angenommen (Steins, 2010). Bisher konnten aber weder

evolutionäre oder physiologische Prozesse oder morphologische Gegebenheiten die These von den biologisch bedingten, angeborenen bildungsrelevanten Geschlechtsunterschieden untermauern (Jöstl, Kollmayer, Finsterwald, Schober, & Spiel, 2015).

Biologische Theorien können daher weder situations- und entwicklungsabhängige Variationen in Geschlechterunterschieden, noch Unterschiede zwischen Personen innerhalb einer Geschlechtergruppe erklären (Hannover, 2008).

Besonders die lang vorherrschende These der niedrigeren allgemeinen kognitiven Fähigkeiten von Mädchen wird mittlerweile abgelehnt (Callahan, C. M., 1991 zit. nach Dresel, Heller, Schober, & Ziegler, 2001).

Ebenfalls können psychoanalytische Theorien nicht als empirisch belegt betrachtet werden und finden in der wissenschaftlichen Diskussion daher keine Berücksichtigung mehr (Hannover, 2008).

Wesentlich wichtiger sind soziale und kulturelle Ansätze zur Erklärung.

2.2. Soziale Lerntheorie

Lerntheorien gehen davon aus, dass Verhalten, das belohnt wird, in der Zukunft häufiger gezeigt werden wird. Zweiter bedeutender Aspekt in den Lerntheorien ist das Lernen am Modell. Beobachtetes Verhalten wird dann gezeigt, wenn das Modell für sein Verhalten Belohnung bekommen hat (Hannover, 2008). Die Theorie des Modelllernens geht auf Albert Bandura zurück, der in seinen neueren Arbeiten auch die Bedeutung der kognitiven Prozesse betont (Bandura, 1992; 2001, zit. nach Berk, 2011). Ahmt demnach ein Bub seinen Vater in geschlechtstypischer Weise nach und bekommt dafür Lob, wird er das gezeigte Verhalten häufiger wieder zeigen. Gleichgeschlechtliche Personen erhalten besonders viel Aufmerksamkeit. Die Beobachtung von Art und Häufigkeit der Verhaltensweisen von Männern und Frauen führt zur Erkenntnis, welches Verhalten typisch für welches Geschlecht ist (Carli & Bukatko, 2000). Leaper (2000) konnte in seiner Studie zeigen, dass Mädchen, die im Spiel Haushaltsarbeiten erledigen, von ihren Müttern viel mehr unterstützenden emotionalen Zuspruch bekamen (in Berk, 2011). Diese positive Verstärkung einer Verhaltensweise, die durch Beobachtung vom Kind verstärkt Frauen zugeschrieben wurde, kann nun dazu führen, dass das Kind dieses Verhalten

noch häufiger zeigt. Einfluss haben aber neben den Eltern auch Lehrkräfte und Gleichaltrige (Maccoby, 1990). „Zusammenfassend sprechen die Befunde über Sozialisationspraktiken im Elternhaus und in der Schule dafür, dass Mädchen und Jungen wahrscheinlicher für geschlechtstypisiertes Verhalten belohnt werden und häufiger Gelegenheit zur Beobachtung geschlechtstypisierten Verhaltens haben“ (Hannover, 2008; S.352).

Für die Annahme der sozialen Lerntheorie spricht, dass geschlechtsangemessenes Verhalten bereits sehr früh (Kindergartenalter) beobachtet werden kann (Berk, 2011).

Warum sich die Geschlechter aber in ihrer Identität oder Selbstwahrnehmung und in ihren Präferenzen unterscheiden, dafür braucht es einen anderen Erklärungsansatz. Diesen liefern kognitive Theorien.

2.3. Kognitive Entwicklungstheorie

Kognitive Theorien besagen, dass kognitive Konzepte (das Wissen über das eigene Geschlecht) die Geschlechtertypisierung steuern und damit die Präferenzen und das Verhalten beeinflussen (Hannover, 2008). In der kognitiven Entwicklungstheorie von Kohlberg (1966) geht dieser davon aus, dass Kinder sich aktiv in Geschlechterrollen sozialisieren. Die Selbstwahrnehmung des eigenen Geschlechtes entsteht dabei vor dem geschlechtstypischen Verhalten. Bereits im Kindergartenalter kann ein Kind sein eigenes Geschlecht zuverlässig benennen und lernt, dass sich das Geschlecht einer Person auch mit der Zeit nicht ändert. Im Volksschulalter hat ein Kind dann so weit Geschlechtskonstanz erworben, dass es weiß, dass das Geschlecht einer Person über die Zeit und über Situationen hinweg unveränderbar ist. Ein Bub, der also mit Puppen spielt oder sich wie ein Mädchen kleidet, bleibt ein Bub. Wenn ein Kind sich selbst als Bub oder Mädchen wahrnimmt, beginnt es dann auch sich geschlechtstypisch zu verhalten (Kohlberg, 1966).

2.4. Geschlechtsschematheorie

Diese Theorie von Bem (1981) versucht Aspekte der sozialen Lerntheorie und der kognitiven Entwicklungstheorie miteinander zu verbinden. Auch hier wird von einer aktiven Rolle des Kindes beim Erwerb von Geschlechtsrollen ausgegangen. Es kommt zu einer Wechselwirkung mit der Umwelt. Das Kind

beobachtet seine Umgebung, nimmt die unterschiedlichen Geschlechter wahr, ihre sozialen Rollen, die Assoziationen mit dem jeweiligen Geschlecht und wie die Gesellschaft mit den Geschlechtern umgeht. Dies hilft ihm bei der Einordnung und Verarbeitung von Informationen über die Geschlechter und seine eigene Person. Das Kind organisiert diese Informationen in geschlechtsbasierten Kategorien (Geschlechtsschemata). Dies geschieht ohne, dass sich das Kind dieses Prozesses bewusst ist. Nimmt sich das Kind als Bub oder Mädchen wahr, beginnt es gezielt geschlechtstypische Informationen aufzunehmen. Außerdem beginnt es sein Verhalten den kulturellen geschlechtstypischen Normen anzupassen. Es orientiert sich dabei auch an den Eltern oder der Schule. Gleichzeitig passt das Kind die Geschlechtsschemata aber auch an sein Verhalten und seine Vorlieben an. Geschlechtsschemata differieren dadurch auch von Kind zu Kind und ändern sich im Laufe seiner Entwicklung.

„Aber warum werden bestimmte Verhaltensweisen und Eigenschaften so stabil als für ein Geschlecht „richtig“ oder passend angesehen (Jöstl et al., 2015; S. 16)?“ Eine Antwort auf diese Frage kann das Konzept der Geschlechterstereotype liefern.

3 Geschlechterstereotype

Ein *Stereotyp* ist die kognitive Komponente eines Vorurteils. Es ist eine Verallgemeinerung über eine Gruppe von Menschen hinweg, wobei allen Mitgliedern mehr oder weniger dieselben Merkmale zugeschrieben werden. Nicht berücksichtigt wird dabei die tatsächliche Variation der Merkmale unter den Mitgliedern. (Aronson, Wilson, & Akert, 2014) Ein Stereotyp muss aber nicht zwangsläufig negativ sein. Es „kann ein nützliches, adaptives geistiges Werkzeug sein, um die soziale Welt zu organisieren“ (Aronson et al., 2014, S.513). Die Welt um uns ist hoch komplex und uns fehlt die Zeit, zu allen Themen differenzierte, ausgereifte Einstellungen zu haben. Für einige Themen, die uns wichtig erscheinen, nehmen wir uns Zeit und Energie, um hoch differenzierte Einstellungen zu haben, bei anderen Themen nutzen wir aber - da wir nur eine begrenzte Fähigkeit zur Informationsverarbeitung haben – einfache, lückenhafte

Überzeugungen und Faustregeln (Aronson et al., 2014). Dabei gilt das „Gesetz der geringsten Anstrengung“ (Allport, 1954). Stereotype können uns demnach helfen, kognitive Ressourcen zu sparen.

Da aber individuelle Unterschiede innerhalb einer Gruppe manchmal „negiert“ werden, können auch an sich positive Stereotype für eine Gruppe für deren Mitglieder einschränkend und negativ sein (Aronson et al., 2014).

Stereotype geben also vor, wie jemand ist (deskriptiv) oder sein soll (präskriptiv) (Eckes, 2008). Zahlreiche Vorgaben gibt es dazu zu Frauen und Männern, sogenannte Geschlechterstereotype, die in Folge näher beleuchtet werden sollen.

Geschlechterstereotype sind Sets von Überzeugungen zu Persönlichkeitseigenschaften (Interessen, Rollen, Kompetenzen) von Frauen und Männern (Ashmore & Del Boca, 1979). Sie beinhalten eine deskriptive Komponente, die beschreibt, welche Charakteristika Frauen und Männer haben und eine präskriptive Komponente, die beschreibt, welche Charakteristika Frauen und Männer besitzen sollen (Burgess & Borgida, 1999).

Typische Charakteristika, die Frauen zugeschrieben werden, sind solche, die mit Fürsorge für andere (communion) in Verbindung stehen. Sie werden als verständnisvoll, warm, emotional, feminin, gemeinschaftsorientiert beschrieben. Männern hingegen werden Attribute wie durchsetzungsstark, dominant, selbstbehauptend oder maskulin (agency) zugeschrieben. Diese Zuordnung erweist sich als sehr stabil über die Zeit und ist universell (Eckes, 2008). „Allein die in vielen Kulturen gleichermaßen anzutreffende Unterteilung von Spielwarenabteilungen in rosafarbene Puppenecken „für Mädchen“ und blau oder dunkel gehaltene Technik- und Autoecken „für Jungen“ lassen universelle Geschlechtsstereotype bezüglich bestimmter thematischer Bereiche deutlich werden“ (Seidel & Krapp, 2014, S. 154). Diesen typischen Zuordnungen begegnet man täglich in der Werbung (die fürsorgliche Mutter, die sich um die aufgeschürften Knie, das Essen, die „gesunde“ Jause ihrer Kinder kümmert; die besten Freundinnen, die lachend Kaffee trinken; der Mann, der alleine ein Haus baut, renoviert; der rasante, sportliche, autobegeisterte Mann), in Kinderbüchern und auch in Schulbüchern. So zeigte sich bei einer Analyse von Schulbüchern für die erste bis vierte Schulstufe, dass darin nach wie vor Männer häufiger in

Berufssituationen, Frauen häufiger in der Familie und in der Freizeit dargestellt werden (Finsterwald & Ziegler, 2007).

Eine Erklärung für die konsistente Zuschreibung von fürsorglichen Merkmalen zu Frauen und von aufgabenbezogenen Merkmalen zu Männern kann die Theorie der sozialen Rolle liefern. Diese geht auf Eagly (1987) zurück, die meinte, dass Menschen annehmen, dass Männer und Frauen die Merkmale aufweisen, die charakteristisch für ihre jeweilige Rolle sind. Die bereits oben erwähnten fürsorglichen Merkmale von Frauen ergeben sich also daraus, dass diese überwiegend als Hausfrauen tätig sind, oder in Berufen mit niedrigem sozialen Status. Männer übernehmen häufiger die Rolle des Ernährers und arbeiten in Berufen mit hohem sozialen Status. Daher werden ihnen andere Merkmale zugeschrieben. Durch die Beobachtung von Menschen in ihren sozialen Rollen werden diesen unmittelbar die für diese Rolle typischen Eigenschaften zugewiesen (Eagly, 1987).

Auch wenn zu einer großen Zahl von Männern und Frauen diese Zuordnungen passen, treffen sie nicht auf alle Individuen zu und sind vor allem in der Ausprägung unterschiedlich. Geschlechterstereotype sind also nicht grundsätzlich falsch, übersehen aber die Breite der einzelnen Merkmale (z.B. ein wenig emotional bis stark emotional,...) und vor allem, dass eine Person sowohl emotional als auch durchsetzungsstark sein kann (Aronson et al., 2014).

Es kann in Folge der wahrgenommenen Einstellungen aus der Umwelt zur „self-fulfilling-prophecy“ kommen. Menschen haben eine Erwartung über eine andere Person, die ihren Umgang mit dieser beeinflusst. Diese wiederum verhält sich unbewusst nun so, dass die ursprünglichen Erwartungen erfüllt werden (Scherr, Madon, Gyll, Willard, & Spoth, 2011; Madon, Willard, Gyll, & Scherr, 2011). So nehmen zum Beispiel SchülerInnen die Einstellungen ihrer Lehrkraft wahr und reagieren auf die Einstellungen dieser. Diese Wahrnehmungen und Reaktionen können dadurch zu einer mehr oder weniger guten schulischen Leistung führen. Dies ist unabhängig davon, wie genau die Einstellung der Lehrkraft ist (Friedrich, Flunger, Nagengast, Jonkmann, & Trautwein, 2015). Jussim, Eccles & Madon (1996) konnten zeigen, dass bei Mädchen der self-fulfilling-prophecy Effekt zwei- bis dreimal größer war als bei Buben. Es erfüllt sich also, was die Lehrkraft unbewusst vermittelt.

Geschlechterstereotype finden sich sehr häufig im Schulkontext bei SchülerInnen wieder. „Mädchen strengen sich mehr an“, „Schülerinnen sind schlecht im Rechnen“, „Buben sind schlechte Leser“, „Schüler sind faul“. Dazu gibt es bereits zahlreiche Studien, die diese Einstellungen widerspiegeln (siehe: Dresel et al., 2001; Fennema, Peterson, Carpenter & Lubinski, 1990; Retelsdorf, Schwartz & Asbrock, 2015).

Schneiden Schüler bei einem Test in Mathematik schlecht ab, wird dies von Lehrkräften in der Grundschule häufig auf mangelnde Anstrengung zurückgeführt, bei Schülerinnen aber auf mangelnde Fähigkeit (Tiedemann, 2002). Dieses Ergebnis zeigte sich auch im Sekundarschulbereich bei einer Studie von Espinoza, Luz Fontes und Arms-Chavez (2014) mit 64 Lehrkräften. Zusätzlich zu diesem Stereotyp schätzten Lehrkräfte ihren besten Schüler im Gegensatz zur besten Schülerin als konkurrenzfähiger, logisch denkender, abenteuerlustiger ein. Schülern wurde auch mehr Freude an Mathematik zugeschrieben und dass sie freiwillig mehr an mathematische Probleme herangehen würden und unabhängiger im mathematischen Denken seien (Fennema et al., 1990).

Vor allem Mädchen vertrauen auf die Einschätzung der Lehrkräfte, um ihre eigenen Fähigkeiten einzuschätzen (Dickhäuser & Meyer, 2006). Welche Probleme durch diese Typisierungen von Schulkindern entstehen, wird in Kapitel 6 noch ausführlich behandelt.

Aber auch andere die Einstellungen anderer SozialisationsagentInnen kann einen Einfluss ausüben. Die nächsten beiden Kapitel befassen sich näher mit Befunden zu den Einflüssen von Eltern und Lehrkräften.

4 Einfluss der Eltern

Eltern haben einen großen Einfluss auf die Weitergabe von Geschlechterstereotypen in ihre Kinder (Carli et al., 2000). Dieser nimmt im Laufe der Entwicklung zwar ab und es kommen andere Modelle und Einflüsse dazu (Lehrkräfte, Freunde, Peers,...), doch sind die Eltern die ersten prägenden Personen für ein Kind (Gelman & Nguyen, 2004). Sie beeinflussen ihre Kinder durch die Auswahl des Spielzeuges (Puppen versus Autos,...), der Kleidung, der

Freunde oder der Kinderbücher. So empfinden Eltern geschlechterstereotype Spielsachen für ihr Kind erwünschter als Spielsachen, die für das andere Geschlecht typisch sind, auch wenn sie keine geschlechterstereotypen Einstellungen angeben (Kollmayer, Schultes, Schober, & Spiel, 2016 in Kollmayer, Schober, & Spiel, 2016). Deshalb ist zu Hause mehr geschlechterstereotypes Spielzeug vorhanden und dies wird beim gemeinsamen Spiel bevorzugt verwendet (Hodosi, 2014). Eltern fördern bei Buben auch eher deren Selbständigkeit und Leistung, während sie bei Mädchen Nähe, Unselbständigkeit und Einhaltung der Regeln forcieren (Berk, 2011). Dies führt dazu, dass Mädchen eher hausfrauliche Tätigkeiten und Buben verstärkt praktische Tätigkeiten trainieren (Li & Wong, 2016). Eltern trauen ihren Kindern Unterschiedliches zu und setzen auch andere Erwartungen in ihr Kind, je nach Geschlecht des Kindes. So konnten Dresel et al. (2001) mittels Elternbefragungen zur Eignung für Studienfächer folgendes zeigen: Mädchen werden von ihren Eltern als geeigneter für Lehramt an Grundschulen und Sprachstudien gesehen, Buben als geeigneter für Studien zu Mathematik oder Maschinenbau. Dies zeigte sich umso stärker, je genderbezogen konservativer die Eltern waren. Diese Befunde passen auch zu Studien, die zeigten, dass Mädchen mehr sprachlich unterstützt werden und mit ihnen häufiger über Emotionen gesprochen wird (Gelman et al., 2004), während Buben vermehrt Sachinformationen erhalten (Crowley, Callanan, Tenenbaum, & Allen, 2001). Eltern trauen demnach ihren Söhnen weniger im sprachlichen Bereich, Töchtern weniger im mathematisch-naturwissenschaftlichen Bereich zu und erwarten in diesen Fächern in der Schule auch schlechtere Ergebnisse (Dresel et al., 2001). Dies kann wieder im Sinne der „self-fulfilling-prophecy“ zu tatsächlich schlechteren Schulleistungen im entsprechenden Fach führen.

