

Geschlechtsspezifische Unterschiede in Mathematikleistungen: Welchen Einfluss haben Persönlichkeitseigenschaften auf die Lösungswahrscheinlichkeit von Matura-Aufgaben

EVA SATTLBERGER, WIEN; JAN STEINFELD, WIEN & PHILIPP GEWESSLER, WIEN

Als Teil des Aufgabenqualitätsprozesses der standardisierten schriftlichen Reifeprüfung in Mathematik wird jährlich eine empirische Überprüfung potentieller Prüfungsaufgaben durchgeführt. Im Zuge dieser Feldtestungen wurde 2017 begleitend ein Fragebogen vorgegeben, um das akademische Selbstkonzept, die Selbstwirksamkeit und die Geschlechtsstereotype der teilnehmenden Schülerinnen und Schüler zu erheben. Die Ergebnisse der Analysen zeigen unter anderem, dass neben strukturellen Faktoren insbesondere das selbstberichtete Selbstkonzept stark mit der Performanz zusammenhängt. Wird der a priori bestehende Geschlechtsunterschied in den erhobenen Persönlichkeitsvariablen mitberücksichtigt, verkleinert sich die vorhergesagte Leistungsdifferenz zwischen Schülerinnen und Schülern wesentlich. Die Resultate stützen Ergebnisse aus der Feldtestung 2016, indem sie den Zusammenhang von Testleistung und akademischem Selbstkonzept replizieren. In diesem Beitrag werden die Analysen vorgestellt und die Ergebnisse im Kontext von Geschlechtsunterschieden in der Mathematikleistung diskutiert.

1. Einleitung

Die Daten der Ergebnisse der standardisierten schriftlichen Reifeprüfung (SRP) in Mathematik an AHS in Österreich werden seit der flächendeckenden Einführung im Jahr 2014/2015 erhoben. Dabei zeigen sich von Beginn an große Unterschiede in den gezeigten Leistungen der Mädchen und Burschen¹. Tabelle 1 zeigt die Negativquoten (Anteil der Nicht genügend) der vergangenen vier Haupttermine (jeweils im Mai des Schuljahres).

Schuljahr	Gesamt	weiblich	männlich
2014/15	10.5 %	12.6 %	7.6 %
2015/16	21.8 %	25.4 %	16.7 %
2016/17	11.9 %	14.3 %	8.3 %
2017/18	22.4 %	25.5 %	17.9 %

Tabelle 1: Prozentueller Anteil an Nicht genügend bei der SRP in Mathematik AHS differenziert nach Geschlecht in den Hauptterminen 2015-2018

Diese Ergebnisse überraschen insofern nicht, da auch PISA 2015 Unterschiede in den Mathematikleistungen der Mädchen und Burschen aufzeigt (Suchań & Breit, 2016). Im Vergleich dazu fallen die Unterschiede in TIMSS (2011) und in der Bildungsstandardüberprüfung M8 (BiSta 2017) moderat aus. Bei allen Testungen wird davon ausgegangen, dass die Instrumente an sich „genderfair“ sind, die geschlechtsspezifischen Unterschiede also nicht durch die Überprüfung selbst hervorgerufen werden. Als mögliche Einflussfaktoren für die gezeigten Leistungen können einerseits Persönlichkeitseigenschaften bzw. -merkmale angenommen werden, welche mit zunehmendem Alter stärker entwickelt sind (siehe Ausführungen weiter unten), andererseits werden Leistungen auch durch Zuschreibungen der Bezugspersonen beeinflusst (Lehrer/innen und Eltern haben an Mädchen und Burschen unterschiedliche Leistungserwartungen). Bei der SRP wird zudem davon ausgegangen, dass die unterschiedliche Grundgesamtheit der Mädchen und Burschen einen Einfluss auf die Ergebnisse hat (vgl. Neuwirth, 2015). Aus Abbildung 1 ist zu entnehmen, dass deutlich mehr Mädchen eine Reifeprüfung abschließen als Burschen, während das Geschlechterverhältnis bei PISA und Bildungsstandards noch annähernd ausgeglichen ist.

¹ Auch im Fach Englisch werden Unterschiede zuungunsten der Mädchen in allen Auswertungsbereichen festgestellt.

Unter Persönlichkeitseigenschaften werden in diesem Artikel Einschätzungen und Einstellungen bzgl. unterschiedlicher Aspekte, die eigene Person betreffend, verstanden. Dabei spielt das Selbstkonzept eine wesentliche Rolle. Das (bereichsspezifische) Selbstkonzept einer Person wird als mentale Repräsentation der eigenen Person gesehen, es werden Vorstellungen, Einschätzungen und Bewertungen, die die eigene Person betreffen, zusammengefasst (vgl. Moschner, 2001). Selbstbeschreibungen können sich auf einzelne Facetten der Person (*Ich zeige in Mathematik gute Leistungen.*) oder auf die gesamte Person (*Ich wünschte, ich wäre jemand anders.*) beziehen. Die Entwicklung des Selbstkonzepts einer Person ist von ihrer jeweiligen Umwelt abhängig, die tatsächliche Leistung wird durch eine positive Bewertung der eigenen Leistungsfähigkeit stark beeinflusst und umgekehrt.

