

ESE

**Emotionale und Soziale Entwicklung
in der Pädagogik der Erziehungshilfe und bei Verhaltensstörungen**



**Heft 4 (2022)
Soziales Lernen und Bildung**

Bibliografie:

Viktoria Pöchmüller, Clemens Hillenbrand
und Ute Koglin:

Vornamen beeinflussen Verhaltensurteile –
Eine experimentelle Studie zu Beurteilungsfehlern
in der Einschätzung des Sozialverhaltens.

Emotionale und Soziale Entwicklung (ESE), 4 (4), 108-119.
<https://doi.org/10.35468/5950-09>

Gesamtausgabe online unter:

<http://www.es-e-zeitschrift.net>
<https://doi.org/10.35468/5950>

ISSN 2629-0170

**Vornamen beeinflussen Verhaltensurteile –
Eine experimentelle Studie zu Beurteilungsfehlern
in der Einschätzung des Sozialverhaltens**

*Viktoria Pöchmüller, Clemens Hillenbrand
und Ute Koglin*

Abstract

Vornamen wecken Assoziationen, die die Wahrnehmung und Beurteilung einer Person beeinflussen können (z.B. Tobisch & Dresel, 2017). Das Ziel der Studie ist die Analyse des Einflusses von Vornamensstereotypen bzgl. des sozioökonomischen Status auf die Verhaltensbeurteilung durch Lehrkräfte. Die Studie folgt einem 2x2-Design. Die Ergebnisse einer multivariaten Varianzanalyse zeigen den signifikanten Einfluss von Vornamensstereotypen zu Gunsten solcher Namen, mit denen ein hoher sozioökonomischer Status assoziiert wird. Auf dieser Grundlage wird die Relevanz diagnostischer Kompetenzen von Lehrkräften diskutiert.

Keywords

Vornamensstereotype, Verhaltensbeurteilung, diagnostische Kompetenzen, multivariate Varianzanalyse

Einleitung

Eine hinreichende Qualität und Güte diagnostischer Urteile ist eine notwendige Bedingung effektiven (sonder-)pädagogischen Handelns. Nur auf Grundlage einer systematischen, kriteriengeleiteten Diagnostik der Kompetenzen und Lernvoraussetzungen von Schüler:innen lässt sich ein adaptives Unterrichts- und Förderangebot gestalten und eine optimale Passung zwischen Anforderungen und Schülermerkmalen herstellen (Tröster, 2019).

Um den Ansprüchen des systematischen Vorgehens von Diagnostik gerecht zu werden, benötigen Lehrkräfte diagnostische Kompetenzen, die die Grundlage zuverlässiger Urteile bilden (Tröster, 2019). Denn Lehrkrafturteile sind nicht per se objektiv, reliabel und/oder valide (Hesse & Latzko, 2017), sondern unterliegen oftmals dem Einfluss nicht urteilsrelevanter Merkmale (Kaiser, Möller, Helm & Kunter, 2015). Für Schülergruppen, die mit bestimmten stereotypen Assoziationen besetzt sind, sind weniger zuverlässige Urteile zu erwarten (Karst & Bonefeld, 2020).

Die vorliegende Studie prüft den Einfluss stereotypischer Vornamensassoziationen, die mit dem sozioökonomischen Status von Schüler:innen einhergehen, auf die Verhaltensbeurteilungen durch Lehrkräfte. Hierfür werden $N = 125$ Lehrkräfte um die Einschätzung des Sozialverhaltens fiktiver Schülerprofile gebeten. Auf Basis der Ergebnisse einer multivariaten Varianzanalyse diskutiert der Beitrag die Bedeutsamkeit standardisierter Diagnosekriterien sowie die Relevanz diagnostischer Kompetenzen von Lehrkräften zur Prävention verzerrter Verhaltensurteile.

2 Vornamensstereotype als Einflussfaktor auf die Verhaltensbeurteilung

Stereotype beschreiben eine kognitive Komponente von Vorurteilen und bezeichnen Wissen über geteilte Assoziationen mit einem bestimmten Merkmal (Werth, Seibt & Mayer, 2020). Solche Vorurteile helfen, die soziale Welt zu strukturieren (Karst & Bonefeld, 2020). Vorurteile sind zumeist an äußere, besonders auffällige Merkmale (z.B. Geschlecht) geknüpft (Kawakami, Amodio & Hugenberg, 2017).

Vornamen spielen eine zentrale Rolle bei der Aktivierung von Stereotypen (Anderson-Clark, Green & Henley, 2008; Kleen & Glock, 2020). Sie können als soziale Identität einer Person bezeichnet werden. Denn mit dem Vornamen werden bereits bestimmte Eigenschaften verbunden und der Trägerin bzw. dem Träger zugeschrieben (Kleen & Glock, 2020). Diese Assoziationen betreffen in erster Linie das Geschlecht oder Alter einer Person und darüber hinaus spezifische Personeneigenschaften wie bspw. die Attraktivität (Rudolph, Böhm & Lummer, 2007) oder den sozioökonomischen Status (Tobisch & Dresel, 2017).

2.1 Urteilsverzerrungen durch Aktivierung von Vornamensstereotypen

Die nationale und internationale Sozialforschung zeigt, dass Lehrkräfte oftmals allein aufgrund des Vornamens einer Schülerin bzw. eines Schülers urteilen (z.B. Anderson-Clark et al., 2008; Tobisch & Dresel, 2017). Dabei werden unterschiedliche Stereotype durch den Vornamen aktiviert. Die bisherige Forschung konzentriert sich auf die Einflussfaktoren Geschlecht, Migrationshintergrund und Assoziationen bzgl. des sozioökonomischen Status.

Vornamensstereotype sind kein spezifisch auftretendes Phänomen. Einerseits sind sie kultur- und kontextabhängig (Kleen & Glock, 2020), andererseits bestätigt die empirische Forschung den Einfluss des stereotypisch assoziierten sozioökonomischen Status auf die Leistungs-

beurteilung auch in verschiedenen Kulturkreisen. In der Differenzierung anhand von Vornamen afro-amerikanischer versus westlicher Herkunft werden Kinder, mit deren Name ein höherer sozioökonomischer Status assoziiert wird, von Lehrkräften signifikant besser beurteilt, als dies aufgrund der allgemeinen Intelligenz prädictiert wäre (Alvidrez & Weinstein, 1999). Unabhängig von der tatsächlich gezeigten Leistung werden auch im deutschen Bildungssystem Schüler:innen in Abhängigkeit der mit ihrem Vornamen stereotypisch assoziierten ethnischen Herkunft und dem sozioökonomischen Status beurteilt. Die Assoziation der Zugehörigkeit zu einer ethnischen Minderheit und eines niedrigen sozioökonomischen Status beeinflussen die Beurteilung der Schülerleistung signifikant negativ (Tobisch & Dresel, 2017). Lorenz, Gentrup, Kristen, Stanat und Kogan (2016) stellen heraus, dass Lehrkrafturteile sozialen Verzerrungen unterliegen. Sowohl hinsichtlich der Deutsch- als auch der Mathematikleistung scheinen Lehrkrafterwartungen systematisch durch den sozioökonomischen Status der Schulanfänger:innen beeinflusst. Dies geschieht unabhängig vom Geschlecht oder einem Migrationshintergrund der Schüler:innen zu Gunsten der Kinder mit einem höheren sozioökonomischen Status. Lehrkräfte scheinen ihre Einschätzungen somit u.a. an sozialen Stereotypen auszurichten (Lorenz et al., 2016).

Dass Lehrkrafturteile nicht per se objektiven Kriterien entstammen, zeigen die berichteten Studienergebnisse. Stereotype zu unterschiedlichen Schülermerkmalen beeinflussen Lehrkrafterwartungen und -urteile (z.B. Kaiser et al., 2015; Tobisch & Dresel, 2017). Jedoch werden nicht nur die Erwartungen und das Urteil der wahrnehmenden Person (hier der Lehrkraft), sondern ebenso das Verhalten der stereotypisierten Personengruppe beeinflusst (Petersen & Six, 2020). Der Stereotype-Threat-Ansatz (Steele & Aronson, 1995) weist auf mögliche negative Konsequenzen hin: Personen erleben in Situationen, in denen sie fürchten, aufgrund negativer Stereotype beurteilt zu werden oder durch ihr Verhalten gruppenspezifische negative Stereotype zu verstärken, ein Gefühl der Bedrohung. Wilbert (2010) weist in einer Studie mit Förderschüler:innen im Bereich Lernen nach, dass die Aktivierung des Stereotype-Threat zu negativen Effekten auf die Intelligenztestleistung und zu erhöhter Prüfungsangst führt. Hannover und Kessels (2011) verdeutlichen die komplexen Prozesse anhand des Faktors Geschlecht, denn Lehrkräfte versuchen offenbar „wahrgenommene Geschlechterunterschiede und geschlechtsbezogene Nachteile in Schulfächern entsprechend der gängigen Geschlechterstereotype auszugleichen und somit Jungen in einem stereotyp weiblichen Fach einen Kompetenzvorteil einzuräumen“ (Latsch, 2014, S. 26), sodass Mädchen wiederum Nachteile erfahren. Auch wenn dies in der vorliegenden Studie, in der die Erhebung von Lehrkrafturteilen erfolgte, nicht im Fokus steht, werden damit erschwerte Bedingungen stereotypisierter Gruppen für soziale Situationen ersichtlich.

2.2 Verhaltensbeurteilung in der Schule

Die Verzerrung von Leistungserwartungen und -beurteilungen durch die Aktivierung von Stereotypen beschreibt einen Risikofaktor für die (Lern-) Entwicklung, sofern urteilsirrelevante Merkmale die Leistungsbeurteilung negativ beeinflussen. Denn während Lehrkrafturteile über leistungsrelevante Merkmale die Grundlage für die Anpassung des Unterrichtsangebotes an die individuellen Lernvoraussetzungen der Schüler:innen darstellen (Tröster, 2019), bedingt der Einfluss nicht leistungsrelevanter Merkmale eine Verzerrung des Urteils. In der Folge ist eine adaptive Unterrichtsgestaltung nicht länger gewährleistet (Karst & Bonefeld, 2020). Laut Tobisch und Dresel (2017) verzeichnen Schüler:innen, von denen geringere Leistungen erwartet werden, außerdem tatsächlich einen geringeren Leistungszuwachs.

Schulische Lehr-Lernprozesse unterliegen fortlaufenden Beurteilungssequenzen, wobei auch das schulische Sozialverhalten der Schüler:innen durch die Lehrkraft beurteilt wird (Limbourg & Steins, 2011a). Dabei gibt es Hinweise auf geschlechtsspezifische Unterschiede, wonach Mädchen höhere Problembeurteilungen hinsichtlich emotionaler Probleme und prosozialen Verhaltens verzeichnen, Jungen hingegen bzgl. Verhaltensproblemen und Problemen mit Gleichaltrigen (Lohbeck, Schultheiß, Petermann & Petermann, 2015).

Soziale Kompetenzen beschreiben sogenannte fächerübergreifende Kompetenzen. Diese erweisen sich in der Bewältigung fachlicher Anforderungen als maßgeblich (Klieme, Artelt & Stanat, 2014) und verdeutlichen die Verknüpfung akademischen und sozial-emotionalen Lernens (Leidig, Hennemann & Hillenbrand, 2020). Akademische Lernprozesse werden angenommen, von emotional-sozialen Kompetenzen abhängig zu sein (Zins, Bloodworth, Weissberg & Walberg, 2004), sodass erst die ganzheitliche Betrachtung von Lernprozessen eine systematische Entwicklungsförderung ermöglicht (Leidig et al., 2020). Valide Verhaltensurteile und daran bemessene Unterstützungsangebote besitzen demnach auch Relevanz für die Gestaltung akademischer Lern- bzw. Bildungsprozesse.

Die Herausforderung besteht darin, soziale Kompetenzen explizit zu messen (Klieme et al., 2014). „Kritisch werden in diesem Zusammenhang insbesondere unklare Bewertungskriterien und mangelnde Ausbildung der Lehrkräfte im Hinblick auf entsprechende diagnostische Fähigkeiten betrachtet“ (van Ackeren & Clausen, 2011, S. 479). Unzureichende Vorgaben und fehlende diagnostische Kompetenzen beeinträchtigen ein systematisches Vorgehen im Sinne formeller Diagnostik. Diagnostische Kompetenzen sind keine generalisierbaren Fähigkeiten. Besonders in der Einschätzung affektiver Eigenschaften zeigt sich, dass Lehrkräfte Schülerkompetenzen weniger valide als Schulleistungen beurteilen (Tröster, 2019).

Eine weitere Herausforderung der standardisierten Beurteilung schulischen Sozialverhaltens liegt in der Kontextabhängigkeit spezifischer Verhaltensweisen – soziales Verhalten ist immer auf das Verhalten anderer bezogen (Limbourg & Steins, 2011b). Eine objektive und reliable Verhaltensbeurteilung berücksichtigt die spezifischen Rahmenbedingungen in der konkreten Situation. Um den Herausforderungen im Kontext der Beurteilung spezifischen Verhaltens zu begegnen, werden standardisierte Vorgehensweisen entwickelt, die einen wiederholten Einsatz erlauben, um Entwicklungsverläufe abbilden zu können (Casale, Huber, Hennemann & Grosche, 2019).

Eine objektive Leistungsbeurteilung anhand standardisierter Kriterien beugt Urteilsverzerrungen vor. Allerdings geschieht Diagnostik im (sonder-) pädagogischen Handlungsfeld nicht immer systematisch anhand standardisierter Vorgehensweisen, sondern gestaltet sich als Kombination formeller Diagnostik, die durch ein systematisches Vorgehen zur Informationsgewinnung anhand standardisierter Erhebungsinstrumente gekennzeichnet ist, und informeller unsystematischer Urteilsbildung auf Grundlage subjektiver Einschätzungen und Erwartungen (Rieß & Bolz, 2021). Insbesondere der Einfluss subjektiver Erfahrungen und Erwartungen der Lehrkraft kann die Entstehung systematischer Beurteilungsfehler begünstigen (Kaiser et al., 2015). Dabei dominieren einzelne Merkmale bzw. Informationen die Beurteilung. Doch diese müssen keine für die Urteilsfindung relevanten sein (Tröster, 2019).

3 Fragestellung und Hypothesen

Lehrkrafturteile sind nicht per se objektiv und nicht ausschließlich an urteilsrelevanten Faktoren ausgerichtet (Hesse & Latzko, 2017; Kaiser et al., 2015). Forschungsergebnisse verdeutlichen den Einfluss stereotypischer Assoziationen hinsichtlich des sozioökonomischen Status auf Lehrkrafterwartungen und -beurteilungen (z.B. Lorenz et al., 2016; Tobisch & Dresel, 2017). Das Ziel der vorliegenden Studie besteht in einer Sensibilisierung für den Einfluss (urteils-)irrelevanter Merkmale auf die Beurteilung und einer Diskussion der Notwendigkeit standardisierter Diagnosekriterien für einen bislang vorwiegend unsystematisch erfassten Kompetenzbereich. Dabei wird untersucht, ob die Aktivierung spezifischer Vornamensstereotype, die mit dem sozioökonomischen Status von Schüler:innen assoziiert sind, die Verhaltensbeurteilung von Lehrkräften beeinflusst. Die leitende Fragestellung lautet *Welchen Einfluss nehmen mit dem Schülermerkmal Vorname einhergehende stereotypische Assoziationen bzgl. des sozioökonomischen Status auf die Beurteilung des Sozialverhaltens?*

In Orientierung an dargestellten Forschungsergebnissen liegt zum einen die Annahme zugrunde, dass Schülerprofile mit Vornamen, die stereotypische Assoziationen eines niedrigen sozioökonomischen Status wecken, schlechter beurteilt werden als solche Profile, mit deren Name positive Assoziationen des sozioökonomischen Status einhergehen (H1). Darüber hinaus wird vermutet, dass männliche Schülerprofile signifikant schlechtere Verhaltensbeurteilungen in den ausgewählten Dimensionen erhalten als weibliche (H2).

4 Methodik

4.1 Design

Die Studie folgt einem empirisch quantitativen Vorgehen in einem querschnittlich angelegten experimentellen 2 (sozioökonomischer Status) x 2 (Geschlecht) Design. Dabei wird in einer randomisierten Zuweisung das Sozialverhalten anhand vier fiktiver Schülerprofile eingeschätzt. Die Profile sind inhaltsgeleich, sodass die Manipulation einzig über den Vornamen geschieht. Je zwei Profile sind männlich (Paul, Justin) bzw. weiblich (Marie, Chayenne), von denen je eines einen hohen (Paul, Marie) bzw. niedrigen (Justin, Chayenne) sozioökonomischen Status assoziiert. Erhoben wird der Einfluss der Vornamensstereotype (UV) auf die Einschätzung des Sozialverhaltens (AV) durch die Lehrkräfte. Der multivariate Plan erfasst Hyperaktivität, Verhaltensprobleme und Prosoziales Verhalten als abhängige Variablen (Döring & Bortz, 2016).

In der Durchführung sehen die Lehrkräfte nur die ihnen zufällig zugewiesene Vignette. Im Anschluss an die individuelle Bearbeitung wird aufgeklärt, dass vier parallele Versionen des inhaltsgeichen Bogens Anwendung finden und das Ziel in der Analyse des Einflusses von Vornamensstereotypen auf die Verhaltensbeurteilung besteht.

4.2 Stichprobe

Die Stichprobe umfasst $N = 125$ Lehrkräfte aus niedersächsischen Allgemeinen Schulen. Von den befragten Lehrkräften (weiblich $n = 79$) geben 14,9% an, in der Grundschule zu unterrichten, 23,1% in der Haupt-, Real- oder Oberschule, 28,9% am Gymnasium und 33,1% an der Förderschule. 15,6% der Lehrkräfte berichten, bislang weniger als fünf Jahre Berufserfahrung zu haben, 23% zwischen 6 und 10 Jahren, 32% zwischen 11 und 20 Jahren und 29,5% von mehr als 20 Jahren Berufserfahrung. Die randomisierte Zuweisung des

individuell beurteilten Schülerprofils ergibt folgende Verteilung: Das Schülerprofil *Justin* wird $n = 29$ Mal eingeschätzt, das Äquivalent *Paul* wird $n = 32$ Mal beurteilt. Die Verteilung auf die Schülerinnenprofile geschieht im Umfang von $n = 33$ Beurteilungen des Profils *Chayenne* und $n = 31$ Einschätzungen des Äquivalents *Marie*.

4.3 Erhebungsinstrument

Zur Datenerhebung wurde eine Vignette entworfen, die auf der Vorderseite zunächst ein fiktives Schülerprofil beschreibt. Die Personeninformationen sind inhaltsgleich, einzig die Vornamen, mit denen die Bögen betitelt sind, variieren. In der Wahl der Vornamen finden geschlechtsspezifische männliche und weibliche Namen in gleicher Anzahl Berücksichtigung. Ein Migrationshintergrund soll explizit nicht assoziiert sein. Als Grundlage dienen Namensstatistiken und Bibliografien (z.B. GfS, 2019).

Auf der Rückseite wurden die Lehrkräfte um die Einschätzung des Sozialverhaltens des Profils gebeten. Die Beurteilung geschieht auf einer vierstufigen Ratingskala (1 = „stimme vollkommen zu“ bis 4 = „stimme vollkommen dagegen“) anhand von Items der Skalen Hyperaktivität, Verhaltensprobleme und Prosoziales Verhalten der Lehrkraftversion des SDQ (Goodman, 2005). Vor der Analyse wurden die Items der Skalen Hyperaktivität und Verhaltensprobleme umgepolt, sodass höhere Werte mit höheren Problembeurteilungen einhergehen. Für die Skala Prosoziales Verhalten bedeuten höhere Werte nach einer Umpolung, dass mehr prosoziales Verhalten berichtet wird.

4.4 Auswertung

Zunächst werden deskriptivstatistische Analysen berechnet. Es folgt eine multivariate Varianzanalyse (MANOVA), um den Einfluss des mit dem Vornamen assoziierten sozioökonomischen Status sowie des assoziierten Geschlechts auf die Verhaltensbeurteilung zu analysieren. Die Profile werden hierfür zu Skalen zusammengefasst – Justin und Paul zu einer Skala „männlich“, Chayenne und Marie zu einer Skala „weiblich“, Justin und Chayenne zu einer Skala „niedriger sozioökonomischer Status“ sowie Paul und Marie zu einer Skala „hoher sozioökonomischer Status“.

5 Ergebnisse

5.1 Deskriptive Statistik

Es liegen mittlere Interkorrelationen zwischen den Variablen zum Sozialverhalten der Schüler:innen vor. Die Skala Hyperaktivität korreliert signifikant mit der Skala Verhaltensprobleme ($r = .40, p < .01$) und der Skala Prosoziales Verhalten ($r = -.43, p < .01$). Die Skala Verhaltensprobleme korreliert signifikant mit der Skala Prosoziales Verhalten ($r = -.57, p < .01$).

Anhand bivariater Korrelationen kann weiterhin aufgezeigt werden, dass das Geschlecht und Alter der Lehrkräfte nicht mit den Skalen zum Sozialverhalten der eingeschätzten Schüler:innen variieren (Geschlecht: Hyperaktivität $r = .04, p = .68$; Verhaltensprobleme $r = -.08, p = .36$; Prosoziales Verhalten $r = .04, p = .68$ sowie Alter: Hyperaktivität $r = .04, p = .68$; Verhaltensprobleme $r = .02, p = .85$; Prosoziales Verhalten $r = .07, p = .41$). Es besteht kein Zusammenhang zwischen dem Geschlecht und Alter der Lehrkräfte und deren Einschätzung des Sozialverhaltens.

5.2 MANOVA

Tabelle 1 stellt die deskriptive Statistik für die drei Skalen nach Geschlecht und sozioökonomischem Status dar. Die angeführte Teststatistik zeigt die Ergebnisse der MANOVA auf Skalenebene.

Tab. 1: Mittelwerte und Standardabweichungen für die Skalen Hyperaktivität, Verhaltensprobleme und Prosoziales Verhalten nach Geschlecht und sozioökonomischem Status und Teststatistik der MANOVA

Skala	Mädchen		Jungen		Teststatistik		
	SÖS niedrig	SÖS hoch	SÖS niedrig	SÖS hoch	Geschlecht	SÖS	Interaktion
	M (SD)	M (SD)	M (SD)	M (SD)	F (df=1,121)	F (df=1,121)	F (df=1,121)
HYP	2.97 (0.73)	2.26 (0.77)	2.76 (0.69)	2.31 (0.69)	0.37	20.04***	1.05
VP	2.44 (0.62)	1.89 (0.49)	2.59 (0.57)	1.98 (0.53)	1.50	33.49***	0.06
ProV	2.48 (0.44)	3.01 (0.42)	2.53 (0.39)	3.07 (0.50)	0.57	45.23***	0.01

Anmerkungen: HYP = Hyperaktivität; VP = Verhaltensprobleme; ProV = Prosoziales Verhalten; SÖS = sozioökonomischer Status; *** $p < .001$.

Die MANOVA mit den drei abhängigen Variablen Hyperaktivität, Verhaltensprobleme und Prosoziales Verhalten ist mit $F_{df=3,119} = 4740,21$ und $p < .001$ signifikant. Multivariat zeigt sich ein signifikanter Haupteffekt durch den Faktor sozioökonomischer Status ($F_{df=3,119} = 19,38$, $p < .001$), aber nicht für den Faktor Geschlecht ($F_{df=3,119} = 1,36$, $p = .26$). Eine Interaktion zwischen den Faktoren Geschlecht und sozioökonomischer Status besteht nicht ($F_{df=3,119} = 0,46$, $p = .71$).

Auf Skalenebene wiederholt sich dieses Ergebnismuster (s. Tab. 1). Die Varianzanalysen zeigen für alle drei abhängigen Variablen einen signifikanten Haupteffekt des sozioökonomischen Status auf die Verhaltensbeurteilung (Hyperaktivität: $F_{df=1,121} = 20,04$, $p < .001$; Verhaltensprobleme: $F_{df=1,121} = 33,49$, $p < .001$; Prosoziales Verhalten: $F_{df=1,121} = 45,23$, $p < .001$). So werden Hyperaktivität und Verhaltensprobleme bei Schülerprofilen mit einem Vornamen, der stereotypische Assoziationen eines höheren sozioökonomischen Status weckt, niedriger eingeschätzt. Gleichzeitig erfahren diese Profile eine höhere Einschätzung auf der Skala Prosoziales Verhalten. Für den Einfluss des assoziierten Geschlechts ergeben sich keine Effekte. Dies gilt auch für die Interaktion zwischen Geschlecht und sozioökonomischem Status, die ebenfalls nicht signifikant ist. Die Mittelwertunterschiede zwischen den Gesamtskalen beschreiben in Bezug auf alle Verhaltensdimensionen mit $d \geq |.8|$ starke Effekte (Hyperaktivität $d = .81$; Verhaltensprobleme $d = 1.03$; Prosoziales Verhalten $d = -1.23$) (Cohen, 1988).

6 Diskussion

Trotz aller Beschränkungen der internen und externen Validität der Studie, deuten die Ergebnisse darauf hin, dass Lehrkrafturteile in der Einschätzung ausgewählter Verhaltensdimensionen dem Einfluss aktivierter Vornamensstereotype unterliegen. Der mit dem Vornamen assoziierte sozioökonomische Status nimmt demnach einen signifikanten Einfluss auf die Verhaltensbeurteilung zu Gunsten der Namen, mit denen ein höherer sozioökonomischer Status assoziiert wird. Dies gibt einen Hinweis darauf, dass Stereotype von Lehrkräften gegenüber spezifischen Schülermerkmalen vor dem Hintergrund adaptiver Unterrichtsgestaltung und Chancengleichheit und Bildungsgerechtigkeit gezielt reflektiert werden müssen.

6.1 Ergebnisdiskussion

Die experimentelle Lehrkräftebefragung zeigt den höchst signifikanten Einfluss des mit Vornamen stereotypisch assoziierten sozioökonomischen Status auf die Verhaltensbeurteilung. Die Analysen erlauben die Schlussfolgerung, dass sich Verzerrungen der Verhaltensbeurteilung aufgrund von Vornamensstereotypen bzgl. des sozioökonomischen Status zu Gunsten solcher Namen zeigen, mit denen Assoziationen eines hohen sozioökonomischen Status einhergehen. Eine niedrigere Problembeurteilung zeigt sich in den Profilen mit Namen, die einen hohen sozioökonomischen Status assoziieren. Der mit dem Namen assoziierte sozioökonomische Status hat einen signifikanten Einfluss auf die Verhaltensurteile von Lehrkräften. Ein stereotypisch niedrig assoziierter sozioökonomischer Status wirkt sich signifikant negativ auf die Beurteilung definierter Dimensionen des Sozialverhaltens aus (H1). Entgegen der Annahme, Profile mit einem Vornamen, der geschlechtsspezifisch männliche Assoziationen weckt, würden schlechter beurteilt als Schülerinnenprofile, kann in der vorliegenden Untersuchung kein signifikanter Einfluss des stereotypisch assoziierten Geschlechtes auf die Verhaltensbeurteilung ermittelt werden (H2).

Als eine Schlussfolgerung erscheint eine Sensibilisierung für den Einfluss von Stereotypen auf die Verzerrung von Verhaltensurteilen notwendig. Van Ackeren und Clausen (2011) kritisieren bereits, dass v.a. in der Einschätzung sozialen Verhaltens standardisierte Beurteilungskriterien fehlen, was besondere Herausforderungen an diagnostische Verfahren stellt (Casale et al., 2019). Darüber hinaus benötigen Lehrkräfte spezifische diagnostische Kompetenzen zur validen Beurteilung sozial-emotionaler Kompetenzen. Die Relevanz formeller Urteilsbildung geht mit der Forderung nach einer Schärfung diagnostischer Kompetenzen von Lehrkräften einher. Diese sind Voraussetzung eines reflektierten und datenbasierten Vorgehens in der Urteilsbildung, werden jedoch nicht allein mit zunehmender praktischer Erfahrung erworben. Vielmehr bedarf eine entsprechende Professionalität des systematischen Erwerbs spezifischen Wissens und Könnens (Hesse & Latzko, 2017). Dabei erscheint auch die Sensibilisierung für und die Entwicklung eines differenzierten Verständnisses der zwei Kategorien formeller und informeller Diagnostik relevant. Denn Diagnostik und Urteilsbildung geschehen besonders im (sonder-)pädagogischen Handlungsfeld in Kombination systematischer und informeller Urteilsbildung (Rieß & Bolz, 2021). Die signifikanten Ergebnisse legen nahe, den Einfluss aktivierter Vorurteile auf die informelle Urteilsbildung zu reflektieren.

Diagnostische Kompetenzen sind Grundlage valider Urteilsfindung und wirksamen Unterrichtens (Tröster, 2019). In der Verschränkung akademischen und sozial-emotionalen Lernens zeigt sich die Bedeutsamkeit valider Verhaltensurteile auch für akademische Lernprozesse (Leidig et al., 2020). Daher sollten Verhaltensurteile durch Lehrkräfte möglichst genau sein, damit pädagogische Maßnahmen Wirksamkeit entfalten (Hesse & Latzko, 2017). Die

dargestellten Ergebnisse bestätigen den Einfluss urteilsirrelevanter Merkmale auf die Lehrkräfteeinschätzung jedoch eindrücklich.

Eine objektive Verhaltensbeurteilung anhand systematischer Vorgehensweisen erfährt auch vor dem Hintergrund bestehender Bildungsungleichheit Relevanz. Verzerrungen durch mit Vornamen assoziierte Stereotype des sozioökonomischen Status können negativ auf Bildungsverläufe wirken, wenn davon ausgegangen wird, dass Schüler:innen, denen ein hoher sozioökonomischer Status zugeschrieben wird, eher das Gymnasium besuchen (BPB, 2021). Dabei wird die Beurteilung sozialen Verhaltens u.a. aufgrund der Notwendigkeit sozialer Kompetenzen für die Bewältigung akademischer Lernanforderungen (Klieme et al., 2014) wirksam. Für das (sonder-)pädagogische Handeln ergeben sich an dieser Stelle drei zentrale Ableitungen. Lehrkräfte sind notwendigerweise für den Einfluss von Namensstereotypen auf Verhaltensurteile zu sensibilisieren. Darüber hinaus gilt es, diagnostische Kompetenzen von Lehrkräften insbesondere zur Beurteilung sozialen Verhaltens zu schärfen und deren Erwerb in der Ausbildung zu verankern. Zuletzt wird der Bedarf einer Entwicklung standardisierter Kriterien zur Verhaltensbeurteilung sichtbar. Aufgrund der Kontextabhängigkeit sozialer Verhaltensweisen ist dabei zu reflektieren, dass Items in kurzen Zeitabständen wiederholt einsetzbar sind und der Forderung nach einer Berücksichtigung der spezifischen Situation in der Verhaltensbeurteilung entsprechen (Casale et al., 2019).

Die MANOVA konnte den Einfluss aktivierter Vornamensstereotype bei Lehrkräften auf Verhaltensurteile bestätigen. Dabei zeigt sich eine systematische Verzerrung, die sich in einer höheren Problembewertung von Schülerprofilen niederschlägt, die mit Namen versehen sind, die Assoziationen eines niedrigen sozioökonomischen Status wecken. In der theoretischen Herleitung wurde bereits auf den Stereotype-Threat-Ansatz (Steele & Aronson, 1995) verwiesen. Inwieweit Schüler:innen befürchten, auf Basis negativer Stereotype beurteilt zu werden bzw. mit ihrem Verhalten negative Stereotype für die Gruppe der Vornamensträger:innen zu bestätigen und in welcher Form dies das gezeigte Verhalten und in der Konsequenz die Verhaltensbeurteilung durch die Lehrkraft beeinflusst, eröffnet Perspektiven für die weitere Forschung.

6.2 Limitationen

Wesentliche Limitation erfährt die vorliegende Studie durch die Einschränkungen der externen Validität. Aufgrund der Komplexität des alltäglichen Unterrichtsgeschehens sind die Ergebnisse kaum generalisierbar. Mithilfe der entwickelten Fallvignetten können das komplexe Unterrichtsgeschehen und die Kontextabhängigkeit von Verhalten nicht abgebildet werden. Des Weiteren beeinträchtigt die gemeinsame Betrachtung der Urteile von Lehrpersonen aus unterschiedlichen Schulformen die Generalisierbarkeit der Ergebnisse entscheidend. Eine Übertragung auf alle Lehrkräfte scheint nicht angemessen. Dennoch sind die höchst signifikanten Ergebnisse des Experiments hinsichtlich möglicher Konsequenzen für Schüler:innen und Lernbiografien zu reflektieren.

Die Individualität und Subjektivität der Aktivierung vornamensspezifischer stereotypischer Assoziationen hinsichtlich unterschiedlicher Merkmale (Rudolph et al., 2007), wie dem sozioökonomischen Status, führt auf inhaltlicher Ebene zu einer Beeinträchtigung der Übertragbarkeit der Ergebnisse. So ist davon auszugehen, dass andere Namen andere bzw. ähnliche Assoziationen in anderer Intensität wecken und folglich in einer Erhebung des Einflusses von Vornamensstereotypen auf die Verhaltensbeurteilung zu anderen Ergebnissen führen. Letztlich polarisieren Namensstereotype immer, sodass deren Erhebung und Analyse keinesfalls kritikfrei geschieht (Kaiser, 2010).

7 Fazit

Der Beitrag belegt nachdrücklich den Einfluss von Vornamensstereotypen bzgl. des assoziierten sozioökonomischen Status auf die Verhaltensbeurteilung durch Lehrkräfte, zu Gunsten von Vornamen, die Assoziationen eines höheren sozioökonomischen Status wecken. Der Bedarf einer Schärfung diagnostischer Kompetenzen von Lehrkräften sowie die Notwendigkeit zur Formulierung und Zugrundelegung standardisierter Kriterien und Vorgehensweisen zur Beurteilung eines bisher meist unsystematisch erfassten Kompetenzbereichs konnten bestätigt werden. Lehrkräfte sind für den Einfluss entsprechender Bias zu sensibilisieren. Denn nicht urteilsrelevante, aber prominente (Personen-) Merkmale, wie der Vorname, und damit verbundene Stereotype können Verhaltensurteile entscheidend beeinflussen und verzerren.

Literatur

- Ackeren, I. van, & Clausen, M. (2011). Soziale Kompetenzen im Kontext schulischer Leistungsmessung und -bewertung. In M. Limbourg & G. Steins (Hrsg.), *Sozialerziehung in der Schule* (S. 475–496). VS Springer.
- Alvidrez, J., & Weinstein, R. (1999). Early teacher perceptions and later student academic achievement. *Journal of Educational Psychology*, 91(4), 731–746. doi:10.1037/0022-0663.91.4.731
- Anderson-Clark, T.N., Green, R.J., & Henley, T.B. (2008). The relationship between first names and teacher expectations for achievement motivation. *Journal of Language and Social Psychology*, 27(1), 94–99. doi:10.1177/0261927X07309514
- Bundeszentrale für politische Bildung (BPB). (2021). *Der sozioökonomische Status der Schülerinnen und Schüler*. Abgerufen am 8. November 2021 von <https://m.bpb.de/nachschlagen/datenreport-2021/bildung/329670/der-soziooekonomische-status-der-schuelerinnen-und-schueler>
- Casale, G., Huber, C., Hennemann, T., & Grosche, M. (2019). *Direkte Verhaltensbeurteilung in der Schule: Eine Einführung für die Praxis*. Ernst Reinhardt.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation für Sozial- und Humanwissenschaften* (5. Aufl.). Springer. doi: 10.1007/978-3-642-41089-5
- Gesellschaft für deutsche Sprache e.V. (GfDS). (2019). *Die beliebtesten Vornamen*. Abgerufen am 12. Mai 2019 von <https://gfds.de/vornamen/beliebteste-vornamen/>
- Goodman, R. (2005). *Strengths and Difficulties Questionnaire. Fragebogen zu Stärken und Schwächen (SDQ-Deu): Lehrer 4-17*. Abgerufen am 4. Dezember 2018 von <https://sdqinfo.org/py/sdqinfo/b3.py?language=German>
- Hannover, B., & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze. *Zeitschrift für Pädagogische Psychologie*, 25(2), 89–103. doi: 10.1024/1010-0652/a000039
- Hesse, I., & Latzko, B. (2017). *Diagnostik für Lehrkräfte* (3. Aufl.). Barbara Budrich.
- Kaiser, A. (2010). „Kevin ist kein Name, sondern eine Diagnose!“. Der Vorname in der Grundschule: Klang-, Mode- oder Reizwort? *Die Grundschulzeitschrift*, 24, 26–29.
- Kaiser, J., Möller, J., Helm, F., & Kunter, M. (2015). Das Schülerinventar: Welche Schülermerkmale die Leistungsurteile von Lehrkräften beeinflussen. *Zeitschrift für Erziehungswissenschaften*, 18, 279–302. doi: 10.1007/s11618-015-0619-5
- Karst, K., & Bonefeld, M. (2020). Stereotype, Urteile und Urteilsakkuratheit von Lehrkräften: Eine Zusammenschau im Rahmen des Heterogenitätsdiskurses. In S. Glock & H. Kleen (Hrsg.), *Stereotype in der Schule* (S. 281–308). Springer.
- Kawakami, K., Amodio, D.M., & Hugenberg, K. (2017). Intergroup Perception and Cognition: An Integrative Framework for Understanding the Causes and Consequences of Social Categorization. In J.M. Olson (Hrsg.), *Advances in Experimental Social Psychology* (Bd. 55, S. 1–80). Academic Press. doi: 10.1016/bs.aesp.2016.10.001
- Kleen, H., & Glock, S. (2020). Sag mir, wie du heißt, dann sage ich dir, wie du bist: Eine Untersuchung von Vornamen. In S. Glock & H. Kleen (Hrsg.), *Stereotype in der Schule* (S. 99–131). Springer.
- Klieme, E., Artelt, C., & Stanat, P. (2014). Fächerübergreifende Kompetenzen: Konzepte und Indikatoren. In F. E. Weinert (Hrsg.), *Leistungsmessungen in Schulen* (S. 203–218). Beltz.
- Latsch, M. (2014). „Schlaue Mädchen – dumme Jungs?!“ *Der Einfluss von negativen Stereotypen über Jungen auf Leistung, Motivation und Gruppenprozesse im Klassenzimmer*. Dissertation, Freie Universität Berlin. Abgerufen von <https://refubium.fu-berlin.de/handle/fub188/6695>

- Leidig, T., Hennemann, T., & Hillenbrand, C. (2020). Integration sozial-emotionalen Lernens im (Fach-)Unterricht. *Zeitschrift für Heilpädagogik*, 71(4), 148–159.
- Limbourg, M., & Steins, G. (Hrsg.). (2011a). *Sozialerziehung in der Schule*. VS Springer.
- Limbourg, M., & Steins, G. (2011b). Sozialerziehung in der Schule. In M. Limbourg & G. Steins (Hrsg.), *Sozialerziehung in der Schule* (S. 11–28). VS Springer.
- Lohbeck, A., Schultheiß, J., Petermann, F., & Petermann, U. (2015). Die deutsche Selbstbeurteilungsversion des Strengths and Difficulties Questionnaire (SDQ-Deu-S): Psychometrische Eigenschaften, Faktorenstruktur und Grenzwerte. *Diagnostica*, 61(4), 222–235. doi: 10.1026/0012-1924/a000153
- Lorenz, G., Gentrup, S., Kristen, C., Stanat, P., & Kogan, I. (2016). Stereotype bei Lehrkräften? Eine Untersuchung systematisch verzerrter Lehrererwartungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 68, 89–111. doi: 10.1007/s11577-015-0352-3
- Petersen, L.-E., & Six, B. (Hrsg.). (2020). *Stereotype, Vorurteile und soziale Diskriminierung: Theorien, Befunde und Interventionen* (2. Auflage). Beltz.
- Rieß, B., & Bolz, T. (2021). Diagnostik in der Pädagogik bei Verhaltensstörungen. In H. Ricking, T. Bolz, B. Rieß & M. Wittrock (Hrsg.), *Prävention und Intervention bei Verhaltensstörungen: Gestufte Hilfen in der schulischen Inklusion* (S. 156–171). Kohlhammer.
- Rudolph, U., Böhm, R., & Lummer, M. (2007). Ein Vorname sagt mehr als 1000 Worte. *Zeitschrift für Sozialpsychologie*, 38(1), 17–31. doi: 10.1024/0044-3514.38.1.17
- Steele, C.M., & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811. doi: 10.1037//0022-3514.69.5.797
- Tobisch, A., & Dresel, M. (2017). Negatively or positively biased? Dependencies of teachers' judgements and expectations based on students' ethnic and social backgrounds. *Social Psychology of Education*, 20, 731–752. doi: 10.1007/s11218-017-9392-z
- Tröster, H. (2019). *Diagnostik in schulischen Handlungsfeldern: Methoden, Konzepte, praktische Ansätze*. Kohlhammer.
- Werth, L., Seibt, B., & Mayer, J. (2020). Vorurteile. In L. Werth, B. Seibt & J. Mayer (Hrsg.), *Sozialpsychologie – Der Mensch in sozialen Beziehungen: Interpersonale und Intergruppenprozesse* (2. Auflage) (S. 227–321). Springer.
- Wilbert, J. (2010). Stereotype-Threat Effekte bei Schülern des Förderschwerpunkts Lernen. *Heilpädagogische Forschung*, 4, 154–161.
- Zins, J.E., Bloodworth, M.R., Weissberg, R.P., & Walberg, H.J. (2004). The Scientific Base Linking Social and Emotional Learning to School Success. In J.E. Zins, R.P. Weissberg, M.C. Wang & H.J. Walberg (Hrsg.), *Building Academic Success on Social and Emotional Learning: What does the Research say?* (S. 3–22). Teachers College.