Kinder werden bereits durch Erwartungen und Vorerfahrungen außerhalb der Schule geprägt. Hat nun die Lehrkraft ebenfalls starke geschlechterstereotype Einstellungen, wird der Effekt noch verstärkt.

5 Einfluss der Lehrkräfte

Da Kinder ab dem Alter von sechs Jahren für neun oder mehr Jahre einen großen Teil ihrer Zeit in der Schule verbringen, stellt die Schule neben dem Elternhaus einen bedeutenden Einflussfaktor bei der Entstehung von Geschlechterstereotypen dar.

Die Befunde im Bereich Schule decken sich sehr stark mit den bisher besprochenen zu den Einstellungen der Eltern.

Da die Hypothesen dieser Arbeit auf die spezielle Beurteilung im Lehrfach Physik ausgerichtet sind, werden in den folgenden Abschnitten vor allem Theorien zu den mathematisch-naturwissenschaftlichen Fächern dargestellt.

Bereits in der Grundschule halten Lehrkräfte Buben für begabter in Mathematik als Mädchen und erklären schlechte Leistungen von Mädchen eher durch mangelnde Begabung als mangelnde Anstrengung und mangelnden Fleiß (Fennema et al., 1990). Daran ändert sich auch im Sekundarschulbereich nichts. Ungefähr ein Drittel der Mathematik- und Physiklehrkräfte schätzen Mädchen für weniger begabt in ihrem Unterrichtsfach ein, unabhängig vom eigenen Geschlecht (Ziegler, Kuhn, & Heller, 1998). Auch wenn es keine Geschlechterdifferenzen in standardisierten Tests gibt, halten Lehrkräfte Buben für begabter in Mathematik (Tiedemann, 2000). Dass Mädchen doppelt so viel Lernaufwand in Mathematik haben wie Buben und für dieses Fach weniger begabt sind, davon sind bereits Lehramtsstudierende überzeugt (Rustemeyer, 1999). Sie würden, ebenso wie viele Lehrkräfte, Buben eher an mathematisch-naturwissenschaftlich orientierte weiterführende Schulen empfehlen und Mädchen an sprachlich orientierte (Nürnberger, Nerb, Schmitz, Keller, & Sütterlin, 2016).

Unterschiede zeigen sich auch in der Zuwendung. Buben erhalten im Mathematikunterricht mehr Lob und werden auch häufiger aufgerufen als Mädchen (Keller, 2001). Lehrkräfte trauen den Schülern mehr zu und wollen ihre Schülerinnen möglicherweise unbewusst oder bewusst „schonen“.

Bei der Einstellung der Lehrkräfte zur Begabung ihrer SchülerInnen scheint deren Annahme zur Entstehung von Geschlechterdifferenzen eine nicht unbedeutende

Rolle zu spielen. So zeigt sich, dass zum Beispiel umso eher stereotype Empfehlungen zum Besuch weiterführender Schulen von den Lehrkräften gegeben werden, je stärker diese glauben, dass Verhalten oder Charakter biologisch/genetisch bestimmt sind und nicht sozial erworben (Nürnberger et al., 2016). Damit einher geht auch die Einstellung, dass Hilfe durch die Lehrkraft diese biologisch/genetisch bedingten Geschlechterdifferenzen nicht wettmachen kann. Dies führt dazu, dass Lehrkräfte meinen, dass ihre Hilfe oder Nichthilfe mehr Einfluss auf den schulischen Erfolg/Misserfolg ihres schlechtesten Schülers als ihrer schlechtesten Schülerin hat (Fennema et al., 1990).

Die Einstellung der Lehrkräfte hat aber nicht nur Auswirkungen auf Häufigkeit des Lobens, sondern auch auf die Häufigkeit von Hilfestellungen oder auf schulische Weiterempfehlungen, die angenommen werden können oder auch nicht.

Stereotype Einstellungen der Lehrkräfte beeinflussen die Selbstwahrnehmung, Motivation und Leistung der SchülerInnen (Jacobs & Eccles, 1992). Dies geschieht unter anderem durch den bereits in Kapitel 3 dargestellten Effekt der self-fulfilling-prophecy. Die schulische Leistung hat nun aber wieder Einfluss auf die Entwicklung des Fähigkeitsselbstkonzeptes, wobei sich ein hohes Fähigkeitsselbstkonzept positiv auf die weitere Leistungsentwicklung auswirkt und umgekehrt (Niklas et al., 2012).

6 Auswirkungen

Die Leistungen von SchülerInnen können durch geschlechtsstereotype Einstellungen der Lehrpersonen beeinflusst werden. Zwei Wege der Beeinflussung werden im folgenden Kapitel näher ausgeführt. Der Einfluss des Fähigkeitsselbstkonzeptes und das Phänomen des *stereotype threat*.

6.1. Fähigkeitsselbstkonzept

„Vorstellungen über die Höhe eigener Fähigkeiten (Fähigkeitsselbstkonzept) sind ein zentraler Inhalt selbstbezogenen Wissens, weil Personen Fähigkeiten häufig als wichtige Eigenschaften des Selbst empfinden: Man ist in der Regel lieber fähig als unfähig“ (Dickhäuser, 2006, S. 5). Weiters bestimmen Fähigkeitsselbstkonzepte mit, ob sich eine Person für eine anspruchsvolle

Bildungs- und Berufslaufbahn befähigt sieht oder nicht und sie sind mit geschlechtertypischen Berufswahlpräferenzen verknüpft (Filipp, 2006).

Shavelson, Hubner und Stanton (1976) konzipierten das hierarchische Selbstkonzeptmodell, an dessen Spitze das allgemeine Selbstkonzept steht und auf der zweiten Ebene das akademische Selbstkonzept (Fähigkeitsselbstkonzept), mit fachspezifischen Facetten. Diese Fähigkeitsselbstkonzepte sind generalisierte Einschätzungen der Fähigkeiten, die SchülerInnen aufgrund von Erfahrungen ihrer Kompetenz in den einzelnen Schulfächern erwerben. Diese Erfahrungen sind geprägt durch die Urteile und Rückmeldungen signifikanter Bezugspersonen und durch Kausalattributionen (Shavelson et al., 1976). Im Bereich Schule sind dies vor allem Rückmeldungen durch die Lehrkräfte in Form von Noten (direkt) oder deren Verhalten (indirekt). Noten werden von den SchülerInnen als Vergleichsmöglichkeit mit den anderen oder zwischen den Fächern herangezogen und bilden so relevante Informationsquellen. Diese zwei Informationsquellen liegen dem bedeutenden Erklärungsansatz zur Genese akademischer Selbstkonzepte zugrunde, dem Internal/External Frame of Reference-Model (I/E-Model; Bezugsrahmenmodell) von Marsh (1986). Für den *external frame of reference* werden interindividuelle bzw. soziale Vergleiche angestellt, es wird also die eigene Leistung mit den Leistungen der MitschülerInnen verglichen. Daneben werden die eigenen Leistungen in einem (z.B. sprachlichen) Fach mit den eigenen Leistungen in einem (z.B. mathematischen) Fach verglichen (*internal frame of reference*), also intraindividuelle bzw. dimensionale Vergleiche angestellt (Marsh, 1986).

Es macht für die sozialen Vergleiche aber auch einen Unterschied, ob man zum Beispiel als starke Schülerin/starker Schüler in einer Klasse mit schwachen oder sehr vielen starken SchülerInnen sitzt. Zwei SchülerInnen gleicher Leistungsstärke können unterschiedliche Fähigkeitsselbstkonzepte entwickeln, je nach Leistungsstärke ihrer Klasse. Dieses Phänomen wird als „*big-fish-little-pond-Effekt*“ bezeichnet (Dickhäuser, 2006).

Auch die Attribuierungsprozesse haben einen Einfluss auf das Fähigkeitsselbstkonzept. Unter Attributionen versteht man subjektive Erklärungen oder Ursachenzuschreibungen für bestimmte Sachverhalte oder Ereignisse (Krapp, Geyer, & Lewalter, 2014). Das zweidimensionale Modell der Kausalattributionen nach Weiner (1986) enthält Unterscheidungen bezüglich der

Lokalität einer Ursache und der Stabilitätsannahme der Ursache. Tabelle 1 gibt einen Überblick über typische Kausalfaktoren in schulischen Lernsituationen.

Tabelle 1

Zweidimensionales Modell der Kausalattributionen nach Weiner (1986)

		Lokalität	
Stabilität	stabil	internal Fähigkeit/Begabung	external Aufgabenschwierigkeit
	variabel	Anstrengung	Zufall

Anmerkung. Nach Seidl et. al, 2014

„Wenn Mädchen in bestimmten Fächern gute Zensuren eher auf Anstrengung als auf Begabung zurückführen, kann dies zu niedrigeren Sk¹ eigener Fähigkeiten beitragen“ (Schilling, Sparfeldt & Rost, 2006, S.16). So haben zum Beispiel Mädchen von genderbezogen konservativ denkenden Eltern ein geringeres Fähigkeitsselbstkonzept bezogen auf das Fach Mathematik (Jöstl et al., 2015). „Mathe ist halt nichts für dich“, „Dafür hast du halt kein Talent“, etc. sind Beispiele für implizite Fähigkeitsurteile, die von den Kindern übernommen werden. Wird Misserfolg anhaltend auf stabile und als nicht veränderbar eingeschätzte Faktoren, wie mangelnde Fähigkeiten, attribuiert, führt dies schließlich zu Gefühlen der Hilflosigkeit und Resignation (Krapp et al., 2014). Schreibt das Kind aber mangelnde Anstrengung (internal variabel) als Ursache für schlechtes Abschneiden zu, hat es die Möglichkeit selbst etwas zu verändern. Der Zusammenhang zwischen der kindlichen Fähigkeitseinschätzung und der Einschätzung der Eltern wird immer stärker, je älter die Kinder sind (Wigfield et al., 1997). Während Buben ihre Leistungen gemessen an den Schulnoten im mathematisch-naturwissenschaftlichen Bereich dann eher überschätzen, unterschätzen Mädchen diese (Dresel et al., 2001).

Zusammengefasst, haben Mädchen also bei gleicher Leistung wie Buben in den mathematisch-naturwissenschaftlichen Fächern geringeres Vertrauen in die eigenen Fähigkeiten, weniger Engagement und größere Angst vor Mathematik

¹ Selbstkonzepten

(Jöstl et al., 2015). Diese Angst führt zu einer weiteren bereits vielfach erforschten Auswirkung von Geschlechterstereotypen – dem *stereotype threat* (Steele & Aronson, 1995).

6.2. Stereotype Threat

Wird StudienteilnehmerInnen mitgeteilt, dass ein Test bekannterweise für Frauen schwieriger zu bearbeiten ist als für Männer, schneiden Frauen daraufhin in diesem Test auch wirklich schlechter ab (O'Brien & Crandall, 2003). In einer Studie mit 312 SchülerInnen erhielten diese eine Physikaufgabe, die sie lösen mussten. Wurde zusätzlich angegeben, dass es bei diesem Test Geschlechterdifferenzen gibt, schnitten die Schülerinnen tatsächlich signifikant schlechter ab. Wurde hingegen explizit betont, dass sich bisher bei diesem Test keine Geschlechterdifferenzen gezeigt hätten, schnitten beide Geschlechter gleich gut ab (Marchand & Taasoobshirazi, 2013).

Eine Person, die das Gefühl hat, dass andere Personen zu der Gruppe, zu der sie sich zugehörig sieht, ein negatives Stereotyp haben, hat Angst, auch in dieser stereotypen Weise gesehen zu werden. Auch dann, wenn diese Person dieses Stereotyp gar nicht bestätigt (Spencer, Logel, & Davies, 2016). Es kommt also zur Bedrohung durch ein Stereotyp (*stereotype threat*). Dieses Phänomen geht auf Steele und Aronson zurück, die dieses erstmals in einer Studie zur Leistung afroamerikanischer Studierender beschrieben (Steele et al., 1995). Sie konzentrierten sich in ihrer Forschung auf Afroamerikaner, doch gab es in weiterer Folge zahlreiche Untersuchungen, die den Einfluss in anderen Leistungsbereichen wie Sport (Stone, Lynch, Sjomeling, & Darley, 1999) oder den Einfluss anderer Stereotype wie Alter (Abrams, Crisp, Marques, Fagg, Bedford & Provias, 2008) auf die Leistung zum Gegenstand ihrer Studien hatten. Manche Forscher nehmen an, dass, sobald ein Stereotyp durch *Priming* aktiviert wird, stereotypes Verhalten automatisch folgt. „Priming ist die (indirekte) Voraktivierung eines Gedächtnisinhalts durch einen Hinweisreiz“ (Herkner, 2001, S. 167). Als Beispiel nennt Herkner die Darbietung des Wortes „grün“. Dabei wird nicht nur der Grün-Knoten aktiviert, sondern auch, indirekt durch Erregungsausbreitung, angrenzende Kognitionen (z.B. die Assoziationen „blau“ und „Wiese“). Dar-Nimrod und Heine (2006) gaben Frauen verschiedene Artikel zu lesen, um *stereotype threat* Effekte nach Priming nachzuweisen. Frauen,

welche einen Artikel zum Thema „Geschlechterunterschiede in Mathematik sind genetisch bedingt“ oder generell zum Thema Gender zu lesen bekamen, schnitten schlechter ab als Frauen, die Artikel bekamen, in denen besagt wurde, dass keine Geschlechterunterschiede im mathematischen Bereich existieren.

Nur Individuen, deren soziale Identität durch das Stereotyp betroffen ist, sind durch stereotype threat gefährdet. Bei den anderen kann es sogar zu einer verbesserten Leistung, genannt „*stereotype lift*“ kommen (Spencer et al., 2016).

Wesentlich für die Aktivierung eines Stereotyps ist allerdings, dass man sich der stereotypisierten Gruppe zugehörig fühlt und sich mit ihr identifiziert. Je stärker sich eine Frau zur Gruppe „Frauen“ zugehörig fühlt, umso geringer identifiziert sie sich zum Beispiel mit Mathematik (Mathe = männlich, ich = weiblich, Mathe $\neq$ ich) und umso geringer ist ihre Präferenz für Mathematik. Je mehr sich eine Person aber wiederum mit Mathematik identifiziert (Mathe = ich), umso positiver sind ihre Einstellungen zu Mathematik. Dann wird sich keine Leistungsminderung einstellen (Nosek, Banaji, & Greenwald, 2002).

Beim Phänomen stereotype threat zeigt sich auch der Einfluss der geschlechterstereotypen Einstellung der Eltern als Moderator. Lehnt die Mutter geschlechterstereotype Einstellungen in Bezug auf Mathematik ab, verringerte sich die Mathematikleistung der Tochter in einer stereotype threat Bedingung nicht (Tomasetto, Alparone, & Cadinu, 2011). Fühlt man sich also der Gruppe „Frauen“ zugehörig, teilt aber die gängigen Stereotype zu Frauen und Mathematik nicht und hat sie damit auch nicht präsent, wird die Leistung nicht beeinflusst.

Im Folgenden soll noch ein Aspekt näher beleuchtet werden, der zu verminderter Leistung führen kann, nämlich gesteigerter Erfolgsdruck.

Gesteigerter Erfolgsdruck: Personen möchten gerne beweisen, dass die negativen Stereotype über ihre Gruppe falsch sind. Daher versuchen sie alles, um in Leistungssituationen möglichst gut abzuschneiden. Dies gelingt bei leichten Aufgaben durch die zusätzliche Motivation gut, nicht aber bei schwierigen (Spencer et al., 2016).

In einer Studie mit Frauen und Männern zu diesem Phänomen erhielt jeder Teilnehmer zwei mathematische Tests (leicht und schwierig). Die weiblichen

Teilnehmer, denen die Information gegeben wurde, dass bekannt ist, dass es Geschlechterdifferenzen bei diesem Test gibt, schnitten beim einfachen Test besser und beim schwierigen Test schlechter ab als Teilnehmerinnen ohne Information darüber. Männer beeinflusste die Information über Geschlechterdifferenzen nicht (O'Brien et al., 2003). Wieso es zu schlechterer Leistung bei schwierigen Aufgaben kommt, kann mit dem Prozessmodell von Schmader, Forbs und Johns (2008) erklärt werden. Für die Erbringung von Leistung spielt das Arbeitsgedächtnis eine Rolle. Dieses koordiniert Informationen und muss gleichzeitig die Interferenz von störenden Informationen verhindern. Versucht nun eine Person bei einer Aufgabe das Stereotyp nicht zu bestätigen, unternimmt sie Anstrengungen angstbezogene und stereotype Gedanken und Gefühle zu unterdrücken. Außerdem kommt es zu einer physiologischen Stressreaktion, die sich direkt auf die präfrontale Verarbeitung auswirkt. Des Weiteren versucht die Person, die eigene Leistung zu überwachen. All dies verbraucht Ressourcen des Arbeitsgedächtnisses, die dann für gute Leistungserbringung fehlen.

Wie in Kapitel 5 besprochen, haben Lehrkräfte großen Einfluss auf ihre SchülerInnen. Sie können ihre geschlechterstereotypen Einstellungen bewusst oder unbewusst vermitteln. Neben Aussagen können sie diese auch über die Bewertung und Benotung der Leistung ihrer SchülerInnen transportieren. Wie Benotung in der Schule erfolgt und wieso Lehrkräfte zu unterschiedlichen Noten kommen können, wird im folgenden Kapitel dargestellt.

7 Benotung/Beurteilung

Wie im Kapitel 6 bereits ausführlich dargestellt, werden Noten von den SchülerInnen als Vergleichsmöglichkeiten mit sich selbst in verschiedenen Fächern und mit anderen in gleichen und unterschiedlichen Fächern herangezogen. Aus diesen Vergleichen bilden sie ihr Fähigkeitsselbstkonzept. In Österreich gibt es ein fünfstufiges Notensystem, das im Grundschulbereich teilweise durch eine verbale Benotung ersetzt oder ergänzt und in der Neuen

Mittelschule ab der siebenten Schulstufe zu einem siebenstufigen System erweitert wird.

Der Leistung des Schülers/der Schülerin wird eine Zahl zugeordnet. Die Zahl sagt aber wenig über die tatsächliche Leistung aus (Holmeier, 2012). So schnitten österreichische Frauen beim Eignungstest schlechter ab als Männer, trotz gleicher Schulnoten (Spiel, Schober, & Litzenberger, 2008). Die Lehrkraft kann sich zur Bewertung der Leistung von SchülerInnen an verschiedenen Normen/Kriterien als Vergleich orientieren. Dabei lassen sich *soziale, kriteriale und individuelle Bezugsnormen*. Vergleicht die Lehrkraft die Leistung eines Schülers/einer Schülerin mit der Leistung der Klasse, spricht man von *sozialer* Bezugsnorm (Holmeier, 2012). Hierbei kommt wieder der „*big-fish-little-pond-Effekt*“ zum Tragen. Es kommt auch zu keiner eigentlichen Beurteilung der Leistung des/der Einzelnen, sondern nur zu einem Vergleich mit der Klasse. Bei einer Beurteilung nach der *kriterialen* Bezugsnorm wird die Leistung mit einem vorher festgesetzten Kriterium verglichen (Holmeier, 2012). Bei der dritten Möglichkeit, der *individuellen* Bezugsnorm, ist die vorherige Leistung der Schülerin/des Schülers ausschlaggebend. Hier hängt die Note nur davon ab, ob sich die Leistung verbessert oder verschlechtert hat, oder ob sie gleichgeblieben ist. Dieselbe Leistung kann also je nach Bezugsnorm ganz unterschiedlich beurteilt werden (Holmeier, 2012).

Dass sich Lehrkräfte in den Normen, die sie zur Bewertung von Schülerleistungen heranziehen unterscheiden (Fischbach, Brunner, Krauss, & Baumert, 2015, S. 3), kann erklären, wie es zu unterschiedlichen Noten bei tatsächlich gleicher Leistung kommen kann. Zusätzlich fließen aber auch noch andere, leistungsfremde Kriterien, in die Benotung mit ein (Ingenkamp & Lissmann, 2008).

8 Die Rolle von Erfahrung/Expertise

Für die Handlungskompetenz von Lehrkräften stellt die Einschätzung der Wichtigkeit von unterschiedlichen Einzelleistungen und auch von Unterrichtsereignissen (wie Mitarbeit) eine zentrale Bedeutung dar (Schäfer & Seidel, 2015). Posner (1988) beschreibt Experten als Personen, die dauerhaft

fachkundige Leistung in einer Domäne erbringen (zit. nach Dünnebier, Gräsel, & Krolak-Schwerdt, 2009). Für Expertise sind drei Kriterien nötig: Erfahrung mit mehrjähriger Praxis, Effizienz und bereichsspezifisches Wissen (Krems, 1994). Der Expertiseansatz wendet sich der professionsspezifischen inneren Struktur des Wissens von Lehrenden zu. Der Ansatz zeigt, dass Anfänger oder Laien nicht über die gleichen spezifischen Wissensschemata verfügen wie Experten (Seidel & Reiss, 2014). Bereits Lehramtsstudierende können im Unterricht eine Reihe relevanter Ereignisse identifizieren, verlieren sich aber schnell in Einzelheiten (Schäfer et al., 2015). Lehrerexperten nehmen im Gegensatz dazu Unterricht in einer komplexen Einheit wahr und markieren diese Einheit mental. Sie richten somit den Blick auf das Wesentliche, behalten den Überblick und können Vorhersagen über zukünftige Prozesse treffen (Seidel et al., 2014). Bei Anfängern basieren die Interpretationen häufig auf naiven Überzeugungen und sie nehmen schnell Übergeneralisationen vor (Seidel et al., 2014). Sie gehen verstärkt heuristisch vor und verlassen sich bei ihren Urteilen auf Stereotype (Dünnebier et al., 2009). Erfahrung hilft auch, bewährte Schemata als Rahmen zu verwenden, um komplexe Informationen über die SchülerInnen besser verstehen und einordnen zu können (Carter, Sabers, Cushing, Pinnegar, & Berliner, 1987).

Aber auch mit Erfahrung benötigt die Lehrkraft einen Referenzrahmen. Wenn kein Bezug auf soziale, individuelle oder kriteriale Norm möglich ist und kaum personenbezogene Merkmale vorhanden sind, dann könnte größerer Bedarf bestehen, sich auf Stereotype zu stützen. Ob Lehrkräfte sich bei der Beurteilung einer SchülerInnenleistung auf Geschlechterstereotype stützen, wenn kaum personenbezogene Informationen vorhanden sind, untersuchte Hofer (2015) in einer länderübergreifenden Studie im Schulfach Physik. Dabei zeigte sich auch gut der Einfluss von Erfahrung. So konnte sie Ergebnisse früherer Studien (u.a. Dünnebier et al., 2009; Krolak-Schwerdt, Böhmer, & Gräsel, 2012) replizieren, die aufzeigten, dass acht bis zehn Jahre Lehrerfahrung nötig sind, um Flexibilität in der Informationsverarbeitung zu erlangen.

Studie Hofer (2015)

Es existieren bereits zahlreiche Studien zum Einfluss von Geschlechterstereotypen im Fach Mathematik, aber nur wenige speziell im Fach

Physik, einem klassisch naturwissenschaftlichen Bereich. Hofer konzentrierte sich daher in ihrer Studie auf dieses Schulfach. Dazu wurde 780 Physiklehrkräften der Sekundarstufe in Österreich, Deutschland und der Schweiz eine Frage aus einem Physiktest und eine fiktive SchülerInnenantwort vorgegeben. Diese war für alle gleich. Manipuliert wurden lediglich zwei Faktoren: das SchülerInnengeschlecht und die Spezifikation auf sprachlichen versus naturwissenschaftlichen Bereich der Schülerin/des Schülers. Weitere Informationen zu den SchülerInnen waren nicht gegeben. Es zeigte sich, dass Physiklehrkräfte SchülerInnen signifikant schlechter beurteilten. Dieser Effekt verschwand mit zunehmender Lehrerfahrung (nach ungefähr zehn Jahren). Einzig bei männlichen Lehrkräften in Deutschland zeigte sich auch bei geringer Lehrerfahrung kein Unterschied in der Beurteilung. Eine Erklärung für diesen Effekt konnte nicht gefunden werden.

Nur das SchülerInnengeschlecht hatte einen Einfluss auf die Beurteilung. Die Spezifikation des Schülers/der Schülerin auf entweder den sprachlichen oder naturwissenschaftlichen Zweig zeigte keinen Effekt. Diese schien keine soziale Kategorie zu aktivieren, die im Beurteilungsprozess verwendet wird.

Aber nicht alle Lehrkräfte, gleichgültig ob mit oder ohne Lehrerfahrung, lassen sich in ihrer Beurteilung von Stereotypen leiten. Worin unterscheiden sich die, die stereotyp denken, von denen, die dies nicht, oder weniger, tun? Das nächste Kapitel befasst sich mit einem Persönlichkeitsmerkmal, das ein verstärktes Verlassen auf Stereotype bei der Urteilsbildung bedingen kann.

9 Need for Cognitive Closure

Bei dem Vorgehen, sich bei der Beurteilung von anderen vermehrt auf Stereotype zu stützen, könnte ein aus der sozialpsychologischen Forschung kommendes Konstrukt ebenfalls eine wesentliche Rolle spielen, nämlich *need for cognitive closure* (NFC). NFC beschreibt eine Dimension individueller Differenzen in Zusammenhang mit der Motivation bezüglich Informationsverarbeitung und Urteilsbildung (Webster & Kruglanski, 1994). Es bezeichnet das Bedürfnis nach kognitiver Geschlossenheit, also einem Streben nach Antworten, Struktur und

der Vermeidung von Ungewissheit und Ambiguität (Kruglanski A. W., 1990). Personen mit hohem Bedürfnis nach kognitiver Geschlossenheit (hohem NFC) akzeptieren als Antwort auf eine Frage jede beliebige Antwort, nur um Ungewissheit und Ambiguität zu vermeiden. Das Gegenteil von Bedürfnis nach kognitiver Geschlossenheit ist das Bedürfnis kognitive Geschlossenheit zu vermeiden. Diese beiden werden als Endpunkte eines Kontinuums gesehen (Webster et al., 1994).

Need for cognitive closure ist sowohl eine Persönlichkeitsvariable, als auch eine situationale Variable. Personen mit einem hohen Bedürfnis nach kognitiver Geschlossenheit ziehen generell weniger Informationen zur Urteilsbildung heran und vertrauen verstärkt ihren eigenen Urteilen. Individuelle Differenzen können durch verschiedenartige kulturelle Normen und Sozialisationspraktiken entstehen. So könnte ein Elternhaus, das viel Wert auf Einhaltung der Regeln, Ordnung und Struktur legt, zur Entwicklung eines größeren Bedürfnisses nach kognitiver Geschlossenheit beitragen (Webster & Kruglanski, 1997). Andererseits können Situationsvariablen, wie Zeitdruck oder Müdigkeit, das Bedürfnis nach kognitiver Geschlossenheit bei jeder Person erhöhen (Schlink & Walther, 2007). Auch in Situationen von hoher Komplexität können sich Personen schneller eine Antwort wünschen, als sie es in anderen Kontexten tun würden (Webster et al., 1994).

Personen mit hohem NFC verlassen sich mehr auf Stereotype, da Stereotype eine vorexistierende Wissensstruktur repräsentieren, die bereits fertig für den Gebrauch ist, während die Verarbeitung von vorliegenden Informationen weiterer aufwändigerer Prozesse bedarf (Kruglanski & Webster, 1996). Sie rufen auch mehr stereotypen-konforme Informationen ab, wenn sie eine Gruppe charakterisieren sollen und beurteilen die Gruppe stereotyper und homogener. Des Weiteren haben sie auch eine größere Präferenz für Informationen, die mit ihren stereotypen Erwartungen übereinstimmen (Dijksterhuis, van Knippenberg, Kruglanski, & Schaper, 1996). Sun, Zuo, Wu und Wen (2016) konnten in einer Studie zur Perspektivenübernahme zeigen, dass diese bei Personen mit hohem NFC zu einer verstärkten Verwendung von Stereotypen führt, während bei Personen mit niedrigem NFC die Stereotypenverwendung reduziert wurde. Bei Personen mit hohem NFC sind während der Perspektivenübernahme die

Stereotype der Zielgruppe salienter und sie beurteilen dann dementsprechend stereotyper.

Zusammenfassend zeigt sich also, dass geschlechterstereotype Beurteilung in der Schule unter anderem eine Folge von geringer Lehrerfahrung und/oder einem großen Bedürfnis nach kognitiver Geschlossenheit sein kann.

10 Fragestellungen und Hypothesen

Hofer (2015) hat unterschiedliche Bewertungen in der Beurteilung von Schülerinnen und Schülern aufgezeigt. Die Antworten von Schülerinnen auf eine Frage aus dem Bereich der Physik wurden signifikant schlechter beurteilt als von Schülern. Ausgehend davon versucht die vorliegende Arbeit zu klären, ob diese Beurteilungsunterschiede repliziert werden können, und ob das sozialpsychologische Konstrukt des Bedürfnisses nach kognitiver Geschlossenheit (need for cognitive closure) ebenfalls einen Einfluss auf die Beurteilung von SchülerInnen im Unterrichtsfach Physik hat.

Die Hauptforschungsfrage dieser Arbeit lautet:

Unterscheiden sich Physik-Lehramtsstudierende in der Benotung von Schülerinnen und Schülern bei einer Physikaufgabe?

Hofer befragte in ihrer Studie Physiklehrpersonen, die bereits unterrichten. Keine Unterschiede bei der Beurteilung von Schülerinnen und Schülern zeigten sich erst bei Lehrkräften ab einer Lehrerfahrung von ungefähr zehn Jahren. Im Gegensatz zu den fertig ausgebildeten Lehrkräften haben Studierende im Fach „Physik für Lehramt“ keine bis wenig Unterrichtserfahrung. Unterrichtserfahrung kann dadurch entstehen, dass in Österreich LehramtsstudentInnen bereits gegen Ende des Studiums unterrichten dürfen, da in bestimmten Fächern fertig ausgebildete Lehrkräfte fehlen.

Durch die geringe Lehrerfahrung sollte sich bei Physiklehramtsstudierenden ebenfalls eine unterschiedliche Beurteilung von Schülerinnen und Schülern

zeigen. Da sich in der Untersuchung von Hofer keine großen Unterschiede zwischen Österreich, der Schweiz und Deutschland zeigten, wurde die nachfolgende Studie nur in Österreich durchgeführt. Aus den Befunden von Hofer (2015) ergibt sich folgende Hypothese:

Hypothese 1: Physik-Lehramtsstudierende in Österreich beurteilen die Antwort auf eine Physikaufgabe von einer Schülerin schlechter als von einem Schüler.

Vorurteile sind Einstellungen, deren Objekte Außengruppen sind (Herkner, 2001). Die stereotype Vorstellung von mangelnder Begabung für Naturwissenschaften wird vermehrt Frauen zugeschrieben. Demnach wären männliche Lehramtsstudierende in der Außengruppe und müssten daher stärkere stereotype Einstellungen haben und Schülerinnen schlechter benoten als weibliche Lehramtsstudierende, wenn stereotype Einstellungen, wie von Hofer (2015) postuliert, eine Rolle spielen. Hofer (2015) fand allerdings keinen Unterschied in der Beurteilung zwischen männlichen und weiblichen Lehrkräften. Zusätzlich zeigten in ihrer Studie männliche Lehrkräfte in Deutschland sogar keinerlei stereotypen Effekt.

Die zweite Hypothese lautet daher wie folgt:

Hypothese 2: Männliche Physik-Lehramtsstudierende in Österreich zeigen signifikante Unterschiede in der Beurteilung der Antworten von Schülerinnen auf eine Physikaufgabe zu weiblichen Lehramtsstudierenden.

Lehramtsstudien in Österreich müssen immer in zumindest zwei Fächern absolviert werden. Die Wahl des Zweitfaches ist unabhängig vom Erstfach. Es können also mit einem naturwissenschaftlichen Erstfach auch geisteswissenschaftliche Fächer kombiniert werden. Zur Beantwortung der Frage, ob die Wahl des Zweitfaches die stereotype Sichtweise beeinflusst, gibt es keine Studien. Denkbar wäre, dass die Kombination von naturwissenschaftlichem mit geisteswissenschaftlichem Fach eine eventuell vorhandene stereotype Denkweise vermindert. Dies führt zur dritten Hypothese, die wie folgend lautet:

Hypothese 3: Lehramtsstudierende in Österreich, deren Zweitfach kein naturwissenschaftliches ist, zeigen signifikante Unterschiede in der Beurteilung der Antwort von Schülerinnen auf eine Physikaufgabe zu Lehramtsstudierenden, deren Zweitfach ebenfalls ein naturwissenschaftliches ist.

Um die Aktivierung eines eventuell vorhandenen Geschlechtsstereotypes zu verstärken, manipulierte Hofer (2015) auch die Schwerpunktwahl der SchülerInnen. Es zeigte sich bei ihrer Studie wider Erwarten aber kein signifikanter Effekt durch die Vorgabe sprachlicher versus naturwissenschaftlicher Schwerpunkte. Denkbar wäre aber, dass Schülerinnen, die ihren Schwerpunkt auf Sprachen legen, für weniger talentiert im naturwissenschaftlichen Bereich gesehen werden. Daher soll in dieser Arbeit nochmals der Einfluss der Schwerpunktsetzung überprüft werden:

Hypothese 4: Es zeigen sich Unterschiede in der Beurteilung der Antwort auf eine Physikfrage von SchülerInnen mit Schwerpunktsetzung im sprachlichen Bereich zu der von SchülerInnen mit naturwissenschaftlicher Schwerpunktsetzung durch Physik-Lehramtsstudierende.

Als Erweiterung zur Studie von Hofer (2015) soll auch der Einfluss eines Persönlichkeitsmerkmals überprüft werden. Personen mit hohem Bedürfnis nach kognitiver Geschlossenheit (NFC) haben die Tendenz sich bei mangelnder Information auf Stereotype zu verlassen.

Hypothese 5: Es gibt einen Zusammenhang zwischen dem Wert in der Skala NFC von Physik-Lehramtsstudierenden und der Beurteilung der Antwort von Schülerinnen auf eine Physikfrage.

III Empirischer Teil

11 Methode

Der folgende Teil der Arbeit beschreibt die Durchführung der Untersuchung, das verwendete Erhebungsinstrument und die Stichprobe.

11.1 Untersuchungsdesign und Vorgehen

Da es sich um eine Replikationsstudie handelt, war geplant, den Fragebogen von Hofer zu adaptieren, um die Fragestellungen zu testen, und diesen von ungefähr 80 bis 100 Lehramtsstudierenden des Unterrichtsfaches Physik ausfüllen zu lassen. Teilnehmen konnten Physik-Lehramtsstudierende aller Fakultäten in Österreich, die das Lehramtsstudium Physik anbieten, wobei alle drei derzeit angebotenen Möglichkeiten des Studiums (Diplom, Bachelor, Master) berücksichtigt waren.

Start der Erhebung sollte noch vor den Prüfungswochen am Ende des Sommersemesters sein und sie sollte bis in den späten Herbst hinein dauern.

Die Studie wurde von Anfang Juni bis November 2016 durchgeführt. Wie auch in der Ursprungsstudie wurde ein Online-Fragebogen mit dem Softwarepaket SoSciSurvey² verwendet. Die Rekrutierung der Teilnehmer erfolgte über soziale Medien (Facebook), Studienrichtungsvertretungen und die Studienprogrammleitungen der Fakultäten in Wien, Graz, Linz, Innsbruck und Salzburg. Sowohl die Studienrichtungsvertretungen als auch die Studienprogrammleitungen erhielten ein Email mit der Bitte, den link für den Fragebogen an die Studierenden des Faches Lehramt-Physik weiterzuleiten. Zusätzlich bekamen sie die Information, dass die Studie im Rahmen einer Masterarbeit im Fach Psychologie am Institut für Bildungspsychologie an der Universität Wien durchgeführt wird und die Angaben der Studierenden anonym behandelt werden. Weder die Erstadressaten noch die Teilnehmer selbst sollten über den eigentlichen Gegenstand der Untersuchung informiert werden, um -

² www.soscisurvey.de

nach Hofer (2015) - keine sozial erwünschten Antworten entgegen eventuell vorhandener geschlechtsstereotypen Einstellungen zu provozieren. Um trotzdem Informationen zum Grund der Untersuchung erhalten zu können, sollten die Teilnehmer die Möglichkeit bekommen, diese nach Abschluss der Studie per Email zu erhalten.

Als Studieninteresse wurde daher wie bei Hofer die Frage genannt, wie eine isolierte Bewertung einzelner Prüfungsantworten von Schülerinnen und Schülern durch verschiedene Experten mit einer Bewertung der vollständigen Prüfung einer Schülerin bzw. eines Schülers durch einen einzigen Experten zusammenhängt. Tatsächlich erhielten aber alle Teilnehmer dann den gleichen Fragebogen mit nur einer Prüfungsfrage zur Bewertung. Da sich Lehramtsstudierende wahrscheinlich nicht als Experten fühlen würden, wurde als Grund für ihre Mithilfe die Überprüfung des Einflusses von Erfahrung angegeben.

11.2 Erhebungsinstrument

Der Fragebogen wurde von Hofer zur Verfügung gestellt, an die veränderte Stichprobe adaptiert und erweitert. Er bestand aus drei Teilen.

Im ersten Teil wurden soziodemographische Daten erfragt. Diese umfassten neben dem *Geschlecht* und *Alter* der Person vor allem Angaben zum Studium. Dies waren *Studienart* (Diplom, Bachelor, Master), *Studiendauer* in Semester, *gewähltes Zweifach* und *Lehrerfahrung* und/oder *Nachhilfeerfahrung* mit Länge der Erfahrung in Semester.

Der zweite Teil des Fragebogens, in dem es um die Beurteilung einer Antwort auf eine Physikfrage ging, wurde ident wie von Hofer gestaltet. Hier fand eine automatische randomisierte Zuordnung zu einer der vier Untersuchungsbedingungen statt. In allen Bedingungen gab es exakt die gleiche Information, die gleiche Frage und die gleiche Antwort. Sie unterschieden sich lediglich durch das fiktive *SchülerInnengeschlecht* (männlich/weiblich) und die fiktive *Schwerpunktwahl* (naturwissenschaftlich/sprachlich). Am Anfang gab es eine kurze Information darüber, dass die Fragen und Antworten von teilnehmenden Schulen zur Verfügung gestellt wurden und der abgeprüfte Stoff im Unterricht behandelt worden war. Die Prüfungsfrage lautet wie folgt: *Zwei*

Skateboard-Fahrer mit deutlich unterschiedlicher Masse stehen sich je auf einem Skateboard gegenüber und sind mit einem gespannten Seil verbunden. Der linke und leichtere Skater zieht aktiv am Seil, der schwerere rechte Skater hält es nur fest. Was passiert? Begründe deine Vermutung in ca. fünf bis sechs Sätzen. Reibung ist zu vernachlässigen. Sie stammt aus dem Bereich der Newtonschen Mechanik und behandelte damit ein gängiges Stoffgebiet, das in der Unterstufe bearbeitet wird.

Es folgte die Antwort des Schülers/der Schülerin: *Generell gilt, die Kraft setzt sich zusammen aus der Masse einer Person und der Beschleunigung. Der rechte Skater muss das Seil genauso fest halten, wie der linke Skater am Seil zieht. Auf beide wirkt deshalb eine gleich grosse Kraft, auch wenn nur der linke Skater am Seil zieht. Folglich sollte nichts passieren, da sich die beiden Kräfte aufheben. Da die Masse des linken Skaters allerdings kleiner ist als die Masse des rechten Skaters, ergibt sich für den linken Skater eine grössere Beschleunigung als für den rechten Skater. Daher sollte sich vermutlich doch zumindest der linke Skater ein wenig in Richtung des rechten Skaters bewegen.*

Danach sollte die Antwort mit einer Note nach österreichischem Schulsystem bewertet werden. Dazu gab es einen Schieberegler von 1 („sehr gut“) bis 5 („nicht genügend“), wobei Abstufungen mit einer Dezimalstelle möglich waren. Dies diente einerseits dazu, eine größere Bandbreite an Noten für die Auswertung zu bekommen und andererseits auch um den Beurteilern einen größeren Spielraum in der Bewertung zu geben.

Die Prüfungsfrage wurde von Hofer aus dem Test of Mechanics Conceptual Understanding (bMCU; Hofer, Schumacher, & Rubin, 2015 in Hofer, 2015) entnommen. Dieser ist ein Rasch-skaliertes multiple-choice Test für das konzeptuelle Verständnis der Newtonschen Mechanik, wobei die Skateboarder-Frage ein Problem behandelt, das häufig in Physiklehrbüchern und Tests vorkommt (Hofer, 2015).

Die Antwort des Schülers/der Schülerin wurde im Zuge der Entwicklung des bMCU designt. Sie sollte nicht komplett falsch, aber auch nicht komplett richtig sein, um einen großen Spielraum für Interpretationen der Leistung des Schülers/der Schülerin zu geben. Sie war eine von drei designten Antworten, die

fünf informierten Physiklehrkräften zur Beurteilung gegeben wurde und repräsentierte die durchschnittlichste Leistung (Hofer, 2015).

Zur Absicherung, dass dieses Beispiel und die Frage auch für Physik-Lehramtsstudierende verständlich und die Antwort als weder richtig noch falsch interpretiert wird, wurden Frage und Antwort nochmals zur Beurteilung einer Physiklehrkraft, die eben erst ihr Studium abgeschlossen hatte, vorgelegt. Ebenso wurde ein Pretest von drei (Lehr-)Personen durchgeführt.

Der dritte Teil des Fragebogens diente der Erhebung des Bedürfnisses nach kognitiver Geschlossenheit. Dafür wurde die „Deutsche Kurzskala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit“ (Schlink et al., 2007) verwendet. Diese umfasst 16 Items im 6-stufigen Antwortformat, wobei die Skalenwerte wie folgt benannt wurden: 1 = „stimme gar nicht zu“ bis 6 = „stimme völlig zu“. Mit einer zu erwarteten Bearbeitungsdauer von ungefähr acht Minuten ist dieses Instrument als ökonomisch anzusehen. Die Kurzskala ist eine Entwicklung aus der ursprünglichen „Need for Cognitive Closure Scale“ (Webster et al., 1994), die aber mit 42 Items wesentlich umfangreicher und daher bearbeitungsintensiver ist. In Validierungsstudien erzielte die Kurzskala mit einem Cronbach $\alpha = .78$ eine zufriedenstellende interne Konsistenz.

Beispielitems:

Ich mag es nicht, wenn die Aussage einer Person mehrdeutig ist.

Im Allgemeinen suche ich nicht nach Alternativlösungen für Probleme, für welche ich schon eine Lösung parat habe.

Nach Beendigung der Befragung wurde noch angeboten, eine Emailadresse anzugeben, um die wichtigsten Ergebnisse zugesandt zu bekommen. Diese wurde getrennt vom restlichen Fragebogen gespeichert, sodass die Anonymität nicht verletzt wurde. Auf diesen Umstand wurden die Teilnehmer ausdrücklich hingewiesen. Die Bearbeitung des Fragebogens nahm gesamt 10 bis 15 Minuten in Anspruch.

11.3 Stichprobenbeschreibung

Insgesamt nahmen an der Befragung nur 34 Personen teil (wobei der Fragebogen 172mal angeklickt wurde), wovon 2 nach der ersten Seite abbrachen

und daher ausgeschlossen wurden. Nach Durchsicht der Daten mussten keine weiteren Personen ausgeschlossen werden. Von den 32, die vollständig ausfüllten, waren 12 Frauen (37,5%) und 20 Männer (62,5%). Das Alter betrug durchschnittlich 27,41 Jahre, wobei der Großteil der Teilnehmer zwischen 21 und 27 Jahren alt war. Die Altersspanne betrug 19 bis 55 Jahre.

25 Personen (78%) befanden sich noch im Diplomstudium, 6 im Bachelorstudiengang (19%) und 1 im Masterstudiengang (3%). Die durchschnittliche Studiendauer bisher betrug 7,9 Semester ($SD = 2.98$).

Lehrerfahrung und/oder Nachhilfeerfahrung zu haben, gaben 23 Personen an (71,9%), wobei 22 Befragte über Nachhilfeerfahrung und 14 über Lehrerfahrung verfügten. Die durchschnittliche Semesteranzahl betrug hierbei 1,28 Semester ($SD = 0.46$).

Nur 8 Personen (25%) gaben an, ein nicht naturwissenschaftliches Fach als Zweitstudium gewählt zu haben. Bei 17 war das Zweitstudium das Unterrichtsfach Mathematik (53%).

10 studierten an der Universität Wien, 8 an der Universität Innsbruck, 8 an der Universität Salzburg, 1 an der Universität Graz. Von der Universität Linz schien niemand teilgenommen zu haben. Die Variable, an welcher Universität studiert wird, wurde aber erst nach 5 Teilnehmern hinzugefügt. Dies erschien notwendig, da nicht von allen Studienprogrammleitern primär Rückmeldungen kamen, um gezielt die Bitte um Weiterleitung an die Studierenden wiederholen zu können. Die ersten 5 Teilnehmer füllten am gleichen oder vorigen Tag aus, wie die Teilnehmer der Universität Salzburg. Trotzdem ist nicht eindeutig ersichtlich, von welcher Universität sie stammen. Daher wurde bei ihnen der fehlende Wert belassen.

Von der Studienprogrammleitung der Universität Linz kam auch auf Nachfrage leider keine Rückmeldung. Es ist daher nicht nachvollziehbar, ob der Fragebogenlink weitergeleitet wurde. Der Studienprogrammleiter der Universität Wien gab auch nach telefonischer Rücksprache sein Einverständnis zur Durchführung der Studie an seinem Institut nicht und untersagte die Durchführung. Daher konnte in Wien auch kein Aushang des Links oder eine zusätzliche paper-pencil Befragung stattfinden. Die Teilnehmer aus Wien waren vor der Absage über die Studienrichtungsvertretung erreicht worden.

12 Ergebnisse

Die deskriptiv- und inferenzstatistische Auswertung erfolgte mit der Statistiksoftware IBM SPSS® Statistics 23. Die Hypothesen wurden mit einfachen Varianzanalysen, t-Test für unabhängige Stichproben und Regression getestet. Zusätzlich wurde im Anschluss noch eine schrittweise Regression gerechnet. Aufgrund der geringen Stichprobengröße konnte die vierte Hypothese, *„Die Antwort auf eine Physikfrage von Schülerinnen mit Schwerpunktsetzung im sprachlichen Bereich wird schlechter beurteilt als von Schülerinnen mit naturwissenschaftlicher Schwerpunktsetzung“*, nicht statistisch ausgewertet werden. Es wären in jeder Zelle nur 8 Fälle vorhanden, was eine Berechnung sinnlos macht.

Bei Stichprobenumfängen mit $n > 30$ kann von einer annähernden Normalverteilung der Daten ausgegangen werden (Bortz & Döring, 2006). Da mit $n = 32$ der Stichprobenumfang aber nur gering größer ist, wurde die Normalverteilung der Daten trotzdem mittels Kolmogorov-Smirnov-Test überprüft. Die Variabel „Note gesamt“ weicht nicht von Normalverteilung der Daten ab mit $p = .152$, ebenso die Variabel „Note Schülerinnen“ mit $p = .200$.

Als Signifikanzniveau der Ergebnisse wurde eine Irrtumswahrscheinlichkeit von $\alpha = .05$ gewählt. Zusätzlich wurde bei signifikanten Ergebnissen noch die Effektstärke aufgezeigt.

12.1 stereotype Beurteilung

Hypothese 1: Physik-Lehramtsstudierende in Österreich beurteilen die Antwort auf eine Physikaufgabe von einer Schülerin schlechter als von einem Schüler.

Die Lehramtsstudierenden vergaben auf die Antwort einer Schülerin oder eines Schülers auf eine Physikfrage eine Note von 1,0 bis 5,0. Diese wurde automatisch umgerechnet in eine Zahl von 1 bis 41 (Note = $0,1 \times$ vergebene Zahl + 0,9). Da es vier Versuchsbedingungen gab (Geschlecht m/w, Schwerpunkt sprachlich/naturwissenschaftlich), wurde eine eigene Variable „SchülerInnengeschlecht“ (1 = weiblich, 2 = männlich) erstellt. Danach wurde mit den Variablen „Note“ und „SchülerInnengeschlecht“ eine einfache

Varianzanalyse gerechnet. Die Voraussetzungen der Homogenität der Varianzen, berechnet mit Levene-Test, war mit $F(1, 30) = .982, p = .330$ gegeben. Es zeigte sich ein signifikanter Unterschied in der Benotung von Schülerinnen und Schülern, $F(1,30) = 5.743, p = .012$ (einseitig getestet). Die ermittelte Effektstärke liegt bei $\eta^2 = .161$, was als großer Effekt eingestuft werden kann. Sieht man sich die Mittelwerte an (weiblich: $M = 13.93$ mit $SD = 11.85$; männlich: $M = 23.23$ mit $SD = 10.11$) so erkennt man aber, dass entgegen der Erwartung, die Schüler schlechter beurteilt wurden als die Schülerinnen. Zum besseren Überblick über die Werte siehe Tabelle 2.

Tabelle 2

Notenverteilung nach SchülerInnenengeschlecht

Geschlecht	<i>N</i>	<i>M</i>	<i>SD</i>	<i>min</i>	<i>max</i>
Männlich	17	23.2353	10.10879	6	41
Weiblich	15	19.9333	11.85307	1	36

12.2 Zusammenhang mit dem Geschlecht der „Lehrkraft“

Hypothese 2: Männliche Physik-Lehramtsstudierende in Österreich zeigen signifikante Unterschiede in der Beurteilung der Antworten von Schülerinnen auf eine Physikaufgabe zu weiblichen Lehramtsstudierenden.

Zur Überprüfung dieser Hypothese wurde wieder eine einfache Varianzanalyse berechnet. Diesmal mit dem „Geschlecht der Lehrkraft“ und der Variable „Note Schülerin“. Die Voraussetzungen der Homogenität der Varianzen war mit $F(1,13) = 0.141, p = .713$ gegeben.

Es zeigte sich kein signifikanter Zusammenhang zwischen dem Geschlecht der/des Lehramtsstudierenden und der Note, die einer Schülerin gegeben wurde: $F(1,13) = 1.752, p = .208$.

Allerdings muss beachtet werden, dass die Stichprobe sehr klein ist. Nur 15 Schülerinnen wurden benotet von 7 weiblichen und 8 männlichen Lehramtsstudierenden.

12.3 Zusammenhang mit dem gewählten Zweitfach

Hypothese 3: Lehramtsstudierende in Österreich, deren Zweitfach kein naturwissenschaftliches ist, zeigen signifikante Unterschiede in der Beurteilung der Antwort von Schülerinnen auf eine Physikaufgabe zu Lehramtsstudierenden, deren Zweitfach ebenfalls ein naturwissenschaftliches ist.

Zur Beantwortung dieser Fragestellung wurde zuerst eine neue Variable „geisteswissenschaftliches Zweitfach gewählt“ gebildet. Dazu zählen die Fächer „Deutsch“, „Fremdsprachen“, „Sport“, „Musik“, „Kunst“, „Geschichte“ und „Religion“, wobei nur die Fächer „Sport“ und „Geschichte“ angegeben wurden. Danach wurde ein t-Test für unabhängige Stichproben mit den Variablen „geisteswissenschaftliches Zweitfach“ und „Note Schülerinnen“ gerechnet. Dabei wurde das Verfahren des Bootstrapping, mit jeweils 5000 gezogenen Stichproben, angewendet. Bei diesem werden wiederholte Samples aus dem vorhandenen Datensets gezogen und so Konfidenzintervalle und Signifikanztests berechnet. Dies erschien sinnvoll, da das Datenset sehr klein ist und so nicht von einer Normalverteilung der Residuen ausgegangen werden kann (Field, 2009).

Im t-Test für unabhängige Stichproben ergab sich ein nicht signifikanter Effekt bezüglich des gewählten Zweitfaches, $t(13) = 1.475$, $p = .164$. Die Voraussetzung der Homogenität der Varianzen war mit $F = 0.390$, $p = .537$ gegeben.

Es zeigt sich also kein Unterschied in der Benotung abhängig vom gewählten Zweitfach.

Zur Veranschaulichung der genauen Werte siehe Tabelle 3.

Tabelle 3

Schülerinnennote nach gewähltem Zweitfach

Zweitfach	N	M	SD	KI untere Grenze ^a	KI obere Grenze ^a
geisteswiss.	5	7.8	9.20	1.00	16.25
naturwiss.	10	17	12.30	9.78	24.25

^a Konfidenzintervall unter Bootstrapping

12.4 Zusammenhang mit Schwerpunktsetzung der Schülerin/des Schülers

Hypothese 4: Es zeigen sich Unterschiede in der Beurteilung der Antwort auf eine Physikfrage von SchülerInnen mit Schwerpunktsetzung im sprachlichen Bereich zu der von SchülerInnen mit naturwissenschaftlicher Schwerpunktsetzung durch Physik-Lehramtsstudierende.

Aufgrund der geringen Teilnehmerzahl konnte diese Frage inferenzstatistisch nicht überprüft werden.

12.5 Need for cognitive closure

Hypothese 5: Es gibt einen Zusammenhang zwischen dem Wert in der Skala NFC von Physik-Lehramtsstudierenden und der Beurteilung der Antwort von Schülerinnen auf eine Physikfrage.

Aus den 16 Items der „Deutsche Kurzskala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit“ (Schlink et al., 2007) wurde eine neue Variable „NCC gesamt“ gebildet. Dazu wurden pro Versuchsperson die Werte auf jede Frage addiert. Vorher mussten noch die Items 4, 11 und 12 umgepolt werden.

Danach wurde für die neue Variable die Reliabilität berechnet. Die interne Konsistenz war mit einem Cronbach $\alpha = .739$ zufriedenstellend.

Die Spanne der NCC Gesamtwerte reichte von 16 bis 49. Da der Wert maximal 96 annehmen kann, bewegten sich alle Versuchspersonen in der unteren Hälfte des möglichen Gesamtwertes. Es wurde aus diesem Grund keine Analyse mit Mediansplit vorgenommen.

Zur inferenzstatistischen Analyse des Zusammenhanges zwischen dem Bedürfnis nach kognitiver Geschlossenheit und vergebener Note bei Schülerinnen wurde eine lineare Regressionsanalyse durchgeführt. Hierbei wurde wiederum das Verfahren des Bootstrapping angewendet.

Die Daten wichen nicht von einer Normalverteilung ab, mit einem $p = .200$. Ebenso waren die Voraussetzungen der Linearität und der Normalverteilung der Residuen gegeben.

Ein Regressionseffekt konnte nicht gefunden werden, $F(1,13) = 2.553$, $p = .134$, mit einem R^2 von .164. NCC konnte nicht als Prädiktor für die Schülerinnennote identifiziert werden, $b = -.405$, $t(14) = -1.598$, $p = .134$, $KI^3 [-1.466; 70.607]$.

Zur Veranschaulichung der Regressionsgeraden und der Datenpunkte siehe Abbildung 1.

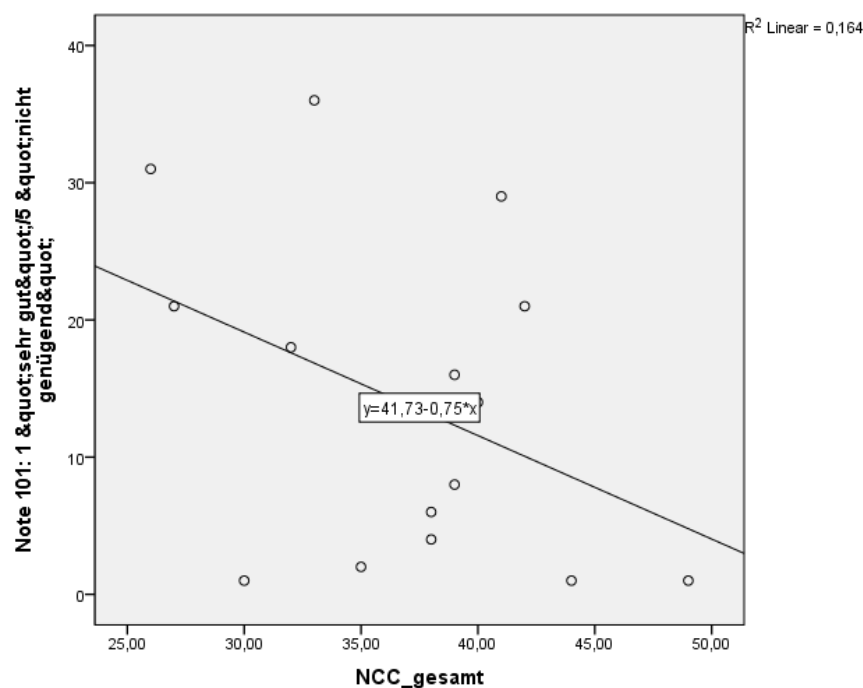


Abbildung 1 Regressionsgerade Schülerinnennote und NCC Studierende

12.6 Mögliche weitere Einflussfaktoren

Um mögliche weitere Einflussgrößen auf die Benotung von SchülerInnen zu identifizieren wurde im Anschluss noch eine schrittweise Regressionsanalyse

³ Unter Bootstrapping

durchgeführt. Dazu wurden folgende unabhängige Variablen schrittweise überprüft:

SchülerInnengeschlecht, geisteswissenschaftliches Zweifach, Geschlecht der „Lehrkraft“, vorhandene Erfahrung und Alter der „Lehrkraft“.

Es zeigte sich auch hier, dass nur das SchülerInnengeschlecht einen Einfluss auf die Benotung hat. $F(1,30) = 5.743$, $p = .023$ mit einem $R^2 = .161$.

Die Korrelationen und Signifikanzen aller getesteter Variablen sind in Tabelle 4 dargestellt.

Allerdings war der Stichprobenumfang mit $n = 32$ wiederum sehr gering.

Tabelle 4

Ergebnisse der schrittweisen Regression

Variable	Korrelation	Signifikanz	N
SchülerInnengeschlecht	.401	.011*	32
geisteswiss. Zweifach	-.243	.090	32
Geschlecht StudentIn	.237	.096	32
Erfahrung	-.017	.463	32
Alter	.240	.093	32

* $p < .05$

IV Diskussion

Ziel dieser Studie war es, mögliche Unterschiede in der Beurteilung von Schülerinnen und Schülern zu untersuchen. Da in der Literatur bisher hauptsächlich Ergebnisse zu Erhebungen im Bereich Mathematik und Frauen zu finden sind, lag der Fokus auf einem naturwissenschaftlichen Fach. Gerade zu Physik gab es bisher kaum Studien. Ausgangspunkt war die Studie von Hofer (2015), die eine schlechtere Beurteilung von Schülerinnen aufzeigte. Da in ihrer Studie Lehrerfahrung als eine wesentliche Einflussgröße identifiziert werden konnte, wurde diesmal die Erhebung mit Lehramtsstudierenden, die noch über wenig bis keine Unterrichtserfahrung verfügen, durchgeführt. Geschlecht der Lehramtsstudierenden, gewähltes Zweitfach und Schwerpunktsetzung der SchülerInnen waren weitere untersuchte Einflussgrößen. Des Weiteren sollte der Einfluss des sozialpsychologischen Konstruktes „Bedürfnis nach kognitiver Geschlossenheit“ (need for cognitive closure) auf eine möglicherweise unterschiedliche Beurteilung von Schülerinnen und Schülern im Unterrichtsfach Physik erhoben werden.

13 Zusammenfassung und Diskussion der Ergebnisse

Keine der aufgestellten Hypothesen konnte bestätigt werden. Die Stichprobengröße war allerdings sehr gering, wobei mögliche Gründe dafür in den Limitationen der Studie noch diskutiert werden.

13.1 Unterschiedliche Beurteilung von Schülerinnen und Schülern

Entgegen der Erwartungen nach der Studie von Hofer (2051) wurden in dieser Studie Schüler ($M = 23.23$, $SD = 10.11$) signifikant schlechter beurteilt als SchülerInnen ($M = 13.93$, $SD = 11.85$). Die durchschnittlichen Noten, die vergeben wurden, waren bei Schülerinnen 2.4, bei Schülern 3.

In der Literatur finden sich zahlreiche Befunde dafür, dass sich Lehrkräfte bei mangelnder Information oder unter Zeitdruck bei der Beurteilung auf Stereotype verlassen (u.a. Fennema et al., 1990; Dresel et al., 2001; Dünnebier et al., 2009). Besonders stark nutzen Novizen, also Anfänger oder Studierende, stereotype Erwartungen für die Benotung von SchülerInnen.

Eine mögliche Ursache für die unterschiedliche Beurteilung könnte demnach sein, dass bei den Teilnehmern der Erhebung Stereotype aktiviert wurden. Hofer (2015) nahm als Grund für die schlechtere Beurteilung der Schülerinnen in ihrer Studie an, dass bei den Lehrkräften das Stereotyp, dass Mädchen weniger begabt im mathematisch-naturwissenschaftlichen Bereich sind, aktiviert wurde. Neben diesem Stereotyp existieren aber noch andere Stereotype zu SchülerInnen. Mädchen werden häufig als fleißiger und sorgfältiger beschrieben als Buben (Spiel, Schober, & Finsterwald, 2011). So konnten Kuhl und Hannover (2012) zeigen, dass Mädchen bei gleicher Leistung bessere Noten erhalten als Buben. Dies zeigt sich auch in der bereits früher erwähnten Evaluation des Aufnahmetest zum Medizinstudium in Österreich (Spiel et al, 2008). Frauen erzielten schlechtere Testergebnisse als Männer, obwohl ihre Schulnoten gleich waren. Die Grundlagen für die Bewertung von Mädchen und Buben in der Schule scheinen demnach nicht ident zu sein. Sehr häufig werden Buben im Zusammenhang mit Schule als problematisch in Bezug auf Verhalten und Leistung gesehen und dargestellt. Jones und Myhill (2004) führten Interviews mit Lehrkräften an englischen Grund- und Sekundarschulen (Schulstufe 1-13) durch. Dabei zeigte sich, dass Lehrkräfte Buben generell als Schüler sehen, die „nichts können, nichts wollen und nichts tun“. Daran scheint sich auch in den letzten Jahren wenig geändert zu haben. 'I have lovely girls in my class, they are fantastic [pause] and my boys are okay too' (Hamilton & Jones, 2016; S. 241). Von der Lehrkraft, von der diese Aussage stammt, wurden in einer beobachteten Unterrichtseinheit 16 positive Kommentare zu Schülerinnen, aber nur 3 zu Schülern gegeben. Auch ADHS wird von Lehrkräften viel häufiger mit Buben in Verbindung gebracht (Hamilton et al., 2016).

Möglicherweise flossen also verstärkt negative Stereotype zu Schülern in die Beurteilung ein und überlagerten so das Stereotyp der „naturwissenschaftlich minderbegabten Schülerin“.

Eine weitere mögliche Erklärung zum entgegengesetzten Ergebnis zur Studie von Hofer (2015) könnte in der Ausbildung der Lehramtsstudierenden zu finden sein. Betrachtet man die Curricula zum Lehramtsstudium Physik, zeigt sich, dass die Genderproblematik bereits Thema ist. So finden sich im Curriculum der Universität Salzburg folgende Rahmenkompetenzen für die Bildungswissenschaften und die Schulpraxis: „Unterricht unter Berücksichtigung

der Heterogenität der Schüler/innen (z.B. Herkunft, Leistungstand, Geschlecht, Interessen und der gesellschaftlichen Herausforderungen (...) unter Anleitung planen und durchführen, reflektieren und evaluieren“. Unter Vernetzungskompetenzen findet sich: „die theoretischen Grundlagen und methodischen Konzepte der Gender Studies und der Relevanz für die Fächer erkennen und schulische Interaktionsprozesse danach ausrichten“ (Salzburg, 2013). Nürnberger et al. (2016) zeigten, dass stereotype Empfehlungen von Lehrkräften umso eher gegeben werden, je stärker diese an genetisch bedingte Begabungsunterschiede der Geschlechter glauben. Wenn es nun in der Ausbildung zur Aufklärung der zukünftigen Lehrkräfte über Ursachen und Entstehung von Geschlechtsunterschieden kommt, könnte dies zu einer Einstellungsänderung führen. Demnach könnte damit die genderstereotype Einstellung zu Mädchen und Naturwissenschaften vermindert worden sein. Da stereotype Einstellungen aber ebenso wie in der Studie von Hofer (2015) nicht erhoben wurden, könnten auch noch andere Gründe für die Beurteilungsunterschiede vorliegen. Es wäre daher sinnvoll in einer weiterführenden Studie auch geschlechtsstereotype Einstellungen mitzuerheben.

13.2 Weitere Zusammenhänge

Weder der Zusammenhang der Beurteilung mit dem Geschlecht der Lehramtsstudierenden, deren gewähltem Zweitfach, ihrem Alter, bereits gewonnener Erfahrung, noch mit der Schwerpunktsetzung der SchülerInnen zeigten einen signifikanten Effekt.

Auch bei Hofer (2015) zeigte sich kein Effekt bei der Schwerpunktsetzung der SchülerInnen und bei dem Geschlecht der Lehrkraft. Der Hinweis zur Schwerpunktsetzung dürfte demnach, wie von Hofer angenommen, nicht ausreichen, das eventuell vorhandene Stereotyp „Mädchen sind weniger talentiert in Physik als Buben“ zu aktivieren.

Carrington, Tymms und Merrell (2008) fanden ebenfalls keinen Zusammenhang des Geschlechts der Lehrkräfte mit der Leistung ihrer SchülerInnen. Zum Einfluss des Geschlechts auf die Leistung liegen aber generell wenige Studienergebnisse vor.

Bezeichnend ist aber, dass 62% der Befragten Männer waren. Dies zeigt sich auch in den Studierendenzahlen. So waren an der Universität Wien im Studienjahr 2015/16 173 Männer, aber nur 78 Frauen für das Bachelorstudium Lehramt Physik inskribiert (Unterrichtsfach Physik - Lehramt Sekundarstufe Allgemeinbildung, 2016). Demgegenüber waren an den pädagogischen Hochschulen in Österreich im gleichen Jahr 11.125 weibliche und nur 3.425 männliche Lehramtsstudierende (Austria, 2016). Dies unterstützt das Bild der Frauen in MINT-Fächern⁴. Auch wenn der Lehrberuf generell als weiblich angesehen wird, sind doch in den naturwissenschaftlichen Fächern wiederum verstärkt Männern zu finden.

Auch der Zusammenhang mit dem gewählten Zweitfach war nicht signifikant. Es macht also anscheinend keinen Unterschied, ob der Lehramtsstudierende sich nur auf naturwissenschaftliche Fächer spezialisiert, oder Physik mit einem geisteswissenschaftlichen Fach kombiniert. Die Stichprobengröße ist hier auch als problematisch anzusehen. Betrachtet man aber die Mittelwerte, so sieht man doch eine Tendenz: generell wurden die Schülerinnen schlechter beurteilt, wenn das Zweitfach auch naturwissenschaftlich war. Es wird also scheinbar ein anderer Beurteilungsrahmen gesetzt. Den gleichen Effekt sieht man auch, wenn man den Zusammenhang des gewählten Schulfaches mit der Note der Schülerinnen und Schüler betrachtet. Auch hier werden SchülerInnen unabhängig vom Geschlecht schlechter benotet, wenn das gewählte Zweitfach ein naturwissenschaftliches ist (naturwiss.: $M = 20.5$, $SD = 11.27$; geisteswiss.: $M = 14.0$, $SD = 16.64$).

Entgegen der Erwartungen zeigte fehlende Erfahrung aber keinen Effekt. Der Großteil der Befragten gab an über Erfahrung zu verfügen, vor allem über Erfahrung im Bereich des Nachhilfeunterrichtes. Hier kann zwar Erfahrung im Unterrichten gewonnen werden, aber nicht in der Benotung von SchülerInnen. Die Befunde von Hofer (2015) konnten also auch hier nicht repliziert werden. Grund für einen fehlenden Effekt könnte auch hier die geringe Stichprobengröße sein.

⁴ Zusammenfassende Bezeichnung für die Fächer Mathematik, Informatik, Naturwissenschaft und Technik

13.3 Bedürfnis nach kognitiver Geschlossenheit

In der Literatur finden sich zahlreiche Hinweise darauf, dass ein hohes Bedürfnis nach kognitiver Geschlossenheit zu einem verstärkten Zurückgreifen auf Stereotype bei der Beurteilung führt (Sun et al., 2016; Dijksterhuis et al., 1996; Kruglanski et al., 1996).

In der vorliegenden Studie zeigte sich kein signifikanter Zusammenhang zwischen dem Bedürfnis nach kognitiver Geschlossenheit und der Benotung von Schülerinnen.

Je höher das Bedürfnis nach kognitiver Geschlossenheit war, umso kleiner war die vergebene Schulnote. Da im österreichischen Schulsystem die Note umso besser ist, je kleiner sie ist, wurde also umso besser beurteilt. Allerdings ist das R^2 mit .164 sehr gering. Es können also nur 16,4% der Streuung der Schülerinnennote durch die Variable NCC erklärt werden.

Als Einschränkung muss auch hier wieder die geringe Stichprobengröße gesehen werden. Außerdem muss bedacht werden, dass der höchste Wert in der NCC Skala 49 war, bei einem maximal zu erreichenden Wert von 96. Der Großteil der Befragten hatte NCC Werte im unteren Bereich. Somit ist nicht eindeutig festzustellen, ob Personen mit Werten deutlich über der Mitte eine andere Tendenz zeigen.

Ein weiterer möglicher Grund wäre auch hier wieder die Aktivierung des Stereotyps der „fleißigen Schülerin“.

Sun et al. (2016) konnten auch zeigen, dass die Perspektivenübernahme die stereotype Sichtweise senken kann, wenn das Bedürfnis nach kognitiver Geschlossenheit niedrig ist. Da der Wert bei keinem Teilnehmer sehr hoch war, könnte auch diese Tatsache eine Rolle spielen. Zur Perspektivenübernahme könnte es dadurch gekommen sein, dass die Teilnehmer als Studierende sich mit der Rolle der SchülerInnen besonders gut identifizieren konnten und daher deren Perspektive unbewusst übernahmen.

14 Limitationen

Die Hauptlimitation dieser Studie besteht sicherlich in der geringen Anzahl von Teilnehmern. Der Fragebogen wurde insgesamt 172mal geöffnet, aber nur 32mal

vollständig bearbeitet. Dies entspricht einer Quote von annähernd 20 %, was für einen Onlinefragebogen durchaus üblich ist.

Betrachtet man die Gesamtstudierendenanzahl an den fünf österreichischen Fakultäten, die das Lehramtsstudium Physik anbieten, sieht man, dass die Zahl mit ca. 750 Studierenden, im Vergleich zu anderen Studiengängen ebenfalls klein ist. Zum Vergleich studierten im Jahr 2013 allein an der Universität Wien knapp 4500 Personen Psychologie (Studierende und Studienzulassungen nach Studienprogrammleitung und Studienrichtung, 2014). Werden Studien für Masterarbeiten mit Studierenden der Psychologie in Wien durchgeführt, beträgt die Teilnehmeranzahl häufig zwischen 150 und 200 Personen, ungefähr 4 % der StudentInnen in diesem Studiengang. Dies entspricht der Prozentzahl der Gesamtpopulation in der vorliegenden Studie.

Weiter muss bedacht werden, dass der Studienprogrammleiter der Universität Wien die Teilnahme seiner Studierenden abgelehnt hat. Diese wäre mit 240 Studierenden (Unterrichtsfach Physik - Lehramt Sekundarstufe Allgemeinbildung, 2016) relativ groß gewesen und auch leichter persönlich erreichbar. Als Grund nannte er die Kenntnis der Ursprungsstudie, wobei er mit der Art und Qualität des Fragebogens nicht einverstanden gewesen wäre.

Allerdings ist auch im Curriculum der Universität Wien zum Lehramtsstudium Physik die Geschlechterforschung Thema (Studienplan für das "Lehramtsstudium" an der Fakultät für Naturwissenschaften, 2002). Auch wenn neue Ergebnisse eventuell nicht in die gewünschte Richtung (Abnahme der Geschlechterstereotype bei Lehramtsstudierenden) weisen könnten, zur Überprüfung von Effekten könnten neue Studien zu diesem Thema Hinweise liefern. Generell zeigte sich aber auch ein sehr geringes Interesse der Studierenden selbst an dieser Studie. Dazu muss erinnert werden, dass den Teilnehmern bereits im Erklärungstext vermittelt wurde, an einer Studie zum Thema „Beurteilung“ teilzunehmen, wobei sie um ihre Mithilfe als zukünftige Experten gebeten wurden. Es war ihnen also bewusst, dass sie nicht Experten auf diesem Gebiet sind. Betrachtet man sich dann noch die Curricula der einzelnen Fakultäten, fällt auf, dass der Bereich des Beurteilens nicht immer einen großen Stellenwert in der universitären Ausbildung hat. So ist der Themenbereich „Leistungsbeurteilung“ im Curriculum der Universität Wien nur ein Unterpunkt eines aus mehreren auszuwählenden Themenbereiches

(Problemfelder des Unterrichts) in „Theorie und Praxis des Lehrens und Lernens“. An der Universität Innsbruck wiederum ist die „Beurteilungskompetenz“ an dritter Stelle der zu erwerbenden allgemein-didaktischen und fächerübergreifenden Kompetenzen zu finden (Studienplan für das Lehramtsstudium an der Universität Innsbruck in den Unterrichtsfächern Biologie und Umweltkunde, Chemie, Geographie und Wirtschaftskunde, Informatik und Informatikmanagement, Mathematik sowie Physik, 2011). Zur Ausbildung gehören auch Schulpraktika, in denen Erfahrungen in der Leistungsbeurteilung gesammelt werden können. Diese werden über den gesamten Verlauf des Studiums absolviert, wobei selbständige Beurteilung sicher nicht zu den primär zu erwerbenden praktischen Kompetenzen gehört. Darüber hinaus hängt die Qualität der Praktika auch von den anleitenden Lehrkräften ab. Möglicherweise bestand daher bei den Personen, die die Einladung zur Untersuchungsteilnahme bekamen, bereits durch den Erklärungstext im Email die Angst, nicht über genug Wissen zu verfügen, um einen positiven Beitrag zur Studie zu leisten. Dies sollte in einer weiterführenden Studie bedacht werden und der Erklärungstext entsprechend umformuliert werden.

Eine weitere Limitation der vorliegenden Studie ist die Verwendung nur eines Items zur Benotung von SchülerInnen. Im realen Schulkontext bestehen Tests im Regelfall aus mehreren Fragen, um so einen größeren Überblick über den derzeitigen Leistungsstand der SchülerInnen zu gewinnen. Das für diese Studie verwendete Testbeispiel war zwar aus einem existierenden Schulbuch gewählt und auch aus einem Bereich der Physik, der bereits sehr früh gelehrt wird, trotzdem sind Verständnisprobleme von Seiten der TeilnehmerInnen nicht auszuschließen.

Für weitere Studien wäre ein Test mit mehreren Prüfungsfragen aus verschiedenen Teilgebieten der Physik zu überlegen.

Zusätzlich sollten in einer weiteren Studie die geschlechtsstereotypen Einstellungen der TeilnehmerInnen erhoben werden.

Um größere TeilnehmerInnenzahl zu erreichen, wäre eine Ausweitung auf Deutschland zu erwägen oder mit allgemeiner gehaltenen Prüfungsfragen, die mehrere Fachgebiete betreffen, eine Ausweitung auf verwandte Studienfächer.

Zuletzt darf auch nicht vergessen werden, dass für die Untersuchung ein Fragebogen verwendet wurde. Dieser ist dem Phänomen der Verfälschbarkeit ausgesetzt (Kubinger & Jäger, 2003). Gerade in diesem Fall bei der Beantwortung von Fragen zu persönlichen Einstellungen könnte eine Antworttendenz in Richtung der Mitte eine Rolle gespielt haben. Darunter versteht man die Neigung einer Person, extreme Antworten zu vermeiden, um seine wirklichen Einstellungen nicht offen zu legen oder aus Unsicherheit, was man antworten soll (Kubinger et al., 2003).

15 Conclusio

In vielen Köpfen sind noch immer geschlechterstereotype Einstellungen stark vorhanden. Zu den Ergebnissen der PISA Studie 2012 bezüglich der Geschlechterunterschiede der Leistung in den Naturwissenschaften schrieb die OECD in einem Bericht folgendes:

In einem ersten Schritt kann es zur Veränderung dieser Denkmuster nötig sein, Mädchen Mathematik auf interessantere Weise näher zu bringen, geschlechtsspezifische Stereotypen in den Schulbüchern zu identifizieren und zu beseitigen, weibliche Vorbilder zu fördern und Lernmaterialien zu verwenden, die Mädchen ansprechen. Längerfristig werden zur Verringerung der geschlechtsspezifischen Leistungsunterschiede in Mathematik konzertierte Bemühungen seitens der Eltern, der Lehrkräfte sowie der Gesellschaft als Ganzes notwendig sein, um die stereotypen Vorstellungen darüber zu verändern, was Jungen und Mädchen jeweils besonders gut können, was sie gerne tun und wozu sie sich selbst in der Lage sehen (OECD, 2013, S. 6) .

Dies trifft nicht nur auf Mathematik zu, sondern auf den gesamten Bereich der Naturwissenschaften. Die Curricula der Lehramtsstudien haben die kritische Betrachtung von Stereotypen bereits teilweise aufgenommen. Die Ergebnisse dieser Studie könnten ein erster Hinweis darauf sein, dass zumindest bei einem Teil der zukünftigen Lehrkräfte ein „Umdenken“ stattgefunden hat.

Demgegenüber zeigen allerdings die Ergebnisse der letzten PISA-Studie, dass es noch ein weiter Weg ist, bis Frauen die gleichen Fähigkeiten in mathematischen und naturwissenschaftlichen Fächern zugeschrieben werden wie Männern.

Abbildungsverzeichnis

Abbildung 1 Regressionsgerade Schülerinnennote und NCC Studierende.....39

Tabellenverzeichnis

Tabelle 1 Zweidimensionales Modell der Kausalattributionen nach Weiner.....	19
Tabelle 2 Notenverteilung nach SchülerInnenengeschlecht.....	36
Tabelle 3 Schülerinnennote nach gewähltem Zweifach	38
Tabelle 4 Ergebnisse der schrittweisen Regression	40

Literaturverzeichnis

- Abrams, D., Crisp, R. J., Marques, S., Fagg, E., Bedford, L., & Provias, D. (2008). Threat inoculation: Experienced and imagined intergenerational contact prevents stereotype threat effects on older people's math performance. *Psychology and Aging, 23*, S. 934-939. doi:10.1037/a0014293
- Allport, G. W. (1954). *The Nature of Prejudice*. Cambridge: Addison-Wesley.
- Aronson, E., Wilson, T., & Akert, R. (2014). *Sozialpsychologie* (8th Ausg.). (M. Reiss, Übers.) Pearson.
- Ashmore, R. D., & Del Boca, F. K. (1979). Sex stereotypes and implicit personality theory: Toward a cognitive-social psychological conceptualization. *Sex Roles, 5*, S. 219-248.
- Austria, S. (2016). *Universitäten, Studium*. Abgerufen am 18. 08 2017 von http://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bildung_und_kultur/formales_bildungswesen/universitaeten_studium/index.html.
- Bem, S. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88* (4), S. 354-364.
- Berk, L. E. (2011). *Entwicklungspsychologie*. München: Pearson.
- Bundesministerium für Bildung (2017). *Geschlechtsspezifische Bildungsfragen / Gleichstellung von Mädchen und Buben*. (B. f. Bildung, Herausgeber) Abgerufen am 16. 08 2017 von <https://www.bmb.gv.at/schulen/unterricht/ba/gs/index.html>.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Auflage). Heidelberg: Springer.
- Bruneorth, M., Lassnigg, L., Vogenhuber, S., Schreiner, C., & Breit, S. (2016). *Nationaler Bildungsbericht Österreich 2015, Band 1*. Graz: Leykam.
- Burgess, D., & Borgida, E. (1999). Who women are, who women should be: Descriptive and prescriptive gender stereotyping in sex discrimination. *Psychology, Public Policy, and Law, 5*, S. 665-692.
- Carli, L., & Bukatko, D. (2000). Gender, communication and social influence: A development perspective. In T. Eckes, & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 295-332). NY: Lawrence Erlbaum Associates.

- Carrington, B., Tymms, P., & Merrell, C. (2008). Role models, school improvement and the 'gender gap'—do men bring out the best in boys and women the best in girls? *British Educational Research Journal*, 34(3), S. 315-327. doi:10.1080/01411920701532202
- Carter, K., Sabers, D., Cushing, K., Pinnegar, S., & Berliner, D. (1987). Processing and using information about students: A study of expert, novice, and postulant teachers. *Teaching and Teacher Education*, 3(2), S. 147-157. doi:10.1016/S0883-0355(02)00004-6
- Crowley, K., Callanan, M. A., Tenenbaum, H. R., & Allen, E. (2001). Parents explain more often to boys than to girls during shared scientific thinking. *Psychological Science*, 12, S. 258-261. doi:10.1111/1467-9280.00347
- Dar-Nimrod, I., & Heine, S. J. (2006). Exposure to scientific theories affects women's math performance. *Science*, Vol.314(5798), S. 435-435.
- Dickhäuser, O. (2006). Fähigkeitsselbstkonzepte: Entstehung, Auswirkung, Förderung. *Zeitschrift für Pädagogische Psychologie*, 20(1/2), S. 5-8.
- Dickhäuser, O., & Meyer, W.-U. (2006). Gender differences in young children's math ability attributions. *Psychology Science*, 48, S. 3-16.
- Dijksterhuis, A., van Knippenberg, A., Kruglanski, A. W., & Schaper, C. (1996). Motivated social cognition: Need for closure effects on memory and judgment. *Journal of Experimental Social Psychology*, 32(3), S. 254-270. doi:10.1006/jesp.1996.0012
- Dresel, M., Heller, K. A., Schober, B., & Ziegler, A. (2001). Geschlechtsunterschiede im mathematisch-naturwissenschaftlichen Bereich. Motivations- und selbstwertschädigende Einflüsse der Eltern auf Ursachenerklärungen ihrer Kinder in Leistungskontexten. In C. Finkberger, & G. W. Schnaitmann (Hrsg.), *Lehren und Lernen im Kontext empirischer Forschung und Fachdidaktik* (S. 270-288). Donauwörth: Auer.
- Dünnebier, K., Gräsel, C., & Krolak-Schwerdt, S. (2009). Urteilsverzerrungen in der schulischen Leistungsbeurteilung: Eine experimentelle Studie zu Ankereffekten. *Zeitschrift für Pädagogische Psychologie*, 23 (3-4), S. 187-195. doi:10.1024/1010-0652.23.34.187
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ.: Lawrence Erlbaum.

- Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen. In R. Becker, & B. Kortendieck (Hrsg.), *Handbuch Frauen- und Geschlechterforschung* (S. 171-182). Wiesbaden: Springer VS.
- Espinoza, P., da Luz Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, 17, S. 105-126. doi:10.1007/s11218-013-9226-6
- Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teacher's attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, 21, S. 55-69. doi:10.1007/BF00311015
- Field, A. (2009). *Discovering Statistics Using SPSS (Third Edition)*. London: SAGE.
- Filipp, S.-H. (2006). Entwicklung von Fähigkeitsselbstkonzepten. *Zeitschrift für Pädagogische Psychologie*, S. 65-72. doi:10.1024/1010-0652.20.12.65
- Finsterwald, M., & Ziegler, A. (2007). Geschlechtsrollenstereotype in Schulbuchabbildungen der Grundschule. In P. Ludwig, & H. Ludwig (Hrsg.), *Erwartungen in himmelblau und rosarot. Effekte, Determinanten und Konsequenzen von Geschlechtsdifferenzen in der Schule* (S. 117-142). Weinheim.
- Fischbach, A., Brunner, M., Krauss, S., & Baumert, J. (2015). Die Bezugsnormorientierung von Mathematiklehrkräften am Ende der Sekundarstufe I: Konvergenz verschiedener Messverfahren und Wirkung auf motivational-affektive Aspekte des Mathematiklernens und Leistung. *Journal for Educational Research Online*, 7 (3), S. 3-27.
- Friedrich, A., Flunger, B., Nagengast, B., Jonkmann, K., & Trautwein, U. (2015). Pygmalion effects in the classroom: Teacher expectancy effects on students' math achievement. *Contemporary Educational Psychology*, 41, S. 1-12. doi:10.1016/j.cedpsych.2014.10.006
- Gelman, S. A., & Nguyen, S. P. (2004). Mother-child conversations about gender: Understanding the acquisition of essentialist beliefs. *Monographs of the society for research in child development*, 69, S. 1-127.
- Hamilton, P. L., & Jones, L. (2016). Illuminating the 'boy problem' from children's and teachers' perspectives: a pilot study. *Education 3 - 13*, 44(3), S. 241.

- Hannover, B. (2008). Vom biologischen zum psychologischen Geschlecht: Die Entwicklung von Geschlechtsunterschieden. In A. Renkl (Hrsg.), *Lehrbuch Pädagogische Psychologie* (S. 339-388). Bern: Huber.
- Herkner, W. (2001). *Lehrbuch Sozialpsychologie*. Bern: Huber.
- Hodosi, T. (2014). *Geschlechtsstereotype von jüngeren und älteren Eltern dreibis sechsjähriger Kinder. Diplomarbeit*. Wien: Universität Wien/ Psychologische Fakultät.
- Hofer, S. I. (2015). Studying gender bias in physics grading: The role of teaching experience and country. *International Journal of Science Education*, 37(37), S. 2879-2905. doi:10.1080/09500693.2015.1114190
- Holmeier, M. (2012). Bezugsnormorientierung im Unterricht im Kontext zentraler Abiturprüfungen. In K. M. Merki (Hrsg.), *Zentralabitur* (S. 237-261). Wiesbaden: Springer VS.
- Ingenkamp, K., & Lissmann, U. (2008). *Lehrbuch der Pädagogischen Diagnostik*. Weinheim und Basel: Beltz.
- Jacobs, J., & Eccles, J. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology*, 63(6), S. 932-944. doi:10.1037/0022-3514.63.6.932
- Jones, S., & Myhill, D. (2004). 'Troublesome boys' and 'compliant girls': gender identity and perceptions of achievement and underachievement. *British Journal of Sociology of Education*, 25(5), S. 547-561. doi: 10.1080/0142569042000252044
- Jöstl, G., Kollmayer, M., Finsterwald, M., Schober, B., & Spiel, C. (2015). Geschlechterstereotype in der Bildungssozialisation - Reflexive Koedukation als Lösungsansatz. In B. Hoyer (Hrsg.), *Migration und Gender : Bildungschancen durch Diversity-Kompetenz* (S. 13-31). Opladen: Budrich UniPress.
- Jussim, L., Eccles, J., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: accuracy and the quest for the powerful self-fulfilling prophecy. *Advances in Experimental Social Psychology*, 28, S. 281-388.

- Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *Journal of Social Psychology*, 14 (2), S. 165-173.
- Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby, *The development of sex differences* (S. 82-172). Stanford: Stanford University Press.
- Kollmayer, M., Schober, B., & Spiel. (2016). Gender stereotypes in education: Development, consequences, and interventions. *European Journal of developmental psychology*, S. 1-17. doi:10.1080/17405629.2016.1193483
- Krapp, A., Geyer, C., & Lewalter, D. (2014). Motivation und Emotion. In T. Seidel, & A. Krapp, *Pädagogische Psychologie* (S. 193-222). Weinheim: Beltz Verlag.
- Krems, J. F. (1994). *Wissensbasierte Urteilsbildung: Diagnostisches Problemlösen durch Experten und Expertensysteme*. Bern: huber.
- Krolak-Schwerdt, S., Böhmer, M., & Gräsel, C. (2012). Leistungsbeurteilungen von Schulkindern: Welche Rolle spielen Ziele und Expertise der Lehrkraft? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(3), S. 111-122. doi:10.1026/0049-8637/a000062
- Kruglanski, A. W. (1990). Lay epistemic theory in social-cognitive psychology. *Psychological Inquiry* 1(3), S. 181-197. doi:10.1207/s15327965pli0103_1
- Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: "Seizing" and "freezing". *Psychological Review*, 103(2), S. 263-283. doi:10.1037/0033-295X.103.2.263
- Kubinger, K. D., & Jäger, R. S. (Hrsg.). (2003). *Schlüsselbegriffe der Psychologischen Diagnostik*. Weinheim, Basel, Berlin: Beltz Verlag.
- Kuhl, P., & Hannover, B. (2012). Differenzielle Benotung von Mädchen und Jungen: Der Einfluss der von der Lehrkraft eingeschätzten Kompetenz zum selbstgesteuerten Lernen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(3), S. 153-162. doi:10.1026/0049-8637/a000066
- Li, R., & Wong, W. (2016). Gender-typed play and social abilities in boys and girls: Are they related? *Sex Roles*, 74(9), S. 399-410. doi:10.1007/s11199-016-0580-7

- Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist*, 45(4), S. 513-520.
- Madon, S., Willard, J., Guyll, M., & Scherr, K. (2011). Self-fulfilling-prophecies: Mechanisms, power, ad links to social problems. *Social and Personality Psychology Compass*, 5(8). doi:10.1111/j.1751-9004.2011.00375.x
- Marchand, G. C., & Taasooobshirazi, G. (2013). Stereotype threat and women's performance in physics. *International Journal of Science Education*, 35(18), S. 3050-3061. doi:10.1080/09500693.2012.683461
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of refence model. *American Educational Research Journal*, 23, S. 129-149. doi: 10.3102/00028312023001129
- Niklas, F., & Schneider, W. (2012). Die Anfänge geschlechtsspezifischer Leistungsunterschiede in mathematischen und schriftsprachlichen Kompetenzen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 44(3), S. 123-138. doi:10.1026/0049-8637/a000064
- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = Male, Me = Female, Therefore Math $\neq$ Me . *Journal of Personality and Social Psychology*, 83(1), S. 44-59. doi: 10.1037/0022-3514.83.1.44
- Nürnberger, M., Nerb, J., Schmitz, F., Keller, J., & Sütterlin, S. (2016). Implicit gender stereotypes and essentialist beliefs predict preservice teachers' tracking recommendations. *The Journal of Experimental Education*, 84(1), S. 152-174. doi:10.1080/00220973.2015.1027807
- O'Brien, L. T., & Crandall, C. S. (2003). Stereotype Threat and Arousal: Effects on Women's Math Performance. *Personality and Social Psychology Bulletin*, 29, S. 782-789. doi:10.1177/0146167203252810
- OECD. (2011). *Report on the gender initiative: Gender equality in education, employment and entrepreneurship*. Abgerufen am 2. Mai 2017 von <http://www.oecd.org/education/48111145.pdf>
- OECD (Hrsg.). (2013). *Österreich-Ländernotiz-Ergebnisse aus PISA 2012*. Abgerufen am 4. Juni 2017 von www.oecd.org/pisa/keyfindings/PISA-results-austria-DEU.pdf.
- OECD (Hrsg.). (2015). *Österreich-Ländernotiz-Bildung auf einen Blick*. Abgerufen am 16. 08 2017 von <https://www.oecd.org/austria/Education-at-a-glance-2015-Austria-in-German.pdf>.

- PISA. (2015). *Grundkompetenzen am Ende der Pflichtschulzeit im internationalen Vergleich*. Abgerufen am 13. November 2016 von <https://www.bifie.at/node/3643>
- Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "Michael can't read!" Teacher's gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*, 107, S. 186-194.
- Rustemeyer, R. (1999). Geschlechtstypische Erwartungen zukünftiger Lehrkräfte bezüglich des Unterrichtsfaches Mathematik und korrespondierende (Selbst-)Einschätzungen von Schülerinnen und Schülern. *Psychologie in Erziehung und Unterricht*, 46 (3), S. 187-200.
- Salzburg, U. (2013). Curriculum für das Bachelorstudium Lehramt an der Universität Salzburg.
- Schäfer, S., & Seidel, T. (2015). Noticing and reasoning of teaching and learning components by pre-service teachers. *Journal for Educational Research Online*, 7(2), S. 34-58.
- Scherr, K. C., Madon, S., Gyll, M., Willard, J., & Spoth, R. (2011). Self-verification as a mediator of mothers' self-fulfilling effects on adolescents' educational attainment. *Personality & social psychology bulletin*, 37(5), S. 587-600. doi:10.1177/0146167211399777
- 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), S. 9-18. doi:10.1024/1010-0652.20.12.9
- Schlink, S., & Walther, E. (2007). Kurz und gut: Eine deutsche Kurzskala zur Erfassung des Bedürfnisses nach kognitiver Geschlossenheit. *Zeitschrift für Sozialpsychologie*, 38(3), S. 153-161. doi:10.1024/0044-3514.38.3.153
- Schmader, T., Johns, M., & Forbes, C. (2008). An integrated process model of stereotype threat effects on performance. *Psychological Review*, 115(2), S. 336-356. doi:10.1037/0033-295X.115.2.336
- Seidel, T., & Krapp, A. (Hrsg.). (2014). *Pädagogische Psychologie*, (6.Auflage). Basel: Beltz Verlag.
- Seidel, T., & Reiss, K. (2014). Lerngelegenheiten im Unterricht. In T. Seidel, & A. Krapp, *Pädagogische Psychologie* (S. 253-275). Weinheim: Beltz.

- Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46(3), S. 407-441. doi:10.3102/00346543046003407
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67, S. 415-437. doi:10.1146/annurev-psych-073115-103235
- Spiel, C., Schober, B., & Finsterwald, M. (2011). Brave Mädchen – böse Buben? Genderstereotype in der Bildungssozialisation. In G. Magerl, R. Neck, & C. Spiel (Hrsg.), *Wissenschaft und Gender* (S. 81-97). Wien: Böhlau Verlag.
- Spiel, C., Schober, B., & Litzenberger, M. (2008). *Evaluation der Eignungstests für das Medizinstudium in Österreich*. Projektbericht für das Bundesministerium für Wissenschaft und Forschung, Wien.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of african americans. *Journal of Personality and Social Psychology*, 69(5), S. 797-811. doi:10.1037/0022-3514.69.5.797
- Steins, G. (2010). *Handbuch Psychologie und Geschlechterforschung*. Wiesbaden: VS Verlag für Sozialwissenschaften. doi:10.1007/978-3-531-92180_1
- Stockinger, S., & Schestak, S. (2015). Frauen und Männer in Österreich: Gender Index 2015. (B. f. Frauen, Hrsg.) Wien. Abgerufen am 2. Mai 2017 von https://www.bmb.gv.at/frauen/gender/gender_index_2015.pdf?5oc6ap
- Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. *Journal of Personality and Social Psychology*, 77, S. 1213-1227. doi:10.1037/0022-3514.77.6.1213
- Studienplan für das "Lehramtsstudium" an der Fakultät für Naturwissenschaften.* (2002). Abgerufen am 5. November 2016 von http://www.univie.ac.at/mtbl93/nummer/2001_02_344.pdf.
- Studienplan für das Lehramtsstudium an der Universität Innsbruck in den Unterrichtsfächern Biologie und Umweltkunde, Chemie, Geographie und Wirtschaftskunde, Informatik und Informatikmanagement, Mathematik sowie Physik.* (2011). Abgerufen am 8. November 2016 von <https://www.uibk.ac.at/fakultaeten->

servicestelle/pruefungsreferate/gesamtfassung/la-natwi_stand-01.10.2011.pdf.

Studierende und Studienzulassungen nach Studienprogrammleitung und Studienrichtung. (2014). Abgerufen am 8. Juni 2017 von https://studienlehrwesen.univie.ac.at/fileadmin/user_upload/studienundlehrwesen/Statistische_Daten/studstat_72_2013S.pdf.

studium.at. (2017). Abgerufen am 18. 08 2017 von <https://www.studium.at/627564-medizin-aufnahmetest-gender-gap-heuer-wieder-etwas-groesser>.

Suchan, B., Wallner-Paschon, C., Bergmüller, S., & Schreiner, C. (. (2011). *PIRLS & TIMMS 2011: Schülerleistungen in Lesen, Mathematik und Naturwissenschaften in der Grundschule. Erste Ergebnisse.* Leykam. Abgerufen am 13. November 2016 von <https://www.bifie.at/buch/1742>

Sun, S., Zuo, B., Wu, Y., & Fangfang, W. (2016). Does perspective taking increase or decrease stereotyping? The role of need for cognitive closure. *Personality and Individual Differences*, 94, S. 21-25. doi:10.1016/j.paid.2016.01.001

Tiedemann, J. (2000). Parents' gender stereotypes and teachers' beliefs as predictors of children's concept of their mathematical ability in elementary school. *Journal of Educational Psychology*, 92(1), S. 144-151. doi:10.1037/0022-0663.92.1.144

Tiedemann, J. (2002). Teacher's gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics*, 50, S. 49-62. doi:10.1023/a:1020518104346

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, 47(4), S. 943-949. doi:10.1037/a0024047

Unterrichtsfach Physik - Lehramt Sekundarstufe Allgemeinbildung. (2016). Abgerufen am 8. Juni 2017 von <https://www.studium.at/studieren/universitaet-wien/unterrichtsfach-physik-lehramt-sekundarstufe-allgemeinbildung-bachelor>.

- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67(6), S. 1049-1062.
- Webster, D. M., & Kruglanski, A. W. (1997). Cognitive and social consequences of the need for cognitive closure. *European Review of Social Psychology*, 8(1), S. 133-173.
- Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer.
- Wigfield, A., Eccles, J. S., Yoon, K. S., Harold, R. D., Arbretton, A. J., Freedman - Doan, C., & Blumenfeld, P. C. (1997). Change in children's competence beliefs and subjective task values across the elementary school years: A 3-year study. *Journal of Educational Psychology*, 89(3), S. 451-469.
- Ziegler, A., Kuhn, C., & Heller, K. A. (1998). Implizite Theorien von Mathematik- und Physiklehrkräften zu geschlechtsspezifischer Begabung und Motivation. *Psychologische Beiträge*, 40, S. 271-287.

Herzlich Willkommen

In dieser Studie für meine Masterarbeit interessiere ich mich dafür, wie eine isolierte Bewertung einzelner Prüfungsantworten von SchülerInnen durch verschiedene Lehrpersonen mit einer Bewertung der vollständigen Prüfung durch eine einzige Lehrperson zusammenhängt.

Um auch den Einfluss von Erfahrung und Einstellungen überprüfen zu können, bitte ich auch Sie, als zukünftige Lehrperson, um Ihre Mithilfe.

Daher bitte ich Sie im Folgenden eine einzelne Prüfungsantwort durch die Zuweisung einer Note zu bewerten und im Anschluss noch einige allgemeinere Fragen zu beantworten.

Die Bewertung dient ausschliesslich Forschungszwecken.

Sowohl die SchülerInnen, deren Antworten bewertet werden sollen, als auch Ihre Angaben sind vollkommen anonym.

Besten Dank für Ihre Unterstützung

Zur Einschätzung der Repräsentativität der Auswahl teilnehmender Personen, bitte ich Sie, kurz ein paar Fragen zu Ihrer Person zu beantworten.

Bitte geben Sie hier Ihr Geschlecht an.

- ☐ weiblich
- ☐ männlich

Bitte tragen Sie hier Ihr Alter ein.

Jahre

An welcher Universität studieren Sie?

[Bitte auswählen]

Im wievielten Semester befinden Sie sich im Lehramtsstudium Physik?

im Semester

Bitte wählen Sie hier die Studienart aus, in der Sie studieren.

[Bitte auswählen]

Markieren Sie hier bitte alle Fächer, die Sie gegebenenfalls zusätzlich zu Physik studieren.

Wenn Sie nur Physik studieren, markieren Sie bitte die Option „kein weiteres Fach“.

- ☐ kein weiteres Fach
- ☐ Chemie
- ☐ Biologie
- ☐ Informatik
- ☐ Mathematik
- ☐ Deutsch
- ☐ Fremdsprache(n)
- ☐ Sport
- ☐ Musik
- ☐ Kunst
- ☐ Geschichte oder Erdkunde/Geografie oder Sozialkunde oder Wirtschaftskunde
- ☐ Religion(slehre) oder Ethikunterricht
- ☐ anderes

Haben Sie bereits Lehr Erfahrung in einer Schule oder als NachhilfelehrerIn?

- ☐ ja
- ☐ nein

Bitte geben Sie hier an, wieviele Semester Lehr Erfahrung an einer Schule Sie bereits haben.

Semester

Bitte geben Sie hier an, wieviele Semester Erfahrung Sie bereits als NachhilfelehrerIn haben.

Semester

Vielen Dank für Ihre Angaben

PHP-Code

```
urnDraw('urne', 'NO03');
$zufallszahl = value('NO03_01');
if ($zufallszahl == 1) {
    goToPage('101');
} elseif ($zufallszahl == 2) {
    goToPage('102');
} elseif ($zufallszahl == 3) {
    goToPage('103');
} elseif ($zufallszahl == 4) {
    goToPage('104');
}
```

Im Folgenden sehen Sie eine Prüfungsfrage aus dem Bereich Newtonsche Mechanik und die Antwort einer Schülerin. Sie besucht die Unterstufe eines Gymnasiums mit sprachlichem Schwerpunkt. Bitte benoten Sie ihre Antwort.

Ihr wurde folgende Prüfungsfrage gestellt:

Zwei Skateboard-Fahrer mit deutlich unterschiedlicher Masse stehen sich gegenüber und sind mit einem gespannten Seil verbunden. Der linke und leichtere Skater zieht aktiv am Seil, der schwerere rechte Skater hält es nur fest. Was passiert? Begründe deine Vermutung in ca. fünf bis sechs Sätzen. Reibung ist zu vernachlässigen.

Die Schülerin antwortete:

Generell gilt, die Kraft setzt sich zusammen aus der Masse einer Person und der Beschleunigung. Der rechte Skater muss das Seil genauso fest halten, wie der linke Skater am Seil zieht. Auf beide wirkt deshalb eine gleich grosse Kraft, auch wenn nur der linke Skater am Seil zieht. Folglich sollte nichts passieren, da sich die beiden Kräfte aufheben. Da die Masse des linken Skaters allerdings kleiner ist als die Masse des rechten Skaters, ergibt sich für den linken Skater eine grössere Beschleunigung als für den rechten Skater. Daher sollte sich vermutlich doch zumindest der linke Skater ein wenig in Richtung des rechten Skaters bewegen.

Bitte bewerten Sie die Antwort der Schülerin, indem Sie ihr eine Note nach österreichischem Notensystem zuweisen.

Bewegen Sie den Schieber dazu auf die gewünschte Position.

1 „sehr gut“

5 „nicht genügend“

Falls die Darstellung des Schiebereglers mit Ihrem Browser nicht richtig funktionieren sollte, geben Sie bitte hier die Note (auf eine Kommastelle genau: x,y) ein.

Note

PHP-Code

```
setNextPage('106');
```

Im Folgenden sehen Sie eine Prüfungsfrage aus dem Bereich Newtonsche Mechanik und die Antwort eines Schülers. Er besucht die Unterstufe eines Gymnasiums mit sprachlichem Schwerpunkt. Bitte benoten Sie seine Antwort.

Ihm wurde folgende Prüfungsfrage gestellt:

Zwei Skateboard-Fahrer mit deutlich unterschiedlicher Masse stehen sich gegenüber und sind mit einem gespannten Seil verbunden. Der linke und leichtere Skater zieht aktiv am Seil, der schwerere rechte Skater hält es nur fest. Was passiert? Begründe deine Vermutung in ca. fünf bis sechs Sätzen. Reibung ist zu vernachlässigen.

Der Schüler antwortete:

Generell gilt, die Kraft setzt sich zusammen aus der Masse einer Person und der Beschleunigung. Der rechte Skater muss das Seil genauso fest halten, wie der linke Skater am Seil zieht. Auf beide wirkt deshalb eine gleich grosse Kraft, auch wenn nur der linke Skater am Seil zieht. Folglich sollte nichts passieren, da sich die beiden Kräfte aufheben. Da die Masse des linken Skaters allerdings kleiner ist als die Masse des rechten Skaters, ergibt sich für den linken Skater eine grössere Beschleunigung als für den rechten Skater. Daher sollte sich vermutlich doch zumindest der linke Skater ein wenig in Richtung des rechten Skaters bewegen.

Bitte bewerten Sie die Antwort des Schülers, indem Sie ihm eine Note nach österreichischem Notensystem zuweisen.

Bewegen Sie den Schieber dazu auf die gewünschte Position.

1 „sehr gut“

5 „nicht genügend“

Falls die Darstellung des Schiebereglers mit Ihrem Browser nicht richtig funktionieren sollte, geben Sie bitte hier die Note (auf eine Kommastelle genau: x,y) ein.

Note

PHP-Code

```
setNextPage('106');
```

Im Folgenden sehen Sie eine Prüfungsfrage aus dem Bereich Newtonsche Mechanik und die Antwort einer Schülerin. Sie besucht die Unterstufe eines Gymnasiums mit naturwissenschaftlichem Schwerpunkt. Bitte benoten Sie ihre Antwort.

Ihr wurde folgende Prüfungsfrage gestellt:

Zwei Skateboard-Fahrer mit deutlich unterschiedlicher Masse stehen sich gegenüber und sind mit einem gespannten Seil verbunden. Der linke und leichtere Skater zieht aktiv am Seil, der schwerere rechte Skater hält es nur fest. Was passiert? Begründe deine Vermutung in ca. fünf bis sechs Sätzen. Reibung ist zu vernachlässigen.

Die Schülerin antwortete:

Generell gilt, die Kraft setzt sich zusammen aus der Masse einer Person und der Beschleunigung. Der rechte Skater muss das Seil genauso fest halten, wie der linke Skater am Seil zieht. Auf beide wirkt deshalb eine gleich grosse Kraft, auch wenn nur der linke Skater am Seil zieht. Folglich sollte nichts passieren, da sich die beiden Kräfte aufheben. Da die Masse des linken Skaters allerdings kleiner ist als die Masse des rechten Skaters, ergibt sich für den linken Skater eine grössere Beschleunigung als für den rechten Skater. Daher sollte sich vermutlich doch zumindest der linke Skater ein wenig in Richtung des rechten Skaters bewegen.

Bitte bewerten Sie die Antwort der Schülerin, indem Sie ihr eine Note nach österreichischem Notensystem zuweisen.

Bewegen Sie den Schieber dazu auf die gewünschte Position.

1 „sehr gut“ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ 5 „nicht genügend“

Falls die Darstellung des Schiebereglers mit Ihrem Browser nicht richtig funktionieren sollte, geben Sie bitte hier die Note (auf eine Kommastelle genau: x,y) ein.

Note

PHP-Code

```
setNextPage('106');
```

Im Folgenden sehen Sie eine Prüfungsfrage aus dem Bereich Newtonsche Mechanik und die Antwort eines Schülers. Er besucht die Unterstufe eines Gymnasiums mit naturwissenschaftlichem Schwerpunkt. Bitte benoten Sie seine Antwort.

Ihm wurde folgende Prüfungsfrage gestellt:

Zwei Skateboard-Fahrer mit deutlich unterschiedlicher Masse stehen sich gegenüber und sind mit einem gespannten Seil verbunden. Der linke und leichtere Skater zieht aktiv am Seil, der schwerere rechte Skater hält es nur fest. Was passiert? Begründe deine Vermutung in ca. fünf bis sechs Sätzen. Reibung ist zu vernachlässigen.

Der Schüler antwortete:

Generell gilt, die Kraft setzt sich zusammen aus der Masse einer Person und der Beschleunigung. Der rechte Skater muss das Seil genauso fest halten, wie der linke Skater am Seil zieht. Auf beide wirkt deshalb eine gleich grosse Kraft, auch wenn nur der linke Skater am Seil zieht. Folglich sollte nichts passieren, da sich die beiden Kräfte aufheben. Da die Masse des linken Skaters allerdings kleiner ist als die Masse des rechten Skaters, ergibt sich für den linken Skater eine grössere Beschleunigung als für den rechten Skater. Daher sollte sich vermutlich doch zumindest der linke Skater ein wenig in Richtung des rechten Skaters bewegen.

Bitte bewerten Sie die Antwort des Schülers, indem Sie ihm eine Note nach österreichischem Notensystem zuweisen.

Bewegen Sie den Schieber dazu auf die gewünschte Position.

1 „sehr gut“ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ 5 „nicht genügend“

Falls die Darstellung des Schiebereglers mit Ihrem Browser nicht richtig funktionieren sollte, geben Sie bitte hier die Note (auf eine Kommastelle genau: x,y) ein.

Note

PHP-Code

```
setNextPage('106');
```

Um noch zusätzliche Informationen zur Beantwortung meiner Forschungsfrage zu erhalten, bitte ich Sie im Anschluss nun noch folgende kurze Fragen zu beantworten.

Es geht darum, wie Sie generell mit verschiedenen Situationen umgehen. Dies kann Ihren angestrebten Beruf, Ihr Studium oder auch andere Themen betreffen.

Kreuzen Sie bitte an, wie stark die nachfolgenden Sätze Ihre persönliche Einstellung, Meinung oder Erfahrung ausdrücken.

	stimme gar nicht zu							stimme völlig zu
1) Ich mag es nicht, wenn die Aussage einer Person mehrdeutig ist.	☐	☐	☐	☐	☐	☐	☐	☐
2) Ich finde, nachdem ich eine Lösung für ein Problem gefunden habe, ist es Zeitverschwendung, weitere mögliche Lösungen in Betracht zu ziehen.	☐	☐	☐	☐	☐	☐	☐	☐
3) Ich mag keine unvorhersehbaren Situationen.	☐	☐	☐	☐	☐	☐	☐	☐
4) Ich finde es spannend nicht zu wissen, was das Leben einem bringen wird.	☐	☐	☐	☐	☐	☐	☐	☐
5) Ein Problem aus verschiedenen Blickwinkeln zu betrachten, führt nur zu Verwirrung.	☐	☐	☐	☐	☐	☐	☐	☐
6) Im Allgemeinen suche ich nicht nach Alternativlösungen für Probleme, für welche ich schon eine Lösung parat habe.	☐	☐	☐	☐	☐	☐	☐	☐
7) Ich bevorzuge die Gesellschaft guter Freunde, weil ich weiß, was ich von ihnen zu erwarten habe.	☐	☐	☐	☐	☐	☐	☐	☐
8) Ich fühle mich unbehaglich, wenn ich es nicht schaffe eine schnelle Antwort auf Probleme zu geben, denen ich gegenüber stehe.	☐	☐	☐	☐	☐	☐	☐	☐
9) Ich bevorzuge Tätigkeiten, bei denen stets klar ist, was getan und wie es getan werden muss.	☐	☐	☐	☐	☐	☐	☐	☐
10) Wenn ich ein Problem lösen muss, verschwende ich im Allgemeinen keine Zeit damit, die unterschiedlichen Standpunkte dazu zu erwägen.	☐	☐	☐	☐	☐	☐	☐	☐
11) Ich mag Aufgaben, bei denen noch unklar ist, wie der genaue Lösungsweg aussieht.	☐	☐	☐	☐	☐	☐	☐	☐
12) Ich liebe die Ungewissheit und die Überraschung, die oft im Alltäglichen steckt.	☐	☐	☐	☐	☐	☐	☐	☐
13) Jedwede Lösung eines Problems ist besser, als in einem Zustand der Ungewissheit zu verharren.	☐	☐	☐	☐	☐	☐	☐	☐
14) Ich ziehe Dinge, die ich gewohnt bin, solchen vor, die ich nicht kenne und die ich nicht vorhersagen kann.	☐	☐	☐	☐	☐	☐	☐	☐
15) Im Allgemeinen vermeide ich es, mich an Diskussionen über uneindeutige und umstrittene Themen zu beteiligen.	☐	☐	☐	☐	☐	☐	☐	☐
16) Ich bevorzuge es, mich für die erstmögliche Lösung zu entscheiden, anstelle lange darüber nachzudenken, was für eine Entscheidung ich treffen sollte.	☐	☐	☐	☐	☐	☐	☐	☐

Bei Interesse sende ich Ihnen nach Abschluss der Untersuchung gerne die wichtigsten Ergebnisse per E-mail zu. Falls Sie dies möchten, geben Sie mir bitte Ihre E-mail Adresse bekannt. Ihre E-mail Adresse wird getrennt von der Erhebung gespeichert. Ihre bisher getätigten Eingaben bleiben also absolut anonym!

- ☐ Ich interessiere mich für die **Ergebnisse dieser Studie** und hätte gerne eine Zusammenfassung per E-Mail.

Letzte Seite

Die Umfrage ist zu Ende

Vielen herzlichen Dank für Ihre Unterstützung

Karin Nigischer BSc, Universität Wien

T-Test für unabhängige Stichproben: Note gesamt und geisteswissenschaftliches
Zweifach

Gruppenstatistiken

			Statistik	Bootstrap ^a		
				Standard Fehler	95% Konfidenzintervall	
					Untere	Obere
Geist		Zweit				
Note_Gesamt	7,00	N	24			
		Mittelwert	20,5000	,0382	2,2267	16,2500 24,9086
		Standardabweichung	11,26750	-,29996	1,34928	8,26453 13,55074
		Standardfehler des Mittelwertes	2,29997			
	8,00	N	8			
		Mittelwert	14,0000	,0112	4,4069	5,5560 22,8881
		Standardabweichung	12,63781	-1,05831 ^b	2,49299 ^b	5,67009 ^b 15,16622 ^b
		Standardfehler des Mittelwertes	4,46814			

a. Sofern nicht anders angegeben, beruhen die Bootstrap-Ergebnisse auf 5000 Bootstrap-Stichproben

b. Basierend auf 4998 Stichproben