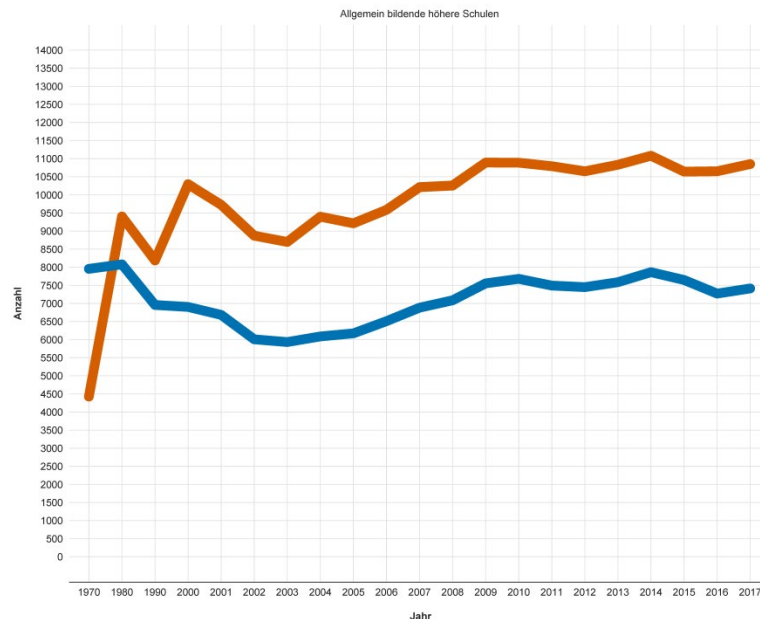


Abb. 1: bestandene Reifeprüfungen von 1970 bis 2017²

Unter dem Begriff der Selbstwirksamkeit wird die Überzeugung verstanden, neue und/oder schwierige Situationen aufgrund eigener Kompetenz bewältigen zu können (vgl. Bandura, 1997). Zur Herausbildung derartiger Erwartungen sind internal-stabile Attributionen auf die eigene Begabung notwendig. Eine empirische Abgrenzung zwischen den Begriffen Selbstwirksamkeit und Selbstkonzept ist aber genau genommen schwierig zu treffen (vgl. Abschnitt 3).

Shavelson (1976) beschreibt ein hierarchisches Modell zur Unterscheidung zwischen schulischem und nicht-schulischem Selbstkonzept. Das Individuum organisiert seine Erfahrungen in Kategorien, um die Komplexität zu reduzieren. Abbildung 2 zeigt das Modell einer zunehmenden Differenzierung der Kategorien im Laufe der zeitlichen Entwicklung. Das schulische (akademische) Selbstkonzept wird dabei u.a. in verbale und mathematische Komponenten unterteilt (beide korrelieren im Hinblick auf das Selbstkonzept einer Person praktisch nicht).

² Quelle: https://www.statistik.at/web_de/statistiken/menschen_und_gesellschaft/bildung_und_kultur/formales_bildungswesen/bildungsabschluss/index.html

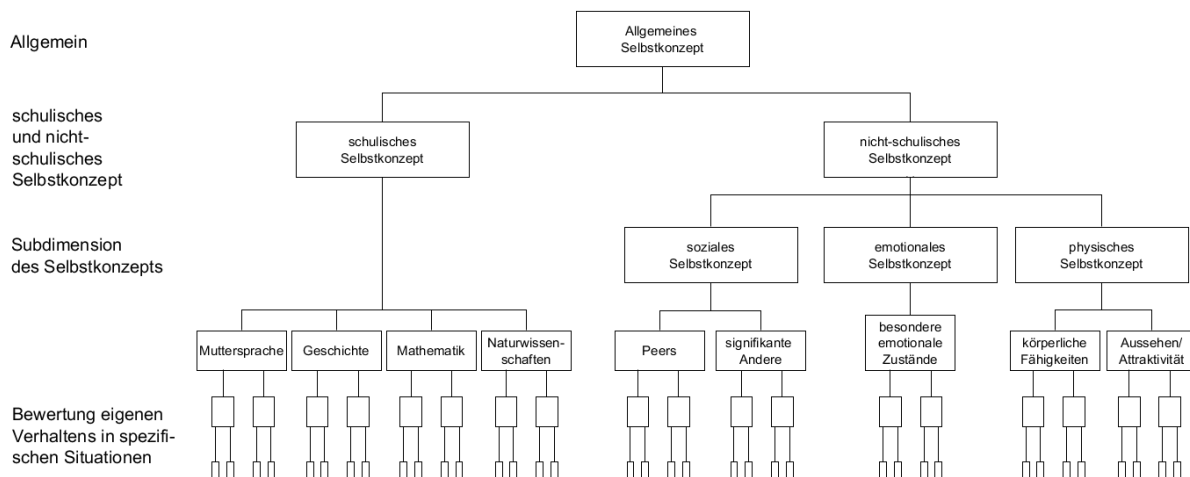


Abb. 2: zunehmende Differenzierung des Selbstkonzepts (nach einer Vorlage von Wild & Möller, 2009)

Geschlechterstereotype haben eine große theoretische und praktische Relevanz für die Entwicklung des Selbstkonzepts einer Person. Differenziert man die Ausprägung von schulbezogenen Selbstkonzepten nach dem Geschlecht, so zeigen sich recht konsistent Unterschiede, die den allgemeinen Geschlechterstereotypen entsprechen (vgl. Marsh & Hattie, 1996; Watt & Eccles, 2008). So lassen sich geschlechtsspezifische Unterschiede einerseits auf das Denken und Handeln von zentralen Bezugspersonen wie Eltern und Lehrerinnen und Lehrern zurückführen (vgl. Frome & Eccles, 1998), sie erwarten dabei im Fachbereich Mathematik von Burschen meist bessere Leistungen als von Mädchen, wodurch das Selbstkonzept der Mädchen negativ und das der Burschen positiv beeinflusst wird. Andererseits scheinen Lehrer/innen zudem bei gleichem Leistungsstand bei Burschen eine höhere Begabung und bei Mädchen dagegen ein stärkeres Ausmaß an Fleiß wahrzunehmen (vgl. Trautwein & Baeriswyl, 2007). Dies beeinflusst wiederum das Selbstkonzept der beiden Geschlechtergruppen unterschiedlich. Es existiert ein sogenannter *stereotype threat*, wahrgenommene Stereotype sind also in weiterer Folge relevant für das zu entwickelnde Selbstkonzept. Geschlechterstereotype spiegeln daher nur teilweise die tatsächlich vorhandenen Leistungsunterschiede wider, Selbstkonzepte werden so auch in Testsituationen sichtbar.

Zentrales Bildungsziel in der Schule ist u.a. die Unterstützung der Schüler/innen bei der Identitäts- und Persönlichkeitsfindung (vgl. Hellmich, 2011). Dabei hat das Selbstkonzept der Schüler/innen eine Schlüsselrolle bei der subjektiven Bewältigung und Bewertung schulischer Lern- und Leistungsanforderungen (vgl. Helmke, 1998), das bedeutet, dass für den Lehr- und Lernkontext bedeutsame Variable durch das Selbstkonzept der Schüler/innen mitbestimmt werden und durch das fähigkeitsbezogene Selbstkonzept (welches u.a. über die Motivation entwickelt wird) die Kompetenzentwicklung in einem Fach beeinflusst wird.

5-8-jährige Kinder wissen, dass ihr Aktivitäten von außen (durch Eltern, Lehrer/innen, Geschwister) beobachtet, bewertet und ggf. gewürdigt werden, es wird also die Perspektive anderer wirksam. 9-12-jährige Kinder nehmen einzelne Selbstrepräsentationen wahr und koordinieren diese (v.a. durch den sozialen Vergleich). Bis weit in die Kindheit hinein ist das Selbstkonzept positiv verzerrt, im Laufe der Entwicklung nimmt aber die Genauigkeit der Selbsteinschätzung – durch die allmähliche Integration verschiedener Informationen über die eigenen Fähigkeiten – zu. Zur gleichen Zeit werden auch schulische Einflüsse bedeutsamer und eine genauere Ausdifferenzierung des eigenen Rollenbildes findet statt. Erste gegen Ende der Jugendzeit existieren relativ stabile Überzeugungen und Werte.

Doch welche Faktoren beeinflussen nun die Höhe des fachbezogenen Selbstkonzepts einzelner Personen? Die sogenannten Determinanten des Selbstkonzepts werden über soziale, temporale, dimensionale und kriteriale Vergleiche ausgeprägt. Im sozialen Vergleich (vgl. Festinger, 1954), welcher sich bereits im Volksschulalter beobachten lässt und über die Folgejahre an Bedeutung gewinnt, messen die Schüler/innen ihre Leistungen an Klassenkameraden, wobei ein Abwärtsvergleich das Selbstkonzept erhöht und ein Aufwärtsvergleich dieses senkt. Im temporalen Vergleich (vgl. Nicholls, 1978) findet ein längs-

schnittlicher Abgleich der eigenen Fähigkeiten und damit eine direkte Information über eine Leistungssteigerung statt. Je günstiger dieser Vergleich ausfällt, desto höher ist auch das wahrgenommene Selbstkonzept, wobei sich Burschen eher mit früheren Leistungen vergleichen als Mädchen dies tun. Beim dimensionalen Vergleich findet ein intraindividueller Vergleich der eigenen Leistungen zwischen mehreren Domänen statt (vgl. Marsh, 1986). Beim kriterialen Vergleich wird als Kriterium das Zeigen einer bestimmten Leistung herangezogen (z.B. die Note Befriedigend in einem bestimmten Fach).

Oft herrscht eine (große) Diskrepanz zwischen aktiver Leistung und der subjektiven Selbsteinschätzung. Wichtig ist dabei die Ursachenzuschreibung zu berücksichtigen. Während Mädchen ihre guten Leistungen eher auf Fleiß zurückführen bzw. sie die gelösten Aufgaben als „leicht“ bezeichnen, schreiben Burschen ihre guten Leistungen dem eigenen Können zu. Dementsprechend werden auch weiterführende Herausforderungen dann gewählt. Wichtig sind diese Erkenntnisse vor allem für Lehrer/innen, da sie dazu beitragen können die Selbstbilder der Schüler/innen besser nachzuvollziehen und allen die entsprechende Förderung zukommen zu lassen.

Empirisch gut belegt ist, dass sich über die gesamte Grundschulzeit geschlechtsspezifische Unterschiede in Bezug auf die Selbstwahrnehmung eigener Fähigkeiten zeigen (vgl. Hellmich, 2011; Helmke, 1998; Köller & Klieme, 2000). Mädchen schätzen sich besser in Deutsch und Burschen besser in Mathematik ein, zudem kann bei Burschen eine höhere Anstrengungsvermeidung nachgewiesen werden als bei Mädchen. Verantwortlich dafür dürften auch im Grundschulalter schon internalisierte Geschlechterstereotype sein. Interessant ist auch, dass sich diese Unterschiede im Selbstkonzept zwar schon zeigen, diese aber noch größer sind als die Unterschiede in den tatsächlich gezeigten Leistungen. Dennoch beeinflusst das Selbstkonzept – wie oben beschrieben – die Lernleistung und die Leistungsentwicklung. Ein hohes Selbstkonzept fördert (über die Motivation) Lernprozesse in bestimmten Domänen und wirkt unterstützend beim Kompetenzerwerb und in Performanzsituationen. Insofern erklären Persönlichkeitsmerkmale leistungsbezogenes Verhalten. Dies konnte auch in der in diesem Artikel in Abschnitt 2 und 3 beschriebenen Feldtestungsstudie repliziert werden.

Das fachbezogene Selbstkonzept von Individuen kann von schulischen und außerschulischen Rückmeldungen beeinflusst werden. Leistungen entwickeln sich dann positiv, wenn Schüler/innen glauben erfolgreich sein zu können und das Fach als interessant und wichtig erachtet wird. Marsh (1986) entwickelte das *Internal External-Frame of Reference-Modell*, welches den Zusammenhang zwischen fachspezifischen Schulleistungen und fachspezifischem Selbstkonzept gut belegt. Das Modell beschreibt das Bedürfnis nach Selbsteinschätzung und der Steigerung der eigenen Fähigkeiten. Ein hohes Selbstkonzept – unabhängig vom Geschlecht – beeinflusst demnach die Leistungsfähigkeit und die tatsächliche Leistung. Während soziale Vergleiche und Bezugsgruppen wenig beeinflusst werden können, wirken sich positive Lernerfahrungen und Leistungsrückmeldungen sowie unterstützendes Verhalten von Lehrerinnen und Lehrern und Eltern förderlich auf die realistische Entwicklung des tatsächlichen Anspruchsniveaus aus. Dabei kommt es in der Schule auf die Wahl der Inhalte und Kontexte, Methoden und die gewährten Freiräume an. Durch bestimmte Lernumwelten (wie z.B. alleinigen Frontalunterricht) werden Burschen bevorzugt, während sich Mädchen eher zurückziehen. Instruktionen gehen dabei meist in Richtung der Burschen, um sie vom Plaudern abzuhalten und Mädchen brauchen mehr Beteiligung am Geschehen, damit sie auch glauben, dass sie das Unterrichtsgeschehen beeinflussen können. Förderlich für eine positive Entwicklung des Selbstkonzepts sind gewisse Faktoren im Bereich Leistungsbeurteilung und -bewertung. Individuelle und kriterienorientierte Bewertungsmaßstäbe (also eine Bezugsnormorientierung bei der Benotung) fördern die Entwicklung des Selbstkonzepts positiv. Unterstützendes Feedbackverhalten und Lob für Anstrengungen (besser als für Fähigkeiten) in einem binnendifferenzierten Unterricht, etwa durch die individuelle Zuweisung von Aufgaben, wirken sich ebenfalls förderlich aus (vgl. von Ow & Husfeldt, 2011).

Prinzipiell gilt, dass das Geschlecht der Lehrperson keinen Einfluss auf die Lernleistung der Schüler/innen hat, viel wichtiger ist die eigene Rolle als Mann bzw. als Frau und das eigene Selbstkonzept. Indirektes *Doing* bzw. *Undoing Gender* findet durch unbewusste Attribuierungen statt. Dabei wird bei gleichem Leistungsstand Mädchen höherer Fleiß und Burschen meist eine höhere Begabung zugeschrieben. Auch die Sprache im Unterricht spielt eine große Rolle, Lehrer/innen spiegeln geschlechtsspezifische Zuschreibungen (im Sinne des stereotype threat) durch unterschiedliche Rückmeldeprozesse an Mäd-

chen und Burschen wider, wobei Burschen weniger von diesen Zuschreibungen abhängig sind als Mädchen. Lehrer/innen sollten daher ihre eigenen Erwartungshaltungen überprüfen und flexible Rollenbilder früh fördern. Studien zeigen, dass Schulungen bei Lehrerinnen und Lehrern hinsichtlich der geschlechtersensiblen Pädagogik wirken, Mathematikleistungen können über eine Zunahme der Selbstwirksamkeitseinschätzung der Schüler/innen gesteigert werden.

2. Datenerhebung

Die Datenerhebung der vorliegenden Studie fand im März 2017 an allgemeinbildenden höheren Schulen in Österreich (AHS) im Rahmen der so genannten Feldtestungen statt. Die Feldtestungen stellen einen Teil der Qualitätssicherung dar, die die gleichbleibend hohe Qualität der für die schriftliche Reife- und Diplomprüfung in Österreich erstellten Aufgaben sicherstellen soll. Unter maturaähnlichen Bedingungen bearbeiten Schüler/innen der Abschlussklassen die Aufgaben. Im Anschluss an die Aufgabenbearbeitung wurde in der Feldtestung ein Fragebogen vorgegeben, der neben technischen Aspekten auch die Persönlichkeitseigenschaften Selbstkonzept, Selbstwirksamkeit und Geschlechtsstereotype abdeckte. Diesen sollten Schülerinnen und Schüler bearbeiten, sobald sie mit den Aufgaben der Testung fertig waren. Der Fragebogen wurde in interdisziplinärer Kooperation erstellt und in den Feldtestungen 2016, 2017 und auch 2018 mit leichten Adaptionen erfolgreich eingesetzt.

Insgesamt wurden 300 Aufgaben 2242 Schülerinnen und Schülern (1308 weiblich, 934 männlich; 30,1 % Gymnasium, 38,8 % Realgymnasium 31,2 % Oberstufenrealgymnasium) vorgegeben. Die Testungen fanden an 56 Standorten, verteilt über alle Bundesländer in Österreich statt. Alle Schüler/innen bearbeiteten dabei 21 Aufgaben in drei Unterrichtseinheiten (inkl. Instruktion und Fragebogen). Damit alle Aufgaben gemeinsam skaliert werden können, wurde ein Multi-Matrix-Design, mit einer dreifachen Verlinkung von jeder Aufgabe verwendet, sodass insgesamt 45 unterschiedliche Testhefte in die Testung eingeflossen sind. Im Anschluss an die eigentliche Aufgabenbearbeitung wurde der oben beschriebene Fragebogen vorgegeben.

3. Auswertung

3.1. Faktorenstruktur des Fragebogens

Für die Analysen des Fragebogens wurde zunächst eine konfirmatorische Faktorenanalyse (mit dem R Paket lavaan, Rosseel, 2012), zur Validierung der Dimensionszuordnung gerechnet. In der Modellierung wurden aufgrund theoretischer Überlegungen Abhängigkeiten zwischen den latenten Dimensionen zugelassen. Das beschriebene Modell weist einen guten Fit auf ($\chi^2(87) = 1951.08$, TLI = 0.96, RMSEA = 0.098 [0.094, 0.102]). Die Korrelation zwischen den Faktoren Selbstkonzept und Selbstwirksamkeit ist wie erwartet hoch, $r = 0.74$, während die Faktorkorrelationen zwischen Geschlechtsstereotypen und Selbstkonzept bzw. Geschlechtsstereotypen und Selbstwirksamkeit gering ausfällt, $r = -0.13$ und $r = -0.18$. Die bei der Modellierung gewonnen gewichteten arithmetischen Mittel jeder latenten Dimension werden in die nachfolgenden Analysen als Prädiktoren der Personenfähigkeit mit aufgenommen. Abbildung 3 ist die Verteilung der Faktorscores je latente Dimension zu entnehmen.

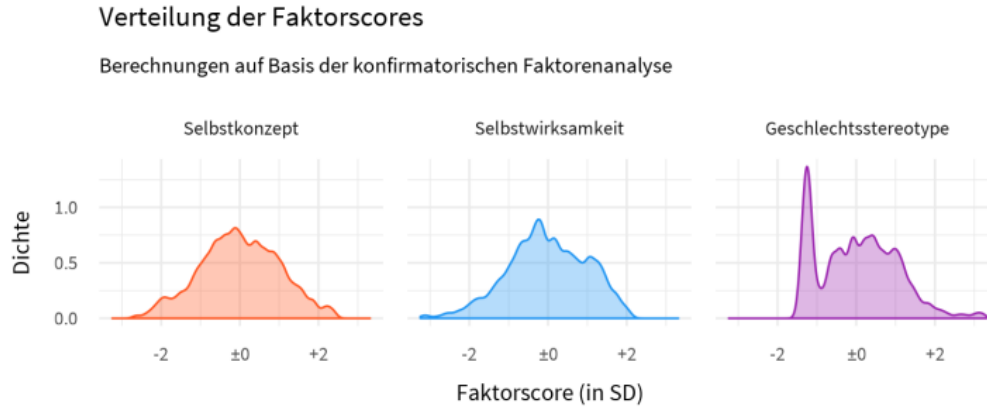


Abb. 3 Verteilung der Faktorscores der drei latenten Dimensionen Selbstkonzept, Selbstwirksamkeit und Geschlechtsstereotype

3.2. Modellierung der Leistungsdaten

Die anschließende Auswertung und Modellierung der Daten erfolgte mit Modellen der Item Response Theory (IRT). In IRT-Modellen werden dabei Personenparameter und Itemparameter auf Basis der jeweils beobachteten Antworten geschätzt. Im einfachsten Fall wird ein einziger Personenparameter (Personenfähigkeit) θ und ein einziger Itemparameter (Itemschwierigkeit) β geschätzt. Es wird angenommen, dass die beobachteten Werte y nur von den Parametern θ und β abhängen. Es gilt

$$y_{jk} \sim \text{Bernoulli}(\text{logit}^{-1}(\theta_j - \beta_k))$$

wobei $j = 1, \dots, J$ der Index über die Personen und $k = 1, \dots, K$ der Index der Items ist.

Für die Analyse der vorliegenden Daten wurde aufgrund von Modellvergleichen ein 2-PL Modell mit einem Interaktionsterm verwendet, in dem neben den oben beschriebenen Parameter ein zusätzlicher Diskriminanzparameter α in das Modell mit aufgenommen wurde.

$$y_{jk} \sim \text{Bernoulli}(\text{logit}^{-1}(\alpha_k \theta_j - \beta_k))$$

Das nachfolgende IRT Modell sind als bayesianisches Modell in Stan (Carpenter et al., 2017) formuliert. Hierfür müssen für alle Modellparameter Priorverteilungen festgelegt werden. In der vorliegenden Studie sind die Personenfähigkeiten von Interesse, daher werden zusätzlich die Itemschwierigkeiten mit

$$\beta_k \sim \text{Normal}(0,1)$$

normiert um das Modell zu identifizieren (Bafumi, Gelman, Park & Kaplan, 2005).

Für die Personenfähigkeiten gilt im vorliegenden Falle

$$\theta_j \sim \text{Normal}(\eta_\theta, \sigma_\theta^2)$$

wobei die Parameter $\eta_\theta, \sigma_\theta^2$ beide aus den Daten geschätzt werden. Für den Diskriminanzparameter gilt

$$\log \alpha_k \sim \text{Normal}(0.5,1)$$

Um den Einfluss der weiterer Variablen auf die Personenfähigkeit zu modellieren, wurde zusätzlich eine latente Regression geschätzt. Sowohl für die Regressionsparameter, als auch für die hierarchischen Varianzparameter der latenten Regression wurden Standardnormalverteilungen als Priorverteilungen verwendet.

Um die hierarchische Datenstruktur zu berücksichtigen, wird ein Mehrebenenmodell mit 3 Ebenen (Personen, Standorte der Schulen und Bundesländer) verwendet. In der latenten Regression werden neben den varying intercepts auf jeder Hierarchieebene und hauptsächlich interessierenden Merkmalen (Geschlecht, Selbstkonzept, Selbstwirksamkeit und Geschlechtsstereotype) auch weitere Variablen wie bspw. die Muttersprache und das Engagement mit einbezogen.

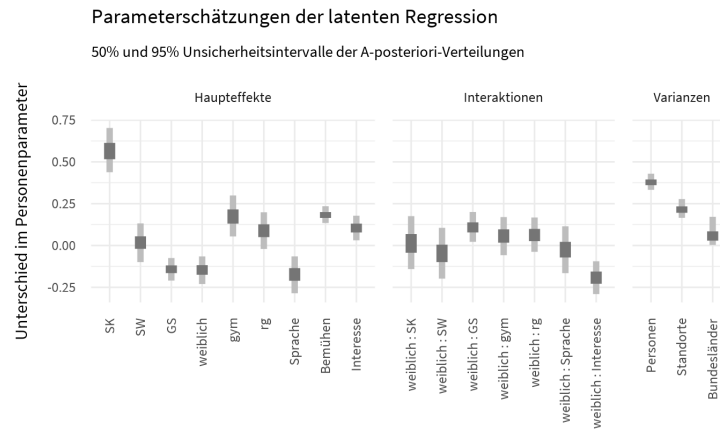


Abb. 4 Parameterschätzung der latenten Regression

Es ist zu erkennen, dass das Selbstkonzept den größten Effekt (der erhobenen Persönlichkeitsmerkmale) auf die Personenfähigkeit aufweist. Man würde also erwarten, dass sich Personen mit hohem und niedrigem Selbstkonzept um etwa 1.13 (95% CI: 0.88, 1.41) in ihrer geschätzten Personenfähigkeit unterscheiden. Darüber hinaus ist kein inhaltlich bedeutsamer Effekt für Selbstwirksamkeit zu beobachten. Ein Effekt in entgegengesetzte Richtung ist für die Geschlechtsstereotype zu finden: Personen mit stark ausgeprägten Geschlechtsstereotypen haben demnach eine geringere geschätzte Personenfähigkeit als Personen mit geringer ausgeprägten Geschlechtsstereotypen. Interessant ist, dass dieser Effekt lediglich für Burschen besteht und für Mädchen durch den Interaktionsterm negiert wird.

3.2.1 Selbstkonzept

Von den inkludierten Prädiktoren weist der Faktorscore des akademischen Selbstkonzepts den größten Effekt auf die Personenfähigkeit auf (siehe auch Abbildung 4). Zur besseren Veranschaulichung dieses Effekts wurden 200 Datensätze aus der Posteriorverteilung gezogen und die mittlere Anzahl gelöster Aufgaben in Abhängigkeit des Selbstkonzepts bestimmt. Das Modell würde etwa für Personen, deren Selbstkonzept im unteren Drittel liegt (vgl. in Abbildung 5 die Ausprägungen (-2, -1] und (+1, +2]), eine durchschnittliche Punktezah von 5.3 (SD = 0.2) vorhersagen. Personen, die ein Selbstkonzept im oberen Drittel berichten, würden laut Modellvorhersage im Mittel 11.1 (SD = 0.2) Aufgaben lösen (vgl. Abbildung 5). Es zeigt sich also auch auf manifester Ebene ein beträchtlicher Zusammenhang zwischen dem Selbstkonzept und der erreichten Punkteanzahl in der Feldtestung.

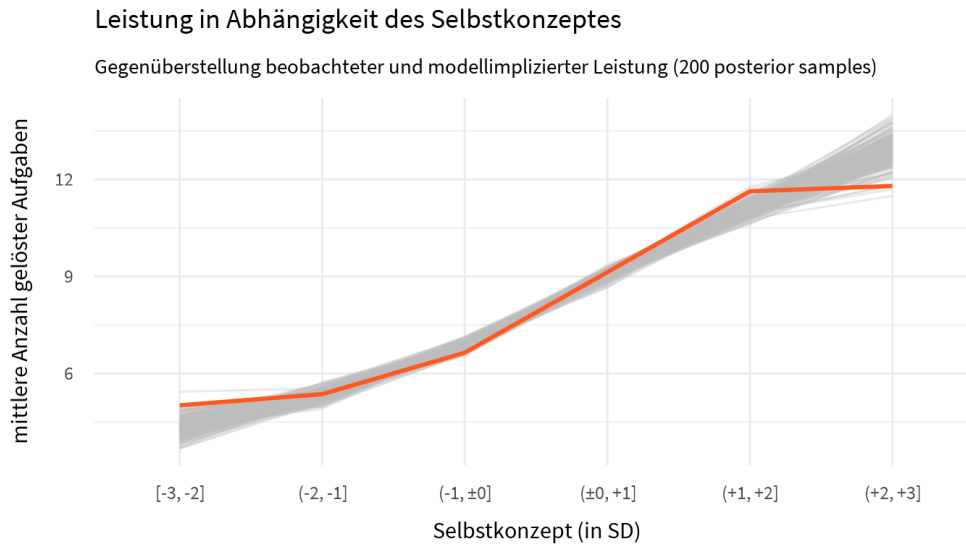


Abb. 5 Darstellung der mittleren Anzahl gelöster Aufgaben in Abhängigkeit der Ausprägung des Selbstkonzeptes (in SD)

3.2.2 Geschlechtsunterschiede

Die beobachtete mittlere Differenz gelöster Aufgaben zwischen Burschen und Mädchen liegt bei etwa 1.4 zu Gunsten der Burschen, während die durch das Modell vorhergesagte Differenz auf $M = 1.35$ ($SD = 0.17$) geschätzt wird. Dementsprechend kann davon ausgegangen werden, dass das Modell die Geschlechtsdifferenz in der Feldtestungsleistung adäquat abbildet (Abbildung 6 hellgraues Histogramm).

Zu beachten ist, dass die vorhergesagte Leistungsdifferenz mit Unterschieden im Selbstkonzept konfundiert ist. So wird etwa bei Mädchen ein durchschnittliches Selbstkonzept von -0.08 ($SD = 0.50$) beobachtet, während Burschen ein durchschnittliches Selbstkonzept von 0.12 ($SD = 0.48$) berichten. Es ist daher anzunehmen, dass die Unterschiede im Selbstkonzept (bzw. Persönlichkeitsmerkmalen allgemein) einen nicht unwesentlichen Anteil der Punktedifferenz ausmachen.

Aus diesem Grund wird in einer Simulation untersucht, welche Leistungsdifferenz man erwarten würde, wenn sich Burschen und Mädchen hinsichtlich des Selbstkonzeptes, der Selbstwirksamkeit und der Geschlechtsstereotype nicht systematisch unterscheiden. Eine unter diesen Bedingungen simulierte Leistungsdifferenz wird durch das Modell auf $M = 0.76$ ($SD = 0.20$) geschätzt. Der Effekt reduziert sich also auf etwa die Hälfte des ursprünglich beobachteten Effekts und läge mit hoher Wahrscheinlichkeit bei einer Differenz von weniger als 1 von 21 vorgegebenen Aufgaben (vgl. Abbildung 6 dunkelgraues Histogramm).

Geschlechterspezifische Leistungsdifferenz

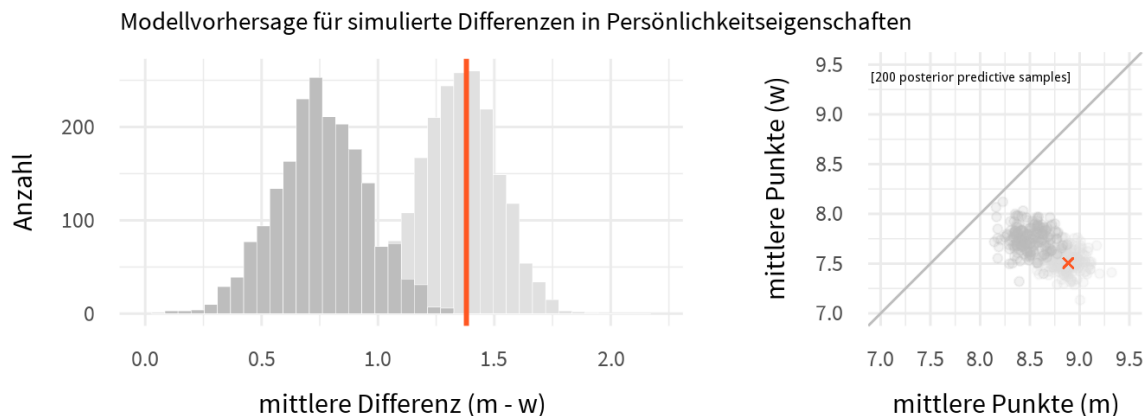


Abb. 6. Darstellung der Geschlechtsdifferenz. Hellgraue Punkte (bzw. Histogramm) beziehen sich auf die Datensimulation mit bestehenden Unterschieden in den Persönlichkeitsvariablen. Dunkelgraue Punkte (bzw. Histogramm) beziehen sich auch die Simulation neuer Daten unter der Annahme keiner systematischen Unterschiede in den berücksichtigten Persönlichkeitsmerkmalen.

4. Zusammenfassung und Ergebnisse

Allgemein beeinflusst das Selbstkonzept einer Person ihre Lernleistung bzw. Leistungsentwicklung. In der oben beschriebenen Studie konnte das Selbstkonzept als Einflussfaktor auf die Schülerleistungen identifiziert werden; von allen berücksichtigten Parametern hat das (akademische) Selbstkonzept den größten Einfluss auf die Mathematikleistung. Burschen berichten dabei ein höheres Selbstkonzept und ein höheres Interesse in Bezug auf mathematische Inhalte als Mädchen. Zudem konnte gezeigt werden, dass Personen mit einem hohen Ausmaß an Geschlechterstereotypen im Vergleich zu Personen mit einem geringen Ausmaß an Geschlechterstereotypen geringere Personenfähigkeit aufweisen.

Zur Leistungsvorhersage erweist sich also das Selbstkonzept als besserer Prädiktor als das Geschlecht oder das Interesse. Da sich das Selbstkonzept einer Person schon im Volksschulalter ausdifferenziert, ist es von großer Bedeutung dieses schon von Kindheitsjahren an bewusst zu stärken. Dadurch steigt die Anstrengungsbereitschaft und ein höheres kognitives Engagement während des Unterrichts kann erreicht werden, entsprechend höhere Leistungen erfolgen (vgl. Helmke & van Aken, 1995; Kessels, 2002).

Da Mädchen ein schlechteres akademisches Selbstkonzept berichten als Burschen ist es für Lehrpersonen wichtig [in ihrer Ausbildung] einen bewussten Umgang mit der Geschlechterthematik zu erlernen, da sie sich ihrer Überzeugungen hinsichtlich Pädagogik, Curriculum und ganzer Schulstrategien bewusst werden (Younger & Warrington, 2008) sowie geschlechterdifferentielles Lehr- und Kommunikationsverhalten erkennen lernen (Ludwig & Ludwig 2007).

Von Ow und Husfeldt (2011) beschreiben mehrere Maßnahmen auf verschiedenen Ebenen, um das Selbstkonzept von Schülerinnen und Schülern bewusst zu stärken. Durch eine Stärkung des Selbstkonzepts steigt die Anstrengungsbereitschaft und ein höheres kognitives Engagement während des Unterrichts kann erreicht werden, daraus können entsprechend höhere Leistungen erfolgen. „Empfohlen werden hierzu inhaltlich-didaktische bzw. sozial-kognitive Interventionsansätze“ (von Ow & Husfeldt 2011). Attributionstrainings sowie die längerfristige Konfrontation mit adäquaten Rollenmodellen sind weitere Ansätze. Zudem schlagen von Ow und Husfeldt (2011, S. 47) die Methode der Reattribution vor: „Das heisst, die Lehrpersonen sollen auf die sichtbaren Fähigkeiten der Schülerinnen und Schüler verweisen und Rückmeldungen so verfassen, dass sie das Selbstvertrauen der Mädchen aufbauen. Ziel ist dabei eine langfristige Verbesserung der selbstbezogenen Kognition der Mädchen.“ (von Ow & Husfeldt, 2011, S. 47f) Diese Methode funktioniert allerdings eher bei Schülerinnen, da Schüler weniger von äußeren Informationen beeinflusst werden. Im organisatorischen Bereich können Unterrichtssituationen geschaffen werden, in denen die Geschlechteridentität der Lernenden möglichst wenig aktiviert wird (z.B. monoedukative Phasen).

Für die Auswertung der Ergebnisse von 2018 und 2019 ist eine Modellvalidierung geplant. Hierzu werden aufgrund der vorliegenden Analysen (vgl. obige Ausführung) und den Angaben im Feldtestungsfragebogen 2018 die Leistungsunterschiede zwischen Burschen und Mädchen bzw. aufgrund des Unterschiedes in der Angabe des (akademischen) Selbstkonzepts in Mathematik vorhergesagt. Für die Feldtestung 2020 ist eine Überarbeitung des Fragebogens geplant um weniger trennscharfe Dimensionen zu ersetzen.

Literatur

- Bafumi, J., Gelman, A., Park, D. K., & Kaplan, N. (2005). Practical issues in implementing and understanding Bayesian ideal point estimation. *Political Analysis*, 13(2), 171-187.
- Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.
- BiSta (2017). Bundesergebnisbericht. Standardüberprüfung 2017 – Mathematik 8. Schulstufe. Abzurufen unter https://www.bifie.at/wpcontent/uploads/2018/02/BiSt_UE_M8_2017_Bundesergebnisbericht.pdf [02.01.2019]
- Bob Carpenter, Andrew Gelman, Matthew D. Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus Brubaker, Jiqiang Guo, Peter Li, and Allen Riddell. 2017. Stan: A probabilistic programming language. *Journal of Statistical Software* 76(1). DOI 10.18637/jss.v076.i01
- Faulstich-Wieland, H. (2008). Schule und Geschlecht. In: Helsper, W.; Böhme, J. (Hrsg.) *Handbuch der Schulforschung*. 673-695. Springer.
- Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2), 117-140.
- Frome, P. M.; Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, 74, 435-452.
- Hellmich, F. (Hrsg.) (2011). *Selbstkonzepte im Grundschulalter. Modelle, empirische Ergebnisse, pädagogische Konsequenzen*. Stuttgart: Kohlhammer.
- Helmke, A., van Aken, M.A.G. (1995). The causal ordering of academic achievement and self-concept of ability during elementary school: A longitudinal study. *Journal of Educational Psychology*, 87, 624-637.
- Helmke, A. (1998). Vom Optimisten zum Realisten? Zur Entwicklung des Fähigkeitskonzeptes vom Kindergarten bis zur 6. Klassenstufe. In *Entwicklung im Kindesalter* (pp. 115-132). Psychologie Verlags Union.
- Kessels, U. (2002). *Undoing gender in der Schule: eine empirische Studie über Koedukation und Geschlechtsidentität im Physikunterricht*. Beltz Juventa.
- Köller, O., & Klieme, E. (2000). Geschlechtsdifferenzen in den mathematisch-naturwissenschaftlichen Leistungen. In *TIMSS/III: Dritte Internationale Mathematik-und Naturwissenschaftstudie. Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn*(pp. 373-404). Leske+ Budrich.
- Köller, O.; Schnabel, K.; Baumert, J. (2000). Der Einfluss der Leistungsstärke von Schulen auf das fachspezifische Selbstkonzept der Begabung und das Interesse. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 32, 70-80.
- Lütke, O.; Köller, O.; Ardelt, C.; Stanat, P.; Baumert, J. (2006) Eine Überprüfung zur Genese akademischer Selbstkonzepte: Ergebnisse aus der PISA-Studie. In: *Zeitschrift für Pädagogische Psychologie*, Vol. 16, Issue 3/4.
- Ludwig, P. H., & Ludwig, H. (2007). Disparitäten bei Erfolgserwartungen zwischen Schülerinnen und Schülern. Eine Einführung. *Erwartungen in himmelblau und rosarot. Effekte, Determinanten und Konsequenzen von Geschlechterdifferenzen in der Schule*. Weinheim/München: Juventa, 7-15.
- Mair, P., Hatzinger, R., & Maier M. J. (2015). *eRm: Extended Rasch Modeling*. 0.15-6. <http://erm.r-forge.r-project.org/>
- Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal*, 23(1), 129-149.
- Marsh, H. W.; Hattie, J. (1996). Theoretical perspectives on the structure of self-concept. In: Bracken, B.A. (Ed.) *Handbook of self-concept*, 38-90. New York: Wiley.
- Moschner, B. (2001). Selbstkonzept. In: Rost, D.H. (Hrsg.) *Handwörterbuch Pädagogische Psychologie*, 629-634. Weinheim: Beltz.
- Neuwirth, E. (2015). Was die Ergebnisse der Zentralmatura wirklich zeigen. Abgerufen am 21.12.2017, von <http://www.neuwirth.priv.at/forblogs/BLERgebnisse.html>
- Nicholls, J. G. (1978). The development of the concepts of effort and ability, perception of academic attainment, and the understanding that difficult tasks require more ability. *Child development*, 800-814.

- Rosseel, Y. (2012). *lavaan: An R Package for Structural Equation Modeling*. Journal of Statistical Software, 48(2), 1-36. <http://www.jstatsoft.org/v48/i02/>.
- Schilling, S.; Sparfeldt, J.; Rost, D. (2006). Facetten schulischen Selbstkonzepts. Welchen Unterschied macht das Geschlecht? *Zeitschrift für Pädagogische Psychologie*, 20 (1/2), 9-18.
- Shavelson, R., Hubner, J. & Stanton, G. (1976). Self-Concept: Validation of construct interpretations. *Review of Educational Research*, 46 (3), 407–441.
- Suchań, B., & Breit, S. (2016). PISA 2015. *Grundkompetenzen am Ende der Pflichtschulzeit im internationalen Vergleich*. Abzurufen unter: https://www.bifie.at/wp-content/uploads/2017/11/PISA15_Kollaboratives-Problemlösen_final.pdf [10.01.2019]
- von Ow, A., Husfeldt V. 2011. *Geschlechterdifferenzen und schulische Leistungen – Eine Übersicht zum Forschungsstand*. Aarau.
- Watt, H.M.G.; Eccles, J.S. (Hrsg.) (2008). *Gender and Occupational Outcomes: Longitudinal assessments of individual, social, and cultural influences*. Washington, D.C.: APA books.
- Wild, E., & Möller, J. (2009). *Pädagogische Psychologie*. Heidelberg: Springer.

Verfasserinnen und Verfasser

Eva Sattlberger
 Bundesministerium für Bildung, Wissenschaft und Forschung
 Sektion III – Bildungsentwicklung und Bildungsmonitoring
 Minoritenplatz 5
 1010 Wien
eva.sattlberger@bmbwf.gv.at

Jan Steinfeld
 Bundesministerium für Bildung, Wissenschaft und Forschung
 Sektion III – Bildungsentwicklung und Bildungsmonitoring
 Minoritenplatz 5
 1010 Wien
jan.steinfeld@bmbwf.gv.at

Philipp Gewessler
 Bundesministerium für Bildung, Wissenschaft und Forschung
 Sektion III – Bildungsentwicklung und Bildungsmonitoring
 Minoritenplatz 5
 1010 Wien
philipp.gewessler@bmbwf.gv.at