Fachbereich Erziehungswissenschaft und Psychologie

der Freien Universität Berlin

Investigating Gender and Gender Stereotyped Behaviors and Traits in Educational Settings from Multiple Perspectives

Dissertation

zur Erlangung des akademischen Grades

Doktorin der Philosophie (Dr. phil.)

Vorgelegt von

(MSc.)

Streck, Hannah

Berlin, Juli 2024

Erstgutachterin: Prof. Dr. Ursula Kessels

Zweitgutachterin: Prof. Dr. Ricarda Steinmayr

Weitere Kommissionsmitglieder:

Prof. Dr. Bettina Hannover, Prof. Dr. Katharina Kluczniok, Dr. Karoline Koeppen

Tag der Disputation: 04.11.2024

Author Contributions

For Manuscript 1 (*Linda's cars vs. Dominik's dolls: How do pedagogical educators in training react to children's violations of gender stereotypes?*), Hannah Streck and Ursula Kessels contributed to the study conception, design, and material preparation. Data collection and analyses were performed by Hannah Streck. The first draft of the manuscript was written by Hannah Streck. Ursula Kessels provided feedback on the manuscript and supervised the research project.

For Manuscript 2 (*Nice, but not smart? Attributional backlash from displaying prosocial behavior in the classroom*), the design and analytical plan was based on a study conducted by Ursula Kessels and Anke Heyder. The additional material was developed by Hannah Streck and Ursula Kessels. Data collection was conducted by Hannah Streck. The first draft of the manuscript was written by Hannah Streck. Ursula Kessels provided feedback on the manuscript and supervised the research project.

Manuscript 3 (*Instrumentality gives girls the edge: Gender-differential relations between instrumentality, achievement motivation, and self-esteem*) uses a dataset collected by Julia Herrmann and Ursula Kessels. Ursula Kessels developed the general idea for the manuscript. Anna K. Nishen analyzed the data and provided the method and results section of the manuscript. Hannah Streck conducted a literature review and wrote the introduction and discussion with Ursula Kessels. The manuscript was written by Hannah Streck, Anna K. Nishen and Ursula Kessels.

Acknowledgments

As “Denglisch” is not (yet) a fully accepted form of communication, I will restrict myself to this section.

An erster Stelle bedanke ich mich herzlich bei Ursula Kessels. Vielen Dank für deine Unterstützung in den letzten Jahren und für die Begleitung dieser Arbeit. Du hast, much to my delight, eine wunderbar großzügige open-door policy – vielen Dank, dass du dir die Zeit genommen hast, um über Wissenschaft, Gender, Lehre, Kultur und everything on my mind zu reden. Ich habe viel gelernt, diese Zeit sehr genossen, und ich hoffe, dass unsere Gespräche für dich auch nur einen Bruchteil so spannend waren wie für mich. Für die wissenschaftliche und persönliche Begleitung, für deine Hilfe bei allen kleinen und großen Schritten dieser Arbeit, für deinen Humor und deine Großzügigkeit bin ich sehr dankbar.

Vielen Dank auch an Ricarda Steinmayr, für die Bereitschaft diese Arbeit zu begutachten. Ich danke dir auch für das Teilen von Daten und Projekten – ich weiß das wirklich zu schätzen. Vielen herzlichen Dank auch an die weiteren Mitglieder meiner Promotionskommission: Professorin Bettina Hannover, Professorin Katharina Kluczniok und Dr. Karoline Koeppen.

An alle jetzigen und ehemaligen Mitglieder des Heterogenitäts-Teams: Henriette, Gülru, Alessia, Maria, Maike, Manon, Amelie. Vielen Dank für eure Unterstützung bei Datensammlung, Dateninput, Prüfung von Literaturverzeichnissen, der Navigation durch die FU-Systeme und noch vieles mehr!

Vielen Dank natürlich auch an Katharina Holder, Anke Heyder und Katharina Odebrecht für die Hilfe und Begleitung in diesen letzten Jahren.

Ich danke an dieser Stelle sehr herzlich allen Teilnehmenden an Schulen, Berufsschulen und der Universität für die Unterstützung dieser Forschung.

Anna und Diana, ihr wisst hoffentlich schon, wie dankbar ich für eure Freundschaft bin. Ohne euren Humor, euren thematischen und wissenschaftlichen Input und eure Ruhe wäre meine Zeit an der Freien Uni Berlin kaum vorstellbar. Ich bin so froh, dass wir uns gegenseitig gefunden haben und dass unsere Freundschaft in drei Dissertationen in Schwarz auf Weiß festgehalten wurde. Nun auch hier: Vielen Dank für alles, ich freue mich schon auf die nächsten Jahre mit euch!

Mom and Dad, I could not have done this without you. Thank you for your encouragement and support, for giving me perspective and making me laugh. Thank you for moving our family to New York, thank you for following me to the UK and giving me a home – I truly do not know where in the world I would be without you. I knew I could count on you for anything, and that has made the past few years possible.

Merle and my new brother-in-law, Joe, thank you for laughing with me and for giving me a home away from home. I cannot wait to visit you in Addis Ababa or Zurich. To my family, I love you. Thank you for everything.

To my friends Frida, Clara, Mila, Atanu, Inga, Hanne, Sarah, Paula, Mona, Micaela, Emma, Claudia, Ritika, Tara, Callum, Charlie and Will: you kept me sane, fed, housed, loved, and happy. Thank you for taking all of my calls, for talking about work when I wanted to, and for distracting me when I needed it.

I could not have finished this dissertation without this community of kind, smart, and thoughtful people. Thank you!

Table of Contents

Summary 3
Zusammenfassung 7
I. Introduction 11
Theoretical Background 16
Gender and Gendered Characteristics 17
Gender Identity and Gender Self-Categorization 17
Gender Stereotypes 18
Descriptive, Prescriptive, and Proscriptive Stereotypes 19
Content of Gender Stereotypes 20
Attribute Self-Perceptions 22
Gender Role Self-Concepts 22
Measuring Gender Role Self-Concepts 23
Gender Schema Theory 25
Gender In/congruence and Associated Outcomes 28
Gender Typicality and Atypicality 29
Adhering to or Violating Gender Stereotypes 30
Masculinity, Femininity, and Androgyny 32
Feminization of School 34
Educators 36
Peers 37
Self 37
Gender and Gendered Characteristics in School 39
Likability & Competence 39
Likability 40
Competence 42
Causal Attributions 44
Self-Esteem, Academic Contingent Self-Esteem, and Achievement Motivation 47
Self-Esteem 47
Academic Contingent Self-Esteem 48
Achievement Motivation 50
Summary and Study Overview 51
References I. 57

II. Empirical Work 77

Manuscript 1: 79

Streck, H., & Kessels, U. (Accepted, pending revision). Linda's cars vs. Dominik's dolls: How do pedagogical educators in training react to children's violations of gender stereotypes*? Archives of Sexual Behavior*.

Manuscript 2: 129

Streck, H., & Kessels, U. (2023). Nice, but not smart? Attributional backlash from displaying prosocial behavior in the classroom. *Social Psychology of Educatio*n, *26*(6), 1621-1649

Manuscript 3: 173

Streck, H., Nishen, A. K., & Kessels, U. (2022). Instrumentality gives girls the edge: Gender-differential relations between instrumentality, achievement motivation, and self-esteem. *Sex Roles, 86*(5), 379-394

III. Discussion 217

Summary of Results 219

Discussing Gender and Gendered Characteristics 222

Relation to Gender Schema Theory 222

Gender as a Status Characteristic 226

Limitations & Future Research Directions 230

Limitations 230

Future Research Directions 233

Multiple perspectives 234

Intersectionality 235

Implications 238

Gender Stereotypes and Gender Role Self-Concepts 238

Gender Salience in Schools 239

Gender Salience Beyond Schools 242

Conclusion 244

References II. 245

IV. Supplementary Documents 259

Eigenständigkeitserklärung 260

Curriculum Vitae 261

Summary

Gender has proven to be a relevant factor in schools, particularly given the finding that girls tend to outperform boys. Empirical findings relating to not just gender but also constructs associated with gender, such as gender stereotypes and gender role self-concepts, can offer insight into academic, social, and personal functioning of children and adolescents. In regarding not academic achievement but rather a variety of variables associated with the educational system, such as perceived likability and competence, attributions of success, and self-esteem and achievement motivation, the present works expands upon the existing literature. Past work has examined gender in relation to these constructs, but research on how gender stereotypes and gender role self-concept relate to perceived likability and competence, attributions of success, and self-esteem constructs and achievement motivation is less common. As the school environment features a variety of unique roles, the perspectives of educators, peers, and students themselves are taken into account in the present works.

In manuscript 1, we gathered gender stereotypes about young children held by German adults in a first study consisting of $N = 397$ participants. This information allowed the construction of vignettes in which gender stereotypical and nonstereotypical three-year-old boys and girls were portrayed. In our main study, pre-service pedagogical educators ($N = 414$) indicated their liking, perceived competence, creativity, self-esteem, prosocial behavior, internalizing and externalizing problems of hypothetical children using a 2 x 2 between participants design (target gender: boy, girl; target stereotyped behavior: masculine, feminine). We hypothesized that counterstereotypical children, particularly feminine boys, would receive the lowest ratings in likability. Further, we hypothesized that ratings of competence would differ between stereotypical and nonstereotypical children and expected the latter children to receive positive ratings in terms of creativity and self-esteem. The analyses revealed no significant difference in liking but showed that masculine girls received better ratings in terms of competence, creativity, and self-esteem compared to feminine girls.

This stands in contrast with past literature, which shows adults displaying negative biases towards feminine boys (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013). The backlash for feminine boys, seen in the past literature, and our findings, showing benefits for masculine girls, highlight the social standing of masculinity (Feinman, 1981).

In manuscript 2, we investigated a peer perspective on how gender stereotyped classroom behavior could impact causal attributions of achievements, based on work by Kessels and Heyder (2020) and the paradox of praise (Meyer et al., 1979; Möller, 2005). Study 2 was an experimental vignette study in which high-achieving students displaying prosocial or nondescript behavior were described. $N = 324$ 9th grade students participated in the 2 x 2 mixed design study, where target gender (girl, boy) was varied between participants and vignette (prosocial, nondescript) was varied within participants. Participants were asked to indicate (1) whether the hypothetical students would receive positive reactions from the teacher, as well as (2) how the students' high achievements could be attributed, and (3) how they would rate targets on social and gender-related outcomes, such as perceived intelligence, masculinity, and femininity. We hypothesized that prosocial students would be perceived as receiving more positive teacher reactions. We further expected prosocial students to receive attributions of greater effort and lower ability, which should be mediated by the expected teacher reaction (Möller, 2005; Weiner & Kukla, 1970). Prosocial students were perceived as more effortful and less able than nondescript students, but the hypothesized teacher reactions did not mediate this association. The findings highlight that social classroom behavior, here feminine stereotyped prosocial behavior, can impact how academic achievements are attributed by peers. The belief that prosocial students are less able than nondescript peers is particularly relevant for female students.

Manuscript 3, a cross-sectional study, established the relation between self-esteem constructs and achievement motivation with gender role self-concept. $N = 355$ 9th grade

students provided information on their gender role self-concept, global self-esteem, academic contingent self-esteem, hope for success, and fear of failure. We hypothesized positive relations between instrumentality and global self-esteem and hope for success, while expecting a negative relation with instrumentality and academic contingent self-esteem and fear of failure. Importantly, gender differential relations were hypothesized, as past work has shown that the association between a masculine gender role self-concept and self-esteem measures can be stronger in female individuals (Cate & Sugawara, 1986; Hirokawa & Dohi, 2007; Whitley, 1988). Structural equation modeling showed a benefit of an instrumental orientation, which was associated with greater self-esteem and hope for success, and lower academic contingent self-esteem and fear of failure. The association between instrumentality and global self-esteem was stronger for girls, for whom this association also indirectly related to lower fear of failure, highlighting how an instrumental gender role self-concept can particularly benefit girls. Our findings that instrumentality can benefit students expand the literature on the feminization of school, which has presented masculinity in school as mostly negative (Lyng, 2009).

The results of all three empirical works highlight a pattern of benefits of masculinity and the disadvantages of femininity from the perspective of educators, peers, and students themselves. The discussion of the present dissertation thus outlines how gender can be understood as a status characteristic (J. Berger et al., 1972; Gerber, 2009), with maleness and masculinity perceived as holding more status than femaleness and femininity. This is seemingly in conflict with past work on the feminization of school, presenting the better fit between femininity and school (Heyder & Kessels, 2013). We discuss some practical and theoretical limitations of the present work and outline future research directions. The implications of the work focus on the feminization of school and on the question of whether reducing gender salience in schools could be advantageous, ultimately highlighting current trends in society regarding focus on gender.

Zusammenfassung

Gender hat sich als ein relevanter Faktor in der Schule erwiesen, insbesondere angesichts der Befunde, dass Mädchen oft bessere Leistungen als Jungen erbringen. Dabei können nicht nur Forschungserkenntnisse zu Gender, sondern auch mit Gender verbundene Konstrukte wie Geschlechterstereotype und Geschlechtsrollen-Selbstkonzepte, Einblicke in die akademische, soziale und persönliche Entwicklung von Kindern und Jugendlichen bieten. Die vorliegenden Arbeiten erweitern die bisherigen Forschungserkenntnisse, indem nicht die akademische Leistung, sondern eine Vielzahl von Variablen betrachtet werden, die im Bildungssystem von Relevanz sind. Hierzu gehören wahrgenommene Sympathie und Kompetenz, Kausalattributionen sowie Selbstwertkonstrukte und Leistungsmotivation. Bisherige Forschung hat die genannten Variablen bereits in Bezug auf Gender untersucht, der Zusammenhang von Geschlechterstereotypen und Geschlechtsrollen-Selbstkonzepten mit wahrgenommener Sympathie und Kompetenz, Kausalattributionen sowie Selbstwertkonstrukten und Leistungsmotivation ist empirisch jedoch weniger etabliert. Da in der schulischen Umgebung viele Rollen vertreten sind, werden in den vorliegenden Arbeiten die Perspektiven von pädagogischen Fachkräften, Peers und Schüler*innen berücksichtigt.

In Manuskript 1 erfassten wir in einer ersten Studie mit $N = 397$ Teilnehmenden, welche Geschlechterstereotype deutsche Erwachsene über junge Kinder vertreten. Aus diesen Informationen erstellten wir Vignetten, in denen geschlechtskonforme und -nonkonforme Jungen und Mädchen dargestellt wurden. In unserer zweiten Studie teilten angehende Erzieher*innen ($N = 414$) ihre Einschätzungen zur wahrgenommener Sympathie, Kompetenz, Kreativität, Selbstwertgefühl, prosozialem Verhalten sowie zu internalisierenden und externalisierenden Problemen von beschriebenen Kindern in einem 2 x 2 between-participants Design (Target Geschlecht: Junge, Mädchen; Target stereotypisiertes Verhalten: maskulin, feminin). Wir erwarteten, dass geschlechtsnonkonforme Kinder, insbesondere feminine Jungen, die niedrigsten Bewertungen in Bezug auf Sympathie erhalten würden. Zusätzlich

stellten wir die Hypothese auf, dass sich die Kompetenzbewertungen zwischen geschlechtskonformen und -nonkonformen Kindern unterscheiden würden. Wir erwarteten, dass nonkonforme Kinder positive Bewertungen in Bezug auf Kreativität und Selbstwertgefühl erhalten würden. Die Analysen demonstrierten keinen signifikanten Unterschied in der Sympathie, zeigten aber, dass maskuline Mädchen in Bezug auf Kompetenz, Kreativität und Selbstwert bessere Bewertungen erhielten als feminine Mädchen. Dies steht im Kontrast zu bisherigen Forschungserkenntnissen, die Vorurteile gegenüber femininen Jungen zeigten (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013). Die negativen Einstellungen gegenüber femininen Jungen, welche sich in bestehenden Befunden zeigen, und unsere Ergebnisse, