

Anna-Katharina Pelkner/Klaus Boehnke

Streber als Leistungsverweigerer?

Projektidee und erstes Datenmaterial einer Studie zu mathematischen Schulleistungen

Zusammenfassung

Wer Schule in Deutschland und in Nordamerika vergleichend beobachtet, dem fallen bald für den Alltag von Schülerinnen und Schülern relevante Unterschiede auf, die in der erziehungswissenschaftlichen Forschung bisher kaum bearbeitet wurden: Hierzu gehört – neben der unterschiedlichen Bewertung des in Nordamerika verpönten, in Deutschland aber fast zum guten Ton gehörenden 'cheating'/Schummelns – die Bewertung guter Schulleistungen durch Peers. Während gute Leistungen in Nordamerika unter Peers hoch im Kurs zu stehen scheinen, werden sie in Deutschland – insbesondere in der Sekundarstufe I – oft mit dem Strebervorwurf belegt. Der hier vorgelegte Text wendet sich der Frage zu, ob der Strebervorwurf Auswirkungen auf tatsächliche Schulleistungen im Bereich der Mathematik hat. Er vertritt die These, dass befürchtete negative Peer-Sanktionen als Folge besonders guter Leistungen in Mathematik bei guten Schülern und insbesondere bei leistungsstarken Mädchen die Ausschöpfung des vorhandenen Leistungspotentials begrenzt und auf Dauer auch die Leistungsfähigkeit reduziert. Vorgestellt wird zum einen die Forschungsidee eines im DFG-Schwerpunktprogramm „Bildungsqualität von Schule“ geförderten Projekts, zum anderen erstes empirisches Material aus zwei Vorerhebungen des Projekts, mit dem die empirische Plausibilität der Forschungsidee untermauert werden soll.

Summary

Do Nerds Refuse Achievement? Project idea and first data from a study on mathematical achievement

Whoever compares schooling in Germany and North America, will soon come across differences with everyday relevance for the school life of students that have rarely been addressed in educational research: Besides the different attitudes towards cheating – being a complete taboo in North America, but a lesser sin in Germany – there is the attitude towards high achievement by peers. Whereas in North America high achievement seems highly valued, high-achieving students in Germany are often accused of being nerds or teacher's pets, particularly in grades 7-10. The present article deals with the extent to which the nerd accusation has consequences for real achievement in mathematics. The hypothesis is formulated that fear of negative sanctions by peers, as a consequence of being particularly good in math, encourages high achievers to reduce their efforts. Particularly among high-achieving girls, this is presumed to even affect the objective achievement potential in math in the long run. Firstly, the research idea for a study funded by the German Research Council (DFG) in the framework of its focal program on "Quality of Education in Schools" is presented. Secondly, first empirical material from two pilot studies is presented in order to underscore the plausibility of the research idea.

1 Einleitung

Mit der Einrichtung des Schwerpunktprogramms „Bildungsqualität von Schule: Fachliches und fächerübergreifendes Lernen im mathematisch-naturwissenschaftlichen Unterricht in Abhängigkeit von schulischen und außerschulischen Kontexten (BIQUA)“ hat die Deutsche Forschungsgemeinschaft (DFG) auf die u.a. durch die TIMS-Studie ausgelöste Diskussion um Bildungsqualität in deutschen Schulen, speziell im mathematisch-naturwissenschaftlichen Unterricht reagiert. Die in Deutschland maßgeblich von Jürgen BAUMERT geleitete Studie hatte erbracht, dass deutsche Schülerinnen und Schüler im internationalen Vergleich in ihren mathematisch-naturwissenschaftlichen Schulleistungen *bestenfalls* einen Mittelplatz unter den an der Studie teilnehmenden Ländern belegen und dass auch die Teilgruppe der deutschen Gymnasialschülerinnen und -schüler weit davon entfernt ist, einen internationalen Spitzenplatz zu erreichen (vgl. BAUMERT u.a. 1997; BAUMERT/BOS/LEHMANN 2000).

Während sich eine Reihe der im Schwerpunktprogramm geförderten Projekte mit Aspekten der Didaktik des mathematisch-naturwissenschaftlichen Unterrichts und mit curricularen Fragen beschäftigen und so Möglichkeiten der Optimierung des Unterrichts speziell in diesen Fächern erforschen, geht es im Projekt „Streber vs. Nerd“ um die Frage, in welchem Zusammenhang Leistungsbewertung durch Peers und mathematische Schulleistungen stehen.

Zentrale These ist dabei, dass gute Schulleistungen in Deutschland im Vergleich zu Nordamerika einem erheblichen Peerdruck unterliegen, oder anders formuliert: dass (sehr) gute deutsche Schülerinnen und Schüler in viel stärkerem Maße als ihre nordamerikanischen Altersgenossen befürchten müssen, wegen eben dieser guten Schulleistungen bei ihren gleichaltrigen Mitschülern und Freunden schlecht angesehen zu sein, als *Streber* diffamiert zu werden. Dieses Erleben, so die weiter im Projekt zu überprüfende These, senkt die Bereitschaft insbesondere begabter Schülerinnen und Schüler der Sekundarstufe I, ihre eigentlich vorhandene Leistungsfähigkeit auszuschöpfen. Zu dieser These wird im Folgenden einerseits ein kurzer – mangels durchgehender Forschungstraditionen im Mosaikartigen verbleibender – Literaturüberblick gegeben, zum anderen wird empirisches Datenmaterial aus zwei am Standort des Projekts in Chemnitz durchgeführten Pilotstudien des vom Zweitautor beantragten und von der Erstautorin durchgeführten Projekts vorgestellt.

2 Der Zusammenhang von mathematischen Schulleistungen und Peerklima im Literaturüberblick

Zu der Frage, ob es einen Zusammenhang zwischen Schulleistungen im mathematischen Leistungsbereich und dem Peerklima in einer Schulklasse gibt, liegen nur wenige kulturvergleichende Studien vor. Eine Ausnahme bilden einige Arbeiten aus der Forschergruppe um Herbert WALBERG. WALBERG, SINGH und TSAI (vgl. 1984) zeigen für indische Schülerinnen und Schüler, dass Peergruppenstatus und Schulleistung nach Ausparialisierung anderer Variablen in einem positiven Zusammenhang stehen. UGUROGLU und WALBERG (vgl. 1986) zeigen für den nordamerikanischen Raum, dass ein (positiver) Zusammenhang zwischen einem unterstützenden Peer-Kontext und Mathematikleistungen

besteht. Die beiden Studien verfolgen in ihrer Anlage allerdings eine andere Kausalitätsannahme als das hier vorgestellte Projekt. Sie gehen davon aus, dass ein hoher Peergruppenstatus gute (Mathematik-)Leistungen *zur Folge* hat. Hier wird hingegen der These nachgegangen, dass (sehr) gute Leistungen im Fach Mathematik in Deutschland zu negativen Peer-Sanktionen *führen*, was im Weiteren dann antizipativ zu einer Reduktion schulischer Performanz führt, während dies tendenziell in Nordamerika nicht der Fall ist.

JUVONEN/MURDOCK (vgl. 1995) arbeiten eine mögliche Altersspezifik negativer Zusammenhänge zwischen guten Schulleistungen und Peer-Status heraus, wenn sie zeigen, dass insbesondere Achtklässler (im Vergleich zu Viert- und Sechstklässlern) sehr zurückhaltend damit sind, ihren Peers mitzuteilen, dass sie 'effortful' sind, d.h., dass sie viel Arbeit in gute Schulleistungen investieren.

LANDSHEER u.a. (vgl. 1998) wenden sich der Frage von Zusammenhängen zwischen Schulleistungen und sozialer Kompetenz zu und fassen zunächst den Forschungsstand dahingehend zusammenfassen, dass beide Bereiche unkorreliert seien. Ihre eigenen empirischen Studien an einer niederländischen Stichprobe erbringen dann allerdings ein anderes, für die Autoren überraschendes Bild: Akademische und soziale Kompetenz erweisen sich als negativ korreliert. Auch wenn soziale Kompetenz und Peer-Akzeptanz keine deckungsgleichen Konzepte sind, so kann man deren negative Korreliertheit in einer mitteleuropäischen Stichprobe u.U. doch als Indiz für die im hier vorgestellten Projekt vertretene These werten, dass (sehr) gute Schulleistungen in einem nicht-angelsächsischen Kulturraum negativ mit Peer-Akzeptanz korreliert sind.

Informelle Gespräche mit Jugendlichen in Deutschland machen deutlich, dass Streber eine der gefürchtetsten Etikettierungen im Peergruppenkontakt ist. Dies gilt insbesondere im Altersbereich der Sekundarstufe I. Wer in deutschen Schulen zu häufig durch besonders gute Leistungen auffällt und wer womöglich auch noch offenbar werden lässt, dass diese Schulleistungen Ergebnis intensiver Lernbemühungen und nicht Produkt einer ‚natürlichen Begabung‘ sind, wird offenbar schnell damit konfrontiert, von Mitschülerinnen und Mitschülern, die i.d.R. in der Sekundarstufe I die relevanteste Peergruppe darstellen, als Streber benannt und tendenziell ausgegrenzt zu werden. Wissen vorwiegend anekdotischer Natur aus dem nordamerikanischen Kulturkreis legt nahe, dass es sich hierbei um ein Kulturspezifikum handelt. Es spricht einiges für die These, dass die negative Sanktionierung von guten Schulleistungen unter Peers für Deutschland typisch ist, während im nordamerikanischen Kontext Schulleistungen unter Peers im Vergleich zu Deutschland positiver bewertet werden. Hierfür spricht schon die Tatsache, dass eine wörtliche Übersetzung von *Streber* ins Englische schwer fällt. Im Oxford-Duden (vgl. 1990) findet sich das Wort ‚swot‘; diese Bezeichnung taucht im englischen Thesaurus von WinWord 2000 nicht auf, es handelt sich also offensichtlich um ein recht ungebräuchliches Wort. Im deutschen Thesaurus werden hingegen für Streber gleich noch eine Reihe von Synonymen (Ehrgeizling, Karrierist, Musterschüler, Opportunist) aufgeführt. Eine weitere in Frage kommende und für die Pilotstudie des DFG-Projekts gewählte Übersetzung, der ‚nurd‘ bzw. ‚nerd‘, hat deutlich andere – zum Teil liebevolle – Konnotationen, die eher dem deutschen Gebrauch des Anglizismus ‚Freak‘ entsprechen (Computerfreak, Mathe-Freak). ‚Overachiever‘ oder ‚teacher’s pet‘ korrespondieren schon eher mit dem deutschen Begriff des Strebers, sind aber auch nicht übermäßig verbreitet. Der übliche Gebrauch des Begriffes ‚nurd‘ legt nahe, dass außergewöhnliche Leistungen zwar den Beigeschmack von Absonderlichkeiten haben können, zumindest in Teilbereichen in Nordamerika unter Peers aber positiv bewertet werden.

Bildungspolitisch ist dies hochgradig relevant. Wenn deutsche Schülerinnen und Schüler der Sekundarstufe I eine massive (negative) soziale Sanktionierung positiver Schulleistungen durch Peers befürchten müssen, so ist zu erwarten, dass hervorragende Leistungen im Sinne einer Selbstbescheidung insbesondere von begabten Schülern und dabei vor allem von Mädchen vermieden werden. Die Güte der Schulleistungen – speziell auch im internationalen Vergleich – wäre dann weniger Resultat von curricularen, didaktischen, schulstrukturellen oder auf Schülerinnen und Schüler, Lehrpersonen oder Eltern bezogenen personalen Bedingungen, sondern vielmehr Ergebnis des leistungsbezogenen Peer-Klimas. Zugespitzt formuliert: Deutsche Schülerinnen und Schüler der Sekundarstufe I vermeiden besonders gute Schulleistungen, um sich die Anerkennung der Peergruppe zu erhalten, nordamerikanische Schülerinnen und Schüler dagegen suchen besonders gute Schulleistungen, um die Anerkennung der Peers zu gewinnen.

Gelegentlich wird in Diskussionen um das Thema davon gesprochen, die ‚Kulturrevolution‘ der Studentenbewegung mit ihrer Ablehnung des Leistungsgedankens sei ‚schuld‘ an der Negativbewertung von Leistung unter deutschen Jugendlichen. Der ‚Streber‘ hat jedoch in Deutschland Geschichte und ist keineswegs erst seit ‘68 *en vogue*.¹ Er hat in Westdeutschland ebenso wie in Ostdeutschland, wo ja die Studentenbewegung nicht oder nur in sehr viel geringerem Maß wirksam war, die gleiche negative Bedeutung in jugendlichen Peergruppen.

Für erste Antworten auf die Frage nach den Gründen für Kulturunterschiede bietet sich der wertetheoretische Ansatz von Shalom SCHWARTZ (vgl. 1992) an. Das SCHWARTZsche Konzept kennt auf der Individualebene zehn motivationale Typen von Werthaltungen, nämlich Hedonismus, Stimulation, Selbstbestimmung, Universalismus, Humanismus, Tradition, Konformität, Sicherheit, Macht und Leistung. Besonders in der Bewertung von Leistung als Wert unterscheiden sich nach dem von SCHWARTZ vorgelegten Datenmaterial deutsche und nordamerikanische Individuen (vgl. SCHWARTZ 1992). Aber nicht nur auf der individuellen Ebene gibt es Unterschiede. Auch in soziologischen Analysen, die sich auf Kulturen beziehen, zeigt SCHWARTZ deutliche Unterschiede zwischen typischen deutschen und typischen nordamerikanischen Werten auf. Er unterscheidet hier sieben kulturelle Wertetypen, nämlich Mastery (Erfolgsorientierung), Hierarchy (Machtorientierung), Embeddedness (kollektive Eingebundenheit), Harmony (Harmonie), Egalitarianism (Egalitarismus), Intellectual Autonomy (intellektuelle Autonomie) und Affective Autonomy (affektive Autonomie). SCHWARTZ (vgl. 1994) zeigt, dass Nordamerika stärker durch Erfolgs- und Machtorientierung, Deutschland stärker durch ein Klima des Egalitarismus und der intellektuellen Autonomie geprägt ist, was sich – etwas gewagt – durchaus auf den Widerspruch von Leistung und Gleichheit zuspitzen lässt. Dieser Widerspruch, so die zweite zentrale These des Projekts, ist letztlich die Ursache für die unterschiedliche Sanktionierung von Leistung in Deutschland und Nordamerika und speziell für die Schülerinnen und Schüler der Sekundarstufe I, für die Peergruppenintegration eine zentrale Entwicklungsaufgabe ist (vgl. NOACK 1990).

Von besonderem Interesse in diesem Kontext ist, wie sich die Peergruppenintegration für Mädchen und Jungen gestaltet und welchen Stellenwert die Sanktionierung durch Peers jeweils einnimmt. In der Weiterführung des SCHWARTZschen Ansatzes liegt es deshalb nahe zu prüfen, inwiefern sich die Wertschätzung guter Schulleistungen bei Mädchen und Jungen bzw. in Stereotypen über Mädchen und Jungen unterscheidet.

Neben der Kulturspezifität die Geschlechterproblematik in den Mittelpunkt eines Projekts zu außerunterrichtlichen Einflüssen auf mathematisch-naturwissenschaftliche Schul-

leistungen zu rücken, wird auch durch die Ergebnisse der TIMS-Studie nahegelegt, die gezeigt hat, dass es sowohl im Bereich der mathematisch-naturwissenschaftlichen Grundbildung als auch im gymnasialen Mathematik- und Physikunterricht substantielle Leistungsunterschiede zwischen Schülerinnen und Schülern gibt (BAUMERT u.a. 1997; BAUMERT/BOS/LEHMANN 2000).

Ziel unserer Untersuchung ist es, diese – wie wir meinen nur vordergründigen – Leistungsunterschiede zwischen Mädchen und Jungen genauer zu untersuchen und aus dem Blickwinkel der Leistungsakzeptanz unter Peers der Frage nachzugehen, *warum* Mädchen schlechter in Mathematik abschneiden als Jungen.

Die Bedeutung von geschlechtsspezifischer Interaktion im Schulalltag für die schulische Entwicklung von Mädchen ist zentral für die seit Jahren geführte Debatte um Koedukation. In Frage gestellt wird, ob Koedukation als „Ansatz zum Aufbrechen von Geschlechtsstereotypen“ (FAULSTICH-WIELAND 1993) geeignet ist oder ob sie nicht vielmehr bestehende gesellschaftliche Ungleichheiten – wie die Geschlechterhierarchie – reproduziert (vgl. PECHTL 1994). Koedukativer Unterricht, so haben verschiedene Untersuchungen gezeigt, wirkt sich keineswegs positiver auf das leistungsbezogene Selbstbewusstsein von Mädchen aus als ein getrenntgeschlechtlicher (vgl. u.a. ENDERS-DRAGÄSER/FUCHS 1988; FRASCH/WAGNER 1982; dagegen kritisch u.a. HAGEMEISTER 1991). Die geschlechtsspezifischen Unterschiede in der schulischen Interaktion führen dazu, dass Jungen mit mehr Selbstvertrauen als Mädchen die Schule verlassen. Obwohl Mädchen formal die besseren Schulleistungen erbringen, weisen sie bereits gegen Ende der Grundschulzeit ein deutlich schlechteres Selbstkonzept in Mathematik auf als Jungen (vgl. TIEDEMANN/FABER 1995). Untersuchungen in amerikanischen High Schools und Colleges haben ergeben, dass stereotype Vorstellungen über die eigene Geschlechtsrolle unter Absolventinnen von Mädchenschulen geringer als bei Absolventinnen koedukativer Einrichtungen ausgeprägt sind (vgl. LEE/BRYK 1986). Für eine neuseeländische Stichprobe wurden problematische Auswirkungen koedukativen Unterrichts bereits Anfang der 1970er-Jahre von JONES und SHALLCRASS (vgl. 1972) belegt. Die Hoffnungen der Frauenbewegung, gleiche Bildung für Jungen und Mädchen könnte die gesellschaftliche Ungleichheit der Geschlechter abbauen, sind nicht in Erfüllung gegangen (vgl. RICHTER 1996). Für eine gerechte und gleichwertige Ausbildung reicht es offenbar nicht aus, dass Mädchen und Jungen im Unterricht die gleichen Mathematikaufgaben rechnen und die gleichen Klassenarbeiten schreiben (vgl. FENNEMA 1990).

Ein wesentlicher Untersuchungsgegenstand der geplanten Studie ist deshalb das geringe Selbstvertrauen von Mädchen in Bezug auf ihre eigene Leistungsfähigkeit. Die TIMS-Studie kommt zu dem Ergebnis, dass das fachspezifische Selbstvertrauen von Schülerinnen auch bei vergleichbarer Fachleistung geringer ist als das von Schülern, die ihrerseits dazu neigen, ihre Fähigkeiten zu überschätzen (BAUMERT u.a. 1997; BAUMERT/BOS/LEHMANN 2000).

Feministische Schulforscherinnen haben in einer Reihe von Beiträgen erörtert, welche Probleme sich für Jungen und Mädchen auf der Grundlage geschlechtsspezifischer Interaktionen und Bewertungshandlungen in der Schule ergeben können (vgl. KRUSE 1993; PFISTER 1993; NYSSSEN 1994, GLUMPLER 1995).

Die Koedukationsfrage lässt sich aber durchaus auch aus einer Peerkontext-Perspektive diskutieren. Im koedukativen mathematisch-naturwissenschaftlichen Unterricht lernen Mädchen (wie Jungen) zusammen mit gleich- *und* gegengeschlechtlichen Peers. Insbesondere die in diesem Lernbereich starken Mädchen sind dadurch nicht nur möglichen

negativen Peer-Sanktionen wegen ‚zu‘ überdurchschnittlicher Leistungen ausgesetzt, sie dürften oft zusätzlich mit Peer-Sanktionen zu rechnen haben, weil sie in einem ‚unweiblichen‘ Bereich gute Leistungen erbringen (vgl. HANNOVER 1999).

Der Erfahrungsraum Schule wird von Mädchen und Jungen unterschiedlich erlebt und wahrgenommen. Wie eine kanadische Studie zeigt, hängen Mädchen hinsichtlich ihrer Selbsteinschätzung in mathematisch-naturwissenschaftlichen Fächern stärker vom Urteil ihrer Mitschüler(innen) als von der Ermutigung durch ihre Lehrkräfte ab; außerdem sind Mädchen im mathematisch-naturwissenschaftlichen Unterricht in stärkerem Maße mit der Belästigung durch Jungen konfrontiert (vgl. MCLAREN/GASKELL 1995). Schule ist offenbar nur vordergründig ein Erfolgsterritorium der Mädchen. HORSTKEMPER (vgl. 1987) hat auf die Abhängigkeit des (geringeren) weiblichen Selbstwerts von schulischen Lernbedingungen hingewiesen. Der Selbstvertrauensabstand zu Lasten der Mädchen vergrößert sich sogar im Verlauf der Schulzeit – trotz ihrer durchweg besseren Schulleistungen. Ähnliche Entwicklungen sind für die ehemalige DDR zu konstatieren (vgl. KRAUSE 1990). Eine Ursache ist in der Ungleichbehandlung von Schülerinnen und Schülern zu suchen. ENDERS-DRAGÄSSER/FUCHS (vgl. 1988) zeigen, dass auch Lehrerinnen im Unterricht – bewusst oder unbewusst – dazu tendieren, den Schein männlicher Überlegenheit aufrecht zu erhalten und von Jungen oftmals nicht fordern, sich über tatsächliche Leistung zu qualifizieren (vgl. auch KREIENBAUM 1989). Gleichzeitig wird das Lernengagement der Mädchen als selbstverständlich vorausgesetzt und bleibt ohne positive Rückmeldung. Auf kritisches Lehrerfeedback reagieren Mädchen zudem sensibler als Jungen (vgl. MEYER/DICKHÄUSER 1998; DWECK u.a. 1978). Die Erwartungen und Sanktionen von Mitschülerinnen und Mitschülern sind sogar noch entscheidender für die Entwicklung des Selbstbewusstseins von Mädchen (vgl. MCLAREN/GASKELL 1995).

Insbesondere im Hinblick auf die mathematisch-naturwissenschaftlichen Fächer stellen seit den frühen 1970er-Jahren amerikanische Soziolog(inn)en und Schulpraktiker(innen) überall dieselben Tendenzen zur Ungleichheit fest: Um die Pubertät herum verlieren die Mädchen ihr Interesse an Mathematik (vgl. KOEPKE 1988). Lehrpersonen tendieren dazu, Mathematik als männliche Domäne zu stereotypisieren (vgl. LI 1999; RUSTEMEYER 1999). Spätestens in der Sekundarstufe I ist eine deutliche Distanz der Mädchen gegenüber dem mathematisch-naturwissenschaftlichen Unterricht zu beobachten (vgl. GÖTZE 1994). Ein zentraler Aspekt ist also die unterschiedliche Einschätzung der eigenen Leistungsfähigkeit bei Mädchen und Jungen.

Um erstes empirisches Material für die Plausibilität der Projektidee vorlegen zu können und zur Erprobung von Erhebungsinstrumenten wurden im Sommer 2000 in Chemnitz zwei Vorerhebungen zum Thema des Projekts durchgeführt. In der ersten Pilotstudie gingen wir mit der Hypothese ins Feld, dass sich zeigen lässt, dass viele Etiketten, mit denen sich Schülerinnen und Schüler der Sekundarstufe I gegenseitig belegen, ‚leistungskonnotiert‘ sind, und dass Ängste vor negativen Bewertungen durch Peers mit Mathematikleistungen korreliert sind. Als ‚leistungskonnotiert‘ bezeichnen wir Etikettierungen, deren Nutzung mit mathematischen Leistungen kovariiert. Gleichzeitig galt es, eine erste Prüfung der These vorzunehmen, dass Ängste vor dem Strebervorwurf bei Mädchen und Jungen unterschiedliche Konsequenzen haben. In der zweiten Pilotstudie ging es darum zu prüfen, ob das Verständnis dessen, was einen Streber ausmacht, sich zwischen Mädchen und Jungen unterscheidet. Wie oben ausgeführt, ist eine unserer zentralen Thesen, dass der Strebervorwurf bei Mädchen eine stärkere, langfristig leistungsmindernde Auswirkung hat. Dies könnte etwas mit Geschlechterstereotypen und mit einer geschlechts-

spezifischen Selbstkonzeptentwicklung zu tun haben. Es könnte seine Ursache aber auch darin haben, dass Mädchen und Jungen unter einem Streber etwas Unterschiedliches verstehen, dass das Konzept Streber also eine geschlechtsspezifische Substanz hat. Der nachfolgende Abschnitt schildert unser Vorgehen in den beiden Pilotstudien und verdeutlicht dabei auch noch einmal unsere Erwartungen.

3 Methode

In Pilotstudie 1 mit $N = 196$ Schülerinnen (51,8 %) und Schülern (48,2 %) im Durchschnittsalter von 14,8 Jahren wurde ein Fragebogen eingesetzt, in dem es um Stereotyping ging. Die Gewinnung der Stichprobe erfolgte durch Praktikanten², die als Schüler der 9. Klasse ein Betriebspraktikum an der Technischen Universität Chemnitz absolvierten. Die Praktikanten waren aufgefordert, für die Befragung selbst Mitschüler als Stichprobe zu akquirieren. Dateneingabe und Rohdatenauswertung waren ebenfalls Bestandteil des Praktikums.

Erfragt wurden Erfahrungen mit stereotypen Personenbezeichnungen, wie sie Schülerinnen und Schüler in der alltäglichen Kommunikation miteinander benutzen. Im Fragebogen wurden Aussagen dazu erbeten, mit welchen von 30 verschiedenen solchen stereotypen Personenbezeichnungen die Befragten schon *selbst belegt worden* sind, welche der 30 Stereotype sie schon für die Benennung anderer *benutzt haben* und in welchem Ausmaß sie *Angst haben*, dass andere Schülerinnen und Schüler die 30 stereotypen Personenbezeichnungen auf sie anwenden. Erfragt wurden Ratings zu folgenden Stereotypen: Streber, Genie, Außenseiter, Schleimer, Angeber, Meckerziege, Intelligenzbestie, Mathe-As, Leseratte, Hohlkopf, Spitzenschüler, Hirnie, Spinner, Labertasche, Stubenhocker, Besserwisser, Klugscheißer, neunmal klug, überschlaue, altklug, Strebersau, Alleswisser, Naturtalent, Wunderkind, superschlaue, vorlaut, strebsam, allwissend, Störenfried und oberflächlich. Ratings wurden jeweils in der Weise erbeten, dass zunächst gefragt wurde, „Bin ich schon mal genannt worden (häufig = 3, manchmal = 2, selten = 1, nie = 0) ...“, danach, „Habe ich schon jemanden aus meiner Klasse genannt (häufig, manchmal, selten, nie)...“, und zu guter Letzt, „Habe ich Angst, genannt zu werden (häufig, manchmal, selten, nie)...“. Im Fragebogen ebenfalls erfragt wurden: die Mathematiknote auf dem letzten Zeugnis, die Selbsteinschätzung der Leistungsfähigkeit in Mathematik („Welche Note würdest Du Dir selbst in Mathematik geben?“), die Selbsteinschätzung der allgemeinen Schulleistungen („Welche Note würdest Du Dir insgesamt als Schüler selbst geben?“). Dazu wurden zwei TIMSS-Aufgaben einbezogen.³

In Pilotstudie 2 wurden $N = 194$ Schülerinnen (55,2 %) und Schülern (44,8 %) im Durchschnittsalter von 14,9 Jahren befragt. Die Gewinnung der Stichprobe, Dateneingabe und Rohdatenauswertung erfolgte auch hier durch Praktikanten.⁴

Erbeten wurden Ratings zu insgesamt 51 Aussagen über Streber; Beispiel: „Streber wollen etwas Besseres sein als ihre Mitschüler“ (vgl. auch Tabelle 6). Die Zustimmung zu den Items wurde mit einer vierstufigen Ratingskala mit den Abstufungen ‚stimmt völlig‘ = 3, ‚stimmt ziemlich‘ = 2, ‚stimmt wenig‘ = 1, ‚stimmt gar nicht‘ = 0 erfragt.

4 Ergebnisse

Wie in der Einleitung ausgeführt, war es Ziel der ersten Pilotstudie herauszufinden, welche der im Fragebogen aufgeführten Stereotype ‚leistungskonnotiert‘ sind. Diese Frage sollte dabei nicht nur für Schülerinnen und Schüler allgemein, sondern auch für Mädchen und Jungen getrennt beantwortet werden.

In Tabelle 1 findet sich zunächst eine Übersicht über Mittelwerte zu den drei Fragemodi zu jedem der 30 Stereotype. Mittels einer multivariaten Varianzanalyse mit Messwiederholung wurde geprüft, ob sich die Mittelwertsvektoren der Antworten zu den drei Fragemodi (Messwiederholungsfaktor) und nach Geschlecht (Zwischensubjektfaktor) unterscheiden. In nachfolgenden univariaten Analysen wurde geprüft, ob sich die Bewertung der einzelnen Stereotype zwischen den drei Fragemodi signifikant ($p \leq 0,05$) unterscheidet. Dies war für alle Stereotype außer ‚altklug‘ der Fall.

Tabelle 1: Deskriptive Statistik

Mit Stereotyp benannt worden	Mittelwert	Stereotyp selbst benutzt	Mittelwert	Angst haben, genannt zu werden	Mittelwert
vorlaut	1,99	Streber	2,54	Außenseiter	1,75
Besserwisser	1,93	Schleimer	2,36	Angeber	1,67
Naturtalent	1,90	Angeber	2,33	Streber	1,63
Streber	1,77	Besserwisser	2,16	Schleimer	1,61
Klugscheißer	1,75	Spinner	2,13	Klugscheißer	1,58
Genie	1,74	Klugscheißer	2,10	Strebersau	1,55
Stubenhocker	1,72	Genie	2,02	Hohlkopf	1,53
Leseratte	1,68	Mathe-As	1,99	Störenfried	1,48
Spinner	1,68	Naturtalent	1,90	vorlaut	1,47
Störenfried	1,67	vorlaut	1,82	Spinner	1,47
Mathe-As	1,62	Störenfried	1,79	Besserwisser	1,47
Schleimer	1,56	Alleswisser	1,78	Stubenhocker	1,47
Angeber	1,54	Außenseiter	1,77	Meckerziege	1,43
Alleswisser	1,51	Meckerziege	1,75	Hirnie	1,39
neunmal klug	1,48	strebsam	1,70	Labertasche	1,37
allwissend	1,47	Stubenhocker	1,69	strebsam	1,36
Labertasche	1,47	Hohlkopf	1,68	neunmal klug	1,35
Spitzenschüler	1,44	Leseratte	1,67	Alleswisser	1,33
strebsam	1,42	Strebersau	1,67	oberschlau	1,31
Wunderkind	1,42	Hirnie	1,65	superschlau	1,31
Intelligenzbestie	1,41	allwissend	1,64	altklug	1,31
superschlau	1,40	oberschlau	1,63	allwissend	1,30
oberschlau	1,39	Spitzenschüler	1,62	Naturtalent	1,30
oberfleißig	1,35	neunmal klug	1,60	Genie	1,29
Meckerziege	1,34	oberfleißig	1,59	oberfleißig	1,29
Außenseiter	1,32	Labertasche	1,58	Intelligenzbestie	1,28
altklug	1,30	superschlau	1,57	Wunderkind	1,28
Hohlkopf	1,30	Intelligenzbestie	1,53	Leseratte	1,27
Hirnie	1,24	Wunderkind	1,49	Spitzenschüler	1,27
Strebersau	1,20	altklug	1,38	Mathe-As	1,23

Die Tabelle macht deutlich, dass die Mittelwerte der eigenen Nutzung der Stereotype am höchsten sind, gefolgt von den Mittelwerten für die Frage danach, ob man mit einem Stereotyp schon belegt worden ist. Die im Durchschnitt niedrigsten Mittelwerte gab es für die Frage, ob man Angst habe, in einer bestimmten Weise benannt zu werden. Die Mittelwertsunterschiede sind multivariat signifikant ($F_{60/82} = 9,06$, $p \leq 0,001$). Gleiches gilt für Geschlechtsunterschiede, die in der Tabelle nicht im einzelnen dokumentiert werden

($F_{30/112} = 3,05$, $p \leq 0,001$). Hier liegen Jungen in der Regel höher als Mädchen. Die Tatsache, dass dies für die Nutzung von Stereotypen und für Erfahrungen, mit den Stereotypen belegt zu werden, der Fall ist, nicht jedoch für die Angstwerte (hier gibt es weniger Geschlechtsunterschiede), kommt in einer signifikanten statistischen Interaktion von Geschlecht und Fragemodus zum Ausdruck ($F_{60/82} = 1,73$, $p = 0,05$).

Top-Plätze bei der eigenen Nutzung nehmen die Stereotype ‚Streber‘, ‚Schleimer‘ und ‚Angeber‘ ein. Dies sind also die Stereotype, mit denen Schülerinnen und Schüler ihre Klassenkameraden am häufigsten belegen. Die eigene Wahrnehmung dessen, als was man von Mitschülerinnen und Mitschülern am häufigsten bezeichnet wird, unterscheidet sich von dieser Reihung. Hier nehmen ‚vorlaut‘, ‚Besserwisser‘ und ‚Naturtalent‘ die ersten drei Plätze ein. Am ehesten angstbesetzt sind hingegen ‚Außenseiter‘, ‚Angeber‘ und ‚Streber‘. Schaut man sich die Stereotype an, die hinsichtlich aller drei Fragemodi auf den ersten 10 Rangplätzen auftauchen, so erweisen sich die Stereotype ‚vorlaut‘, ‚Streber‘ und ‚Klugscheißer‘ als „polyvalent“: Sie werden häufig benutzt und erfahren und sind angstbesetzt.

Ein nächster Auswertungsschritt galt der Frage, welche der Stereotype ‚leistungskonnotiert‘ sind. Hierzu wurden die bereits dargestellten Ratings mit den Schulnoten, mit den TIMSS-Leistungen, mit der Selbsteinschätzung mathematischer Schulleistungen und mit der Selbsteinschätzung allgemeiner Schulleistungen korreliert. Gibt es einen positiven Zusammenhang zwischen einem Stereotypenrating und der Mathematiknote bzw. einer der beiden Selbsteinschätzungen, so bedeutet dies, dass das Stereotyp umso häufiger erfahren bzw. benutzt wurde, je *schlechter* die eigenen Leistungen sind bzw. dass ein Schüler/eine Schülerin umso mehr Angst davor hat, so genannt zu werden, je schlechter seine/ihre Schulleistungen sind. Anders bei den TIMSS-Aufgaben. Hier stehen positive Zusammenhänge dafür, dass ein Schüler/eine Schülerin, der/die die TIMSS-Aufgaben besser gelöst hat, häufiger entsprechende Stereotype benutzt oder erfährt bzw. davor Angst hat, entsprechend genannt zu werden.

Tabelle 2 dokumentiert Korrelationen zwischen der Häufigkeit, mit der Schülerinnen und Schüler erlebt haben, von anderen mit einem bestimmten Stereotyp belegt worden zu sein und verschiedenen Messungen von mathematischen bzw. allgemeinen Schulleistungen. Die kanonische Korrelationsanalyse (SPSS MACRO CANCECORR) zeigt, dass 32,5% der Varianz der Ratings zu den 30 erfragten Stereotypen auf der Basis der mathematischen Leistungsbewertungen erklärt werden kann. Dieser Anteil ist hoch signifikant ($p \leq 0,001$).

Die Tabelle zeigt, dass ausweislich der Einzelanalysen bivariater Korrelationen etwa ein Drittel der Stereotype leistungskonnotiert sind. Mit den Stereotypen ‚Streber‘, ‚Genie‘, ‚Intelligenzbestie‘, ‚Mathe-As‘, ‚Spitzenschüler‘, ‚oberschlau‘, ‚Strebersau‘ und ‚strebsam‘ werden überwiegend Schüler mit guten Zeugnisnoten in Mathematik belegt. Als ‚Außenseiter‘, ‚Hohlkopf‘ oder ‚Spinner‘ werden hingegen Schüler mit schlechten Mathematiknoten überzufällig häufig bezeichnet.

Tabelle 2: Zusammenhänge zwischen Schul- bzw. Mathematikleistungen und der Häufigkeit mit der stereotype Benennungen von Schülerinnen und Schülern erfahren werden^a

	Mathematik-Note auf dem letzten Zeugnis	Gelöste TIMSS-Aufgaben	Wie würdest du dich selbst in Mathe einschätzen?	Welche Note würdest Du Dir insgesamt als Schüler selbst geben?
Streber	- 0,25**	0,11	- 0,09	- 0,13+
Genie	- 0,17*	0,07	- 0,27**	- 0,23**
Außenseiter	0,13+	- 0,10	0,13+	0,16*
Intelligenzbestie	- 0,17*	0,07	- 0,23**	- 0,25**
Mathe-As	- 0,32**	0,13+	- 0,35**	- 0,14+
Leserate	- 0,06	0,13+	0,02	- 0,14+
Hohlkopf	0,23**	- 0,09	0,16*	0,20**
Spitzenschüler	- 0,30**	0,05	- 0,27**	- 0,19**
Spinner	0,23**	0,12	0,18*	0,31**
Labertasche	0,09	0,01	0,11	0,26**
Stubenhocker	0,10	- 0,15*	0,08	0,08
Besserwisser	- 0,03	0,18*	- 0,13+	- 0,03
Klugscheißer	0,04	0,16*	- 0,07	0,08
überschlau	- 0,13+	0,01	- 0,17*	- 0,14+
Strebersau	- 0,18*	0,09	- 0,19**	- 0,11
Naturtalent	- 0,04	0,15*	- 0,20**	- 0,07
Wunderkind	- 0,03	0,07	- 0,17*	- 0,16*
superschlau	- 0,11	0,11	- 0,24**	- 0,15*
vorlaut	0,06	0,14+	- 0,01	0,08
strebsam	- 0,27**	0,01	- 0,27**	- 0,21**
allwissend	- 0,10	0,06	- 0,22**	- 0,13+
Störenfried	0,08	- 0,05	0,02	0,17+
oberflüchtig	- 0,01	- 0,02	- 0,19*	- 0,10

a tabelliert sind nur die Stereotype, die wenigstens für eine unabhängige Variable eine Signifikanz aufweisen

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

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

+ Die Korrelation ist auf dem Niveau von 0,10 (2-seitig) signifikant.

Auch bezogen auf die aktuelle tatsächliche Leistungsfähigkeit in Mathematik, gemessen – wenn auch sehr rudimentär – als Anzahl der richtig gelösten (von zwei) TIMSS-Aufgaben, sind viele Stereotype – wenn auch in etwas geringerem Ausmaß – leistungskonnotiert. Schülerinnen und Schüler mit guten TIMSS-Leistungen werden überzufällig häufig als ‚Mathe-As‘, ‚Leserate‘, ‚Besserwisser‘, ‚Klugscheißer‘, ‚Naturtalent‘ und ‚vorlaut‘ bezeichnet. Schlechte TIMSS-Leister werden überzufällig häufig als ‚Stubenhocker‘ bezeichnet.

Stereotypenerfahrung kovariiert auch mit der Selbsteinschätzung der Mathematikleistung. Schülerinnen und Schüler mit einer hohen Selbsteinschätzung ihrer Mathematikleistung werden überzufällig häufig als ‚Genie‘, ‚Intelligenzbestie‘, ‚Mathe-As‘, ‚Spitzenschüler‘, ‚Besserwisser‘, ‚überschlau‘, ‚Strebersau‘, ‚Naturtalent‘, ‚Wunderkind‘, ‚superschlau‘, ‚strebsam‘, ‚allwissend‘ und ‚oberflüchtig‘ bezeichnet. Schülerinnen und Schüler mit einer schlechten Selbsteinschätzung ihrer Mathematikleistungen werden überzufällig häufig als ‚Außenseiter‘, ‚Hohlkopf‘ oder ‚Spinner‘ bezeichnet.

Mit der Stereotypenwahrnehmung ebenfalls korreliert ist die Selbsteinschätzung der allgemeinen Schulleistung. Schülerinnen und Schüler mit einer hohen allgemeinen Selbsteinschätzung werden überzufällig häufig ‚Streber‘, ‚Genie‘, ‚Intelligenzbestie‘, ‚Mathe-As‘, ‚Leserate‘, ‚Spitzenschüler‘, ‚überschlau‘, ‚Wunderkind‘, ‚superschlau‘, ‚strebsam‘ und

,allwissend‘ genannt. Schülerinnen und Schüler mit niedriger allgemeiner Selbsteinschätzung werden überzufällig häufig ,Außenseiter‘, ,Hohlkopf‘, ,Spinner‘, ,Labertasche‘ und ,Störenfried‘ genannt.

Eine Zusammenschau der Korrelationen der Stereotypenerfahrung mit allen vier Arten der Leistungseinschätzung zeigt, dass ,Genie‘, ,Intelligenzbestie‘, ,Mathe-As‘, ,Spitzen-schüler‘, ,überschlau‘ und ,strebsam‘ in besonderem Maße leistungskonnotiert sind. Sie weisen jeweils mindestens drei signifikante Korrelationen auf: Hohe Leistungen gehen mit überzufällig häufiger Erfahrung dieser Stereotypenzuweisung einher. Es gibt aber auch leistungskonnotierte Stereotype, die besonders häufig mit niedriger Leistung einhergehen: ,Außenseiter‘, ,Hohlkopf‘ und ,Spinner‘.

Tabelle 3: Zusammenhänge zwischen Schul- bzw. Mathematikleistungen und der Häufigkeit mit der andere Schülerinnen und Schülern mit Stereotypen bezeichnen^a

	Mathematik-Note auf dem letzten Zeugnis	Gelöste TIMSS-Aufgaben	Wie würdest du dich selbst in Mathe einschätzen?	Welche Note würdest Du Dir insgesamt als Schüler selbst geben?
Genie	0,18*	0,14*	0,12+	0,05
Außenseiter	0,10	- 0,12+	0,10	0,21**
Schleimer	- 0,09	0,19*	- 0,06	0,01
Angeber	- 0,14+	0,16*	- 0,09	- 0,11
Leseratte	0,21**	- 0,14*	0,14+	0,16*
Hohlkopf	0,06	- 0,03	0,04	0,14+
Spitzenschüler	0,17*	- 0,14*	0,10	0,13+
Spinner	- 0,05	0,17+	0,00	0,08
Besserwisser	- 0,07	0,13+	- 0,03	0,04
überschlau	0,09	- 0,16*	0,08	0,03
altklug	0,00	- 0,08	0,12+	0,00
Strebersau	0,07	- 0,10	0,02	0,12+
Alleswisser	0,04	0,02	0,09	0,17*
Naturtalent	0,03	0,00	0,01	- 0,04
Wunderkind	0,14+	- 0,05	0,13+	0,02
superschlau	0,08	- 0,14+	0,08	0,07

a tabelliert sind nur die Stereotype, die wenigstens für eine unabhängige Variable eine Signifikanz aufweisen

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

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

+ Die Korrelation ist auf dem Niveau von 0,10 (2-seitig) signifikant.

Zusammenhänge zwischen der Leistung und der Tendenz, andere mit Stereotypen zu belegen, sind weniger ausgeprägt. Die kanonische Korrelationsanalyse zeigt hier, dass 23,5% der Varianz der Stereotypenratings auf der Basis der Mathematikleistungsbewertungen vorhergesagt werden kann. Dieser Anteil verfehlt Signifikanz knapp ($p \geq 0,05$). *Tabelle 3* zeigt, dass bei Analyse der bivariaten Korrelationen zwar wiederum etwa ein Drittel der Stereotype leistungskonnotiert ist, die Koeffizienten fallen allerdings geringer aus. Schüler mit schlechten Noten nennen andere überzufällig häufig ,Genie‘, ,Leseratte‘, ,Wunderkind‘, gute Schüler bezeichnen ihre Mitschüler überzufällig häufig als ,Angeber‘.

Die Tendenz, andere mit Stereotypen zu belegen, ist ebenfalls leistungskonnotiert, wenn man sie auf den Zusammenhang mit der aktuellen tatsächlichen Leistungsfähigkeit in Mathematik (gelöste TIMSS-Aufgaben) prüft: Schülerinnen und Schüler mit guten TIMSS-Leistungen bezeichnen andere überzufällig häufig als ,Genie‘, ,Schleimer‘, ,An-

geber', ,Spinner' oder ,Besserwisser'. Schüler mit schlechten TIMSS-Leistungen nennen Mitschüler überzufällig häufig ,Außenseiter', ,Leseratte', ,Spitzenschüler', ,überschlau' und ,superschlau'.

Stereotypennutzung kovariiert auch mit der Selbsteinschätzung der Mathematikleistung. Schülerinnen und Schüler mit einer niedrigen Selbsteinschätzung ihrer Mathematikleistung nennen ihre Mitschülerinnen und Mitschüler überzufällig häufig ,Genie', ,Leseratte', ,altklug' und ,Wunderkind'. Schülerinnen und Schüler mit einer hohen Selbsteinschätzung nutzen hingegen *keines* der 30 Stereotypen überzufällig häufig.

Ebenfalls mit der Stereotypennutzung korreliert ist die Selbsteinschätzung der allgemeinen Schulleistung. Schülerinnen und Schüler mit einer schlechten allgemeinen Selbsteinschätzung nennen andere überzufällig häufig ,Außenseiter', ,Leseratte', ,Hohlkopf', ,Spitzenschüler', ,Strebersau', ,Alleswisser', ,Genie', ,Intelligenzbestie', ,Mathe-As', ,Leseratte', und ,Spitzenschüler'. Schüler mit guter allgemeiner Selbsteinschätzung nutzen allerdings wiederum *keines* der 30 Stereotype überzufällig häufig.

Eine Zusammenschau der Korrelationen der Stereotypennutzung macht zunächst deutlich, dass es einen Unterschied zwischen fremdbewerteten Leistungen und Selbsteinschätzungen als Kovariaten der Stereotypennutzung gibt. Schüler mit guter mathematischer oder allgemeiner (auf schulische Leistungen insgesamt bezogener) Selbsteinschätzung nutzen *keine* Stereotype überzufällig häufig: Wer ein positives Leistungsselbstkonzept hat, hat es offenbar nicht nötig, andere mit derogativen oder Respekt zum Ausdruck bringenden Stereotypen zu belegen. Schlechte Schülerinnen und Schüler hingegen – hier gibt es keine Unterschiede zwischen Leistung und Selbsteinschätzung – nennen ihre Mitschülerinnen und Mitschüler vor allem ,Leseratte' oder ,Spitzenschüler'. Das Stereotyp ,Genie' hat offenbar sowohl derogative als auch ironische Konnotationen, es wird sowohl von Befragten mit guten als auch mit schlechten Mathematikleistungen benutzt. Gute Mathematikschüler nennen ihre Mitschüler darüber hinaus besonders häufig ,Angeber'.

Nachdem bis hierher belegt wurde, dass etwa die Hälfte der im Fragebogen abgefragten Stereotype von Schülerinnen und Schülern manchmal oder häufig ,erlitten' werden, fast alle manchmal oder häufig benutzt werden und immerhin etwa ein Viertel angstbesetzt ist und dass sowohl Nutzung als auch ,Erleiden' der Stereotype vielfach leistungskonnotiert ist, also etwas damit zu tun hat, wie gut Schülerinnen und Schüler in Mathematik sind, soll im nächsten Schritt geprüft werden, ob auch Ängste vor der Benennung mit bestimmten Stereotypen etwas mit den Leistungen in Mathematik zu tun haben. Erst ein solcher Befund würde unsere Annahme plausibel machen, dass Ängste vor der Stereotypisierung – z.B. als Streber – Auswirkungen auf Schulleistungen haben könnten, wie wir es in der Haupterhebung unserer Studie prüfen wollen. Hier erbringt die kanonische Korrelationsanalyse erneut einen Varianzanteil von 23,5% der Angstratings, der auf der Basis der vier Mathematikleistungsbewertungen vorhergesagt werden kann. Dieser Anteil ist allerdings im Falle der Angstratings signifikant ($p \leq 0,05$). Tabelle 4 dokumentiert die einzelnen bivariaten Korrelationskoeffizienten.

Tabelle 4: Zusammenhänge zwischen Schul- bzw. Mathematikleistungen und dem Ausmaß der Angst von anderen mit Stereotypen belegt zu werden^a

	Mathematik-Note auf dem letzten Zeugnis	Gelöste TIMSS-Aufgaben	Wie würdest du dich selbst in Mathe einschätzen?	Welche Note würdest Du Dir insgesamt als Schüler selbst geben?
Streber	-0,20**	0,01	-0,07	-0,02
Streber	-0,20**	0,01	-0,07	-0,02
Genie	-0,02	-0,09	-0,11	-0,14+
Genie	-0,02	-0,09	-0,11	-0,14+
Außenseiter	-0,14+	0,03	-0,08	-0,05
Außenseiter	-0,14+	0,03	-0,08	-0,05
Schleimer	-0,14+	0,03	-0,12	-0,09
Schleimer	-0,14+	0,03	-0,12	-0,09
Angeber	-0,14+	0,05	-0,05	-0,03
Angeber	-0,14+	0,05	-0,05	-0,03
Hirnie	0,05	-0,04	-0,14+	-0,07
Hirnie	0,05	-0,04	-0,14+	-0,07
oberschlau	-0,14+	-0,04	-0,01	-0,02
oberschlau	-0,14+	-0,04	-0,01	-0,02
Alleswisser	0,02	-0,13+	0,08	0,09
Alleswisser	0,02	-0,13+	0,08	0,09
superschlau	0,01	-0,05	0,09	0,12+
superschlau	0,01	-0,05	0,09	0,12+
allwissend	0,05	-0,10	0,10	0,15+
allwissend	0,05	-0,10	0,10	0,15+
Störenfried	-0,12+	0,03	-0,00	-0,04
Störenfried	-0,12+	0,03	-0,00	-0,04

a tabelliert sind nur die Stereotype, die wenigstens für eine unabhängige Variable eine Signifikanz aufweisen

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

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

+ Die Korrelation ist auf dem Niveau von 0,10 (2-seitig) signifikant.

Zunächst einmal macht die Tabelle deutlich, dass praktisch ausschließlich Schülerinnen und Schüler mit *guten* Mathematiknoten Angst vor der Benennung mit Stereotypen zum Ausdruck bringen. Von den 30 vorgegebenen Stereotypen weisen 26 eine negative Korrelation mit der Schulnote (also eine positive Korrelation mit guten Schulleistungen) auf, eine bei Annahme einer Gleichverteilung im Binomialtest hoch signifikante Einseitigkeit ($p < 0,001$). Signifikante Einzelkorrelationen gibt es ebenfalls praktisch nur für den Zusammenhang von Ängsten und guten Schulleistungen. Beim ‚Streber‘-Vorwurf ist der Zusammenhang hoch signifikant. Weitere überzufällige Zusammenhänge finden sich für ‚Außenseiter‘, ‚Schleimer‘, ‚Angeber‘, ‚oberschlau‘ und ‚Störenfried‘. Praktisch unkorreliert sind TIMSS-Leistungen und die Angst vor Stereotypen. Es gibt eine einzige Signifikanz: Schülerinnen und Schüler mit hohen TIMSS-Leistungen haben besonders wenig Angst, ‚Alleswisser‘ genannt zu werden. Befragte mit einer positiven Selbsteinschätzung ihrer Mathematikleistungen haben überzufällig häufig Angst, ‚Hirnie‘ genannt zu werden. Befragte mit einer positiven Einschätzung ihrer allgemeinen Schulleistungen haben überzufällig häufig Angst, ‚Genie‘ genannt zu werden. Schülerinnen und Schüler mit einer schlechten Selbsteinschätzung ihrer allgemeinen Schulleistung haben überzufällig häufig Angst, ‚superschlau‘ oder ‚allwissend‘ genannt zu werden. Insgesamt lässt sich festhalten, dass hinsichtlich der Selbsteinschätzungen und der durch zwei TIMSS-Aufgaben gemessenen objektiven Leistungsfähigkeit in Mathematik kein systematischer Zusammenhang mit der Angst vor der stereotypen Benennung durch Mitschüler besteht. Ein

recht deutlicher Zusammenhang besteht aber zwischen Ängsten und Schulnoten, wobei besonders hinsichtlich des Strebervorwurfs ein hoch signifikanter Zusammenhang besteht.

In einem weiteren Auswertungsschritt sind wir der Frage nachgegangen, ob sich die Leistungskonnotiertheit von Ängsten vor stereotypen Benennungen zwischen Mädchen und Jungen unterscheidet. Hierzu wurden zunächst die Korrelationskoeffizienten zwischen Ängsten und Mathematiknote getrennt für Jungen und Mädchen berechnet. Diese sind in Tabelle 5 dokumentiert.

Tabelle 5: Korrelationen zwischen Stereotypenangst und Mathematiknote nach Geschlechtern getrennt^a

	Mathenote auf dem letzten Zeugnis	
	Mädchen	Jungen
Streber	-0,27**	-0,16
Streber	-0,27**	-0,16
Genie	-0,17+	0,07#
Genie	-0,17+	0,07#
Außenseiter	-0,03	-0,21*
Außenseiter	-0,03	-0,21*
Schleimer*	-0,00	-0,26*#
Schleimer*	-0,00	-0,26*#
Angeber	-0,01	-0,22*
Angeber	-0,01	-0,22*
Mathe-As	-0,22*	0,07#
Mathe-As	-0,22*	0,07#
Leseratte	-0,20*	0,03
Leseratte	-0,20*	0,03
Spitzenschüler	-0,26*	0,04#
Spitzenschüler	-0,26*	0,04#
Spinner	0,17+	-0,16#
Spinner	0,17+	-0,16#
Störenfried	0,00	-0,19+
Störenfried	0,00	-0,19+

a tabelliert sind nur die Stereotype, die wenigstens für eine unabhängige Variable eine Signifikanz aufweisen

Die Korrelationskoeffizienten unterscheiden sich signifikant zwischen Mädchen und Jungen.

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

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

+ Die Korrelation ist auf dem Niveau von 0,10 (2-seitig) signifikant.

Die dokumentierten Korrelationen machen deutlich, dass Mädchen mit guten Mathenoten vor allem Angst davor haben, ‚Streber‘ (höchste Korrelation), ‚Genie‘, ‚Mathe-As‘, ‚Leseratte‘ oder ‚Spitzenschüler‘ genannt zu werden. Mädchen mit schlechten Mathematiknoten haben überzufällig häufig Angst, ‚Spinner‘ genannt zu werden. Jungen mit guten Mathematikleistungen haben besonders häufig Angst davor, ‚Außenseiter‘, ‚Schleimer‘, ‚Angeber‘ oder ‚Störenfried‘ genannt zu werden. Jungen mit schlechten Schulleistungen in Mathematik haben vor keiner stereotypen Benennung überzufällig häufig Angst. Für die Stereotype ‚Genie‘, ‚Schleimer‘, ‚Mathe-As‘, ‚Spitzenschüler‘ und ‚Spinner‘ unterscheiden sich die Korrelationskoeffizienten von Jungen und Mädchen signifikant.

Im nächsten Schritt prüfen wir für Mädchen und Jungen getrennt den Zusammenhang zwischen Schulnote und TIMSS-Leistung einerseits und den zwei Selbsteinschätzungen

andererseits. Bei Mädchen liegt die Korrelation zwischen TIMSS-Leistung und Schulnote nahe bei Null, nämlich $r = -0,07$, $p > 0,05$, die Korrelation zwischen Mathematik-Selbsteinschätzung und Mathematiknote bei $r = 0,56$, $p < 0,01$ und die Korrelation zwischen allgemeiner Selbsteinschätzung und Mathematiknote bei $r = 0,39$, $p < 0,01$. Bei Jungen sind alle drei Korrelationen deutlich höher: Die Korrelation mit den TIMSS-Leistungen liegt bei $r = -0,28$, $p < 0,01$, die mit der Selbsteinschätzung in Mathematik bei $r = 0,72$, $p < 0,01$ und die mit der allgemeinen Selbsteinschätzung bei $r = 0,50$, $p < 0,01$. Bereits durch diese Analysen wird deutlich, dass die Mathematiknoten von Jungen in stärkerem Maße als bei Mädchen Ausdruck ihrer ‚objektiven‘ mathematischen Leistungsfähigkeit sein dürften – als deren sehr groben Indikator man die Lösungsqualität zu den zwei TIMSS-Aufgaben ansehen kann.

Prüfen wir explorativ in einer hierarchischen Regressionsanalyse, welche Variablen die besten Prädiktoren der Mathematiknote sind, so ergibt sich folgendes Bild: Bezieht man nur das Geschlecht in die Vorhersage der Mathematiknote ein, so ergibt sich kein signifikanter Unterschied zwischen Jungen und Mädchen, der β -Koeffizient von Geschlecht auf die Mathematiknote liegt bei $\beta = 0,11$, $p > 0,10$. Die Jungen haben zwar eine geringfügig schlechtere Mathematiknote, doch ist der Unterschied nicht signifikant. Lässt man in einem zweiten Schritt der Regressionsanalyse das Programm entscheiden (METHOD = STEPWISE), welche Stereotypenängste einen bedeutsamen Beitrag zur Vorhersage der Mathematiknote leisten, so ist dies (allein) die Angst vor dem Strebervorwurf, die mit $\beta = -0,21$ zu Buche schlägt. Der Geschlechtsunterschied in der Mathematiknote wird nunmehr marginal signifikant $\beta = 0,12$, $p = 0,09$. Dies bedeutet, die Mädchen hätten bessere Schulnoten als Jungen, wenn Angst vor dem Strebervorwurf nicht stärker mit ihren Noten zusammenhinge als bei Jungen. Prüft man in einem nächsten Schritt, ob sich die Schulnote in Mathematik auch auf der Basis der gelösten TIMSS-Aufgaben vorhersagen lässt, so ist dies der Fall, der β -Koeffizient liegt bei $\beta = -0,16$, $p = 0,03$. Weiterhin signifikant und von der Bedeutung her *größer* bleibt die Vorhersage auf der Basis der Angst vor dem Strebervorwurf: $\beta = -0,21$, $p < 0,01$. Die Angst vor dem Strebervorwurf steht mit der Mathematiknote in engerer Verbindung als die ‚objektive‘ mathematische Leistungsfähigkeit, wie sie ansatzweise mit den beiden TIMSS-Aufgaben gemessen wurde. Der Geschlechtsunterschied bleibt marginal signifikant: $\beta = 0,13$, $p = 0,08$. Bringt man zum Schluss noch die Selbsteinschätzung der Mathematikleistung in die Regressionsgleichung ein, so ergibt sich ein überraschender Befund: Der Geschlechtsunterschied wird nunmehr hoch signifikant: $\beta = 0,16$, $p < 0,01$. Die Angst vor dem Strebervorwurf hat weiterhin einen bedeutsamen Einfluss: $\beta = -0,16$, $p < 0,01$. Die TIMSS-Leistung hingegen verliert ihren signifikanten Einfluss: $\beta = 0,07$, $p = 0,21$. Die Selbsteinschätzung der Mathematikleistung (durch Angabe der Note, die man selbst den eigenen Leistungen in Mathematik geben würde) hat allerdings einen sehr massiven Einfluss: $\beta = 0,60$, $p < 0,001$.

Nachdem Pilotstudie 1, wie unten noch ausführlicher zu diskutieren sein wird, den Beleg erbracht hat, dass es zumindest sinnvoll ist, der Frage nachzugehen, wie nachteilig sich der Strebervorwurf auf mathematische Schulleistungen insbesondere von Mädchen auswirkt, versucht Pilotstudie 2 zu prüfen, ob die womöglich unterschiedliche Wirksamkeit des Strebervorwurfs bei Mädchen und Jungen damit zusammenhängt, was unter einem Streber verstanden wird. Dieser Frage wurde in der zweiten Pilotstudie dadurch nachgegangen, dass zu 51 Stereotypen über das Bild des Strebers Ratings erbeten wurden. Um nun zu prüfen, ob ein zwischen Mädchen und Jungen unterschiedliches Verständnis dessen vorliegt, was unter einem Streber verstanden wird, wurden exploratori-

sche und konfirmatorische Faktorenanalysen berechnet. Zunächst wurden die 51 Items einer unrotierten exploratorischen Hauptkomponentenanalyse unterzogen. Hierbei stellte sich heraus, dass 17 der 51 Items auf der ersten Hauptkomponente laden. Die weiteren 34 Items laden auf 10 weiteren Faktoren. Die nachfolgenden Analysen beziehen sich nur auf die Items der ersten Hauptkomponente.⁵ Für diese wurde in einer mittels AMOS durchgeführten konfirmatorischen Faktorenanalyse geprüft, ob Ladungsäquivalenz zwischen Jungen und Mädchen vorliegt, ob sich also – inhaltlich betrachtet – das Verständnis des Konstrukts ‚Streber‘ zwischen Jungen und Mädchen gleicht. Die Analysen zeigten, dass Ladungsäquivalenz vorliegt. Der χ^2 -Anpassungswert für äquivalente Ladungen liegt bei $\chi^2 = 386,38$, $df = 254$, $p < 0,001$, $\chi^2/df = 1,521$, $RMSEA = 0,058$, $ECVI=3,14$, der für nicht äquivalente Ladungen bei $\chi^2 = 369,89$, $df = 238$, $p < 0,001$, $\chi^2/df = 1,554$, $RMSEA = 0,060$, $ECVI=3,24$. Der Vergleich beider Werte über den χ^2 -Differenzentest zeigt, dass sich durch Freigabe der geschlechtsspezifischen Ladungen für Mädchen und Jungen keine signifikante Modellverbesserung erreichen lässt. *Tabelle 6* zeigt die Ladungsmatrix sowie Mittelwerte und Standardabweichungen getrennt nach Mädchen und Jungen für die 17 Items der 1. Hauptkomponente der 51 in der zweiten Pilotstudie abgefragten Items.

Tabelle 6: Ladungsmatrix zum Konzept ‚Streber‘^a

Item	M		s		Ladung
	♀	♂	♀	♂	
Streber haben es viel schwerer Freunde zu finden.	1,71*	2,03	1,04	0,89	0,70
Streber wollen etwas Besseres sein als ihre Mitschüler.	1,64*	1,95	0,97	1,01	0,69
Mit Strebern will ich nichts zu tun haben.	0,62*	1,02	0,71	0,99	0,66
Viele Streber in einer Klasse machen das Klassenklima kaputt.	1,19*	1,61	1,06	1,24	0,64
Streber wollen vor allem bei Lehrern gut angesehen sein.	1,65*	2,07	1,03	1,01	0,56
Streber sind einsam.	1,49	1,51	0,90	0,94	0,56
Streber und Schleimer ist eigentlich dasselbe.	0,90*	1,28	1,02	1,17	0,52
Ein Streber ist jemand, der sich bei den Lehrkräften einschleimt.	1,88	2,01	0,96	1,02	0,48
Wer als Streber bezeichnet wird, hat nichts zu lachen.	1,50*	1,16	0,98	0,97	0,48
Streber haben viele Freunde.	0,83	0,64	0,69	0,65	-0,46
Ein Streber ist jemand, der sich gegen seine Klassengemeinschaft stellt.	0,69	0,86	0,77	0,86	0,46
Streber sitzen nur zuhause und büffeln für die Schule.	2,04	2,06	0,83	0,92	0,42
Lieber habe ich schlechtere Noten, als dass ich als Streber bezeichnet werde.	0,53	0,69	0,80	0,96	0,39
Wenn man für die Schule oft Dinge extra macht, die man nicht machen muss, ist man ein Streber.	1,08	1,34	0,96	1,23	0,39
Streber sind egoistisch.	1,16	1,33	0,92	1,24	0,38
Streber sind immer Klassenbeste.	1,78	1,76	1,07	1,17	0,35
Ich bin noch nie als Streber bezeichnet worden.	1,45*	2,11	1,25	1,16	0,26

a Mit einem ‘*’ gekennzeichnete Mittelwerte unterscheiden sich zwischen Mädchen und Jungen signifikant ($p < 0,05$).

Die Tabelle macht deutlich, dass die ladungsstärksten Items ($a > 0,60$), also die Items, die am besten Schülervorstellungen über den Streber abbilden, sämtlich etwas mit der sozialen Isolation des Strebers zu tun haben. Weiterhin kommt in der Tabelle zum Ausdruck, dass Jungen in der Regel höhere Werte haben, also den Stereotypen stärker zustimmen, als Mädchen. In neun Fällen gibt es – bei auch dort fast durchgehend höheren Werte der Jungen – keine signifikanten Unterschiede zwischen den Geschlechtern, bei acht Stereotypen gibt es signifikante Unterschiede. Einzig beim Item ‚Wer als Streber bezeichnet wird, hat nichts zu lachen‘ haben die Mädchen höhere Werte.

5 Diskussion

Im Chemnitzer BIQUA-Projekt „Streber vs. Nurd“ soll die These geprüft werden, dass deutsche Schülerinnen und Schüler der Sekundarstufe I stärker als nordamerikanische Schüler aus Angst vor dem Strebervorwurf ihre mathematische Leistungsfähigkeit nicht voll ausschöpfen. Geprüft werden soll weiterhin die These, dass die Angst vor dem Strebervorwurf Mädchen stärker in der Ausschöpfung ihres Leistungspotentials beeinträchtigt als Jungen. Zu der Frage, wie stark kulturelle Effekte im Vergleich zu Geschlechtseffekten wirken, werden keine Hypothesen formuliert. Diese Frage wird explorativ bearbeitet.

Die bisher am Standort Chemnitz durchgeführten Vorerhebungen belegen die Plausibilität der Annahme von Geschlechtsunterschieden in der Bedeutung des Strebervorwurfes für mathematische Schulleistungen. *Prima face* unterscheiden sich Mädchen und Jungen in ihren Mathematiknoten nicht. Berücksichtigt man jedoch statistisch, dass bei Mädchen ein stärkerer Zusammenhang zwischen der Angst vor dem Strebervorwurf und der Mathematiknote besteht und korrigiert die Notenunterschiede um diesen Effekt, so zeigt sich, dass Mädchen nach dieser Korrektur bereits schwach signifikant bessere Mathematiknoten als Jungen haben. Würden Mädchen zudem mehr Vertrauen in die eigenen Mathematikleistungen haben, hätten sie hoch signifikant bessere Mathematiknoten als Jungen.

Wie die signifikante Rohkorrelation zwischen TIMSS-Werten und Mathematiknoten belegt, ist die ‚objektive‘ Leistungsfähigkeit in Mathematik zwar für die Note von Bedeutung, dies gilt aber ausschließlich für Jungen: *Für Mädchen sind TIMSS-Leistung und Schulnote in Mathematik unkorreliert*. Zudem spielt die ‚objektive‘ Leistungsfähigkeit in Mathematik im Vergleich zur Streberangst und vor allem zum mathematischen Selbstwert eine nachgeordnete Rolle. ‚Objektive‘ Leistungsunterschiede sind jedenfalls in keiner Weise dafür verantwortlich, dass es Mädchen nicht gelingt, bessere Mathematiknoten als Jungen zu erzielen.

Pilotstudie 1 hat insgesamt deutlich gemacht, dass viele der von Jugendlichen im Umgang miteinander benutzten Stereotypen leistungskonnotiert sind, also etwas mit den Schulleistungen und speziell auch mit den Mathematikleistungen zu tun haben. Sie hat weiterhin erbracht, dass Ängste, bei Peers als Streber verschrien zu sein, mit Mathematiknoten kovariieren. Je besser die Mathematiknote einer/eines Sekundarstufe-I-Schülerin/-schülers, desto größer ihre/seine Angst, als Streber beschimpft zu werden. Die Vorerhebung hat die Plausibilität der Erwartung belegt, dass die Angst vor dem Strebervorwurf sich bei Mädchen negativer auswirkt: Hätten Mädchen nicht verstärkt Angst vor dem Strebervorwurf, hätten sie vermutlich bessere Mathematiknoten als Jungen.

Die Faktorenanalysen aus Pilotstudie 2 belegen, dass es sich bei der höheren ‚Wirksamkeit‘ des Strebervorwurfs bei Mädchen nicht um ein Artefakt eines zwischen den Geschlechtern unterschiedlichen Verständnisses dessen, was ein Streber ist, handelt. Jungen und Mädchen verstehen – im Lichte der erfragten Stereotype – unter einem Streber dasselbe.

Interessant ist der Befund, dass die Angst vor dem Strebervorwurf auch dann noch als Prädiktor der Mathematiknote relevant ist, wenn man das geringere mathematikbezogene Selbstvertrauen der Mädchen berücksichtigt. Mädchen können offenbar im Zusammenspiel von geringem mathematikbezogenen Selbstwert und Angst vor dem Strebervorwurf ihr eigentlich vorhandenes Leistungspotential nicht voll ausschöpfen.

Festzuhalten ist allerdings, dass die Chemnitzer Pilotstudie 1 das mathematikbezogene Selbstvertrauen sehr unvollkommen über die Note erfasst hat, die eine Schülerin/ein Schüler sich selbst im Fach Mathematik geben würde. In der Hauptuntersuchung sind hierzu umfänglichere Erhebungen notwendig, die sowohl konkret nach dem Selbstvertrauen in die eigene Leistungsfähigkeit in Mathematik fragen, als auch eine Anschlussfähigkeit an die Diskussion zu 'Fear-of-Success' ermöglichen sollten, die in der Geschlechterrollenliteratur sehr kontrovers geführt wird. ISHIYAMA und CHABASSOL (vgl. 1984) haben ein Instrument zur Messung der ‚Furcht vor sozialen Konsequenzen akademischen Erfolgs‘ vorgelegt, mit dem Mädchen zu wesentlich höheren ‚Fear-of-Success‘-Werten kommen als Jungen. Dieses Instrument bietet sich für die Haupterhebung ebenfalls an. – Auch die Messung der ‚objektiven‘ Leistungsfähigkeit in Mathematik sollte in der Haupterhebung umfänglicher gestaltet werden.

Alles in allem bleibt zuletzt noch einmal hervorzuheben, dass es sich hier um den Bericht zweier Pilotstudien mit dem Ziel der Plausibilisierung einer Forschungshypothese gehandelt hat, nicht um die Studie selbst, über die in ersten Ergebnissen an anderer Stelle berichtet wird (vgl. PELKNER/GÜNTHER/BOEHNKE 2002).

Anmerkungen

- 1 Im deutschen Sprachgebrauch taucht das Wort *Streber* zuerst im 16. Jahrhundert auf, wobei der Begriff zunächst die Bedeutung von *Widersacher* hatte. Im 18. Jahrhundert diente er zur Kennzeichnung karrieresüchtiger Beamter. Später wurde der Begriff des *Strebers* dann in der Studentensprache und schließlich in die Sprache der Schule übernommen.
- 2 Für die Datenerhebung und Dateneingabe danken wir Max HOFFMANN, Uwe LINGE und André RÜCKERT. Die Frage, ob die Durchführung der Datenerhebung von Schülern durch Schüler zu einer eingeschränkten Validität von Ergebnissen führt, kann hier nicht ausführlich diskutiert werden. An dieser Stelle kann nur auf gute Erfahrungen verwiesen werden, die der Zweitautor hiermit bereits in einer früheren DFG-Studie gemacht hat (vgl. HOFFMANN u.a. 1998).
- 3 siehe hierzu: [http://www. mpib-berlin.mpg.de/TIMSS-Germany/home-d.htm](http://www.mpib-berlin.mpg.de/TIMSS-Germany/home-d.htm); Testaufgaben TIMSS/III, S. 35+40; Stichwort: Busausflug in einen Wildpark bzw. Glanzi Waschpulver.
- 4 Für die Datenerhebung und Dateneingabe danken wir Andreas FÖRSTER, David NICOLAI und Christian WECK.
- 5 Einzelanalysen zu den weiteren Faktoren fallen in ihrer Substanz genauso aus, wie die unten berichteten zur ersten Hauptkomponente. Der Stichprobenumfang der zweiten Pilotstudie (N = 194) verbietet allerdings gemeinsame konfirmatorische Analysen, so dass auf diese im Rahmen der hier berichteten Vorerhebungen des Projekts verzichtet werden muss.

Literatur

- BAUMERT, J./BOS, W./LEHMANN, R. (Hrsg.) (2000): TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie – Mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn. Bd. 1: Mathematische und naturwissenschaftliche Grundbildung am Ende der Pflichtschulzeit; Bd. 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe. Opladen.
- BAUMERT u.a. 1997 = BAUMERT, J./LEHMANN, R./LEHRKE, M./SCHMITZ, B./CLAUSEN, M./HOSENFELD, I./KÖLLER, O./NEUBRAND, J. (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde. – Opladen.
- Dudenredaktion/Oxford University Press (Hrsg.) (1990): Duden-Oxford Großwörterbuch Englisch: englisch-deutsch; deutsch-englisch. – Mannheim.

- DWECK u.a. 1978 = DWECK, C.S./DAVIDSON, W./NELSON, S./ENNA, B. (1978): Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom and III. An experimental analysis. In: *Developmental Psychology*, Vol. 14, pp. 268-276.
- ENDERS-DRAGÄSSER, U./FUCHS, C. (1988): Interaktionen und Beziehungsstrukturen in der Schule. Eine Untersuchung an hessischen Schulen im Auftrag des hessischen Instituts für Bildungsplanung und Schulentwicklung. Frankfurt: Hessisches Institut für Bildungsplanung und Schulentwicklung.
- FRASCH, H./WAGNER, A. (1982): „Auf Jungen achtet man einfach mehr...“ Eine empirische Untersuchung zu geschlechtsspezifischen Unterschieden im Lehrer/innenverhalten gegenüber Jungen und Mädchen in der Grundschule. In: BREHMER, I. (Hrsg.): *Sexismus in der Schule. Der heimliche Lehrplan der Frauendiskriminierung*. – Weinheim, S. 260-278.
- FAULSTICH-WIELAND, H. (1993): Koedukation – Ein Ansatz zum Aufbrechen von Geschlechtsstereotypen? Referat auf dem niedersächsischen Lehrerinnenfortbildungskurs Nr. 93.02.81 in Hannover (vervielfältigtes Manuskript).
- FENNEMA, E. (1990): Teachers' beliefs and gender differences in mathematics. In: FENNEMA, E./LEDER, G. C. (Hrsg.): *Mathematics and gender* – New York, pp. 169-187.
- GLUMPLER, E. (1995): Feministische Schulforschung. In: ROLFF, H.-G. (Hrsg.): *Zukunftsfelder von Schulforschung*. – Weinheim, S. 133-155.
- GÖTZE, K. (1994): Für ein Physikum einfach zu dumm In: Freitag, 17.Juni 1994, Nr. 25, S. 14.
- HAGEMEISTER, V. (1991): Argumente für die Fortsetzung der Koedukation. Teil 1: Über Erfolge von Mädchen und Mißerfolge von Jungen in unseren Schulen, hrsg. vom Pädagogischen Zentrum. – Berlin.
- HANNOVER, B. (1999): Schulischer Kontext, geschlechtsbezogenes Selbstwissen und Fachpräferenzen. In: HANNOVER, B./KITTLER, U./METZ-GÖCKEL, H. (Hrsg.): *Sozialkognitive Aspekte der pädagogischen Psychologie*. Band I. Dokumentation des 3. Dortmunder Symposions für Pädagogische Psychologie 1998. – Essen, S. 125-139.
- HOFFMANN u.a. 1998 = HOFFMANN, D./BOEHNKE, K./MÜNCH, T./GÜFFENS, F. (1998). Radiohörertypen und Entwicklungsbedarf: Hörfunknutzung im Kontext jugendlicher Entwicklung. In *Medienpsychologie*, Jg. 10 (1), S. 131-148.
- HORSTKEMPER, M. (1987): Schule, Geschlecht und Selbstvertrauen: Eine Längsschnittstudie über Mädchensozialisation in der Schule. – Weinheim.
- ISHIYAMA, F. I./CHABASSOL, D. J. (1984): Fear of Success Consequence scale: Measurement of fear of social consequences of academic success. In: *Psychological Reports*, Vol. 54(2), pp. 499-504.
- JONES, J. C./SHALLCRASS, J. (1972): Coeducation and adolescent values. In: *Journal of Educational Psychology*, Vol. 63(4), pp. 334-341.
- JUVONEN, J./MURDOCK, T. B. (1995): Grade-level differences in the social value of effort: Implications for self-presentation tactics of early adolescents. In: *Child Development*, Vol. 66(6), pp. 1694-1705.
- KOEPKE, M. (1988): Mathe für Mädchen in den USA. Lust auf Risiko. In: *Frauen und Schule*, 8. Jg., H. 24, S. 14-15.
- KRAUSE, C. (1990): Mädchen und Jungen in der Schule der DDR. In: *Die Deutsche Schule*, Beiheft 1, S. 126-138.
- KREIENBAUM, M. A. (1988): Frauen bilden Macht. Fachkongreß Frauen und Schule. In: *Frauen und Schule*, 7. Jg., H. 23, S. 24-25.
- KRUSE, A. M. (1993): Non-koedukativer Unterricht als geschlechtsbewußter Unterricht. Erfahrungen aus Dänemark und England. In: GLUMPLER, E. (Hrsg.): *Erträge der Frauenforschung für die LehrerInnenbildung*. – Bad Heilbrunn, S. 120-148.
- LANDSHEER u.a. 1998 = LANDSHEER, H. A./MAASSEN, G.-H./BISSHOP, P./ADEMA, L. (1998): Can higher grades result in fewer friends? A reexamination of the relation between academic and social competence. In: *Adolescence*, Vol. 33, pp. 185-191.
- LEE, V. E./BRYK, A. S. (1986): Effects of single-sex secondary schools on student achievement and attitudes. In: *Journal of Educational Psychology*, Vol. 78, pp. 381-395.
- LI, Q. (1999): Teachers' beliefs and gender differences in mathematics: a review. In: *Educational Research*, Vol. 41(1), pp. 63-76.
- MCLAREN, A./GASKELL, J. (1995): Now you see it, now you don't: Gender as an issue in school science. In: GASKELL, J./WILLINSKY, J. (Eds.): *Gender informs Curriculum: From Enrichment to Transformation*. – Toronto, pp. 136-156.

- MEYER, W. U./DICKHÄUSER, O. (1998): What are the causes of gender differences in young children's ability attributions? – Bielefeld (unveröff. Manuskript).
- NOACK, P. (1990): Jugendentwicklung im Kontext. Zum aktiven Umgang mit sozialen Entwicklungsaufgaben in der Freizeit. – München.
- NYSSSEN, E. (1994): Aber ich behandle doch Mädchen und Jungen gleich. Über die Notwendigkeit der Frauenforschung in der LehrerInnenbildung. In: GLUMPLER, E. (Hrsg.): Koedukation. Entwicklungen und Perspektiven. – Bad Heilbrunn, S. 162-179.
- PECHTL, S. (1994): Chancengleichheit zwischen Knaben und Mädchen: ein Überblick über die feministische Schulforschung mit Bezugnahme auf das gegenwärtige österreichische Bildungssystem. – Frankfurt am Main.
- PELKNER, A./GÜNTHER, R./BOEHNKE, K. (2002): Die Angst vor sozialer Ausgrenzung als leistungshemmender Faktor? Zum Stellenwert guter mathematischer Schulleistungen unter Gleichaltrigen. In: Zeitschrift für Pädagogik, 45. Beiheft, S. 326-340.
- PFISTER, G. (1993): Der Widerspenstigen Zähmung. Rauman eignung, Körperlichkeit und Interaktion in der Grundschule. – Frankfurt/M., S. 67-83.
- RICHTER, S. (1996): Unterschiede in den Schulleistungen von Mädchen und Jungen. – Regensburg.
- RUSTEMEYER, R. (1999): Geschlechtstypische Erwartungen zukünftiger Lehrkräfte bezüglich des Unterrichtsfaches Mathematik und korrespondierend (Selbst-)Einschätzungen von Schülerinnen und Schülern. In: Psychologie in Erziehung und Unterricht, 46. Jg., S. 187-200.
- SCHWARTZ, S. H. (1992): Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In: Advances in Experimental Social Psychology, Vol. 25, pp. 1-65.
- SCHWARTZ, S. H. (1994): Beyond Individualism/Collectivism: New cultural dimensions of values. In: KIM, U./TRIANDIS, H. C./KAGITCIBASI, C./CHOI, S.-C./YOON, G. (Eds.): Individualism and collectivism: Theory, method and applications. – Newbury Park, pp. 85-119.
- TIEDEMANN, J./FABER, G. (1995): Mädchen im Mathematikunterricht: Selbstkonzept und Kausalattribution im Grundschulalter. In: Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 27. Jg., S. 61-71.
- UGUROGLU, M. E./WALBERG, H. J. (1986): Predicting achievement and motivation. In: Journal of Research and Development in Education, Vol. 19(3), pp. 1-12.
- WALBERG, H.J./SINGH, R.P./TSAL, S.-L. (1984): The achievement productivity of psychological variables: A replication and extension in Rajasthan, India. In: Alberta Journal of Educational Research, Vol. 30(3), pp. 171-178.

Korrespondenzadresse: Prof. Dr. Klaus Boehnke, International University Bremen, Campus Ring 1, 28759 Bremen, e-mail: K.Boehnke@iu-bremen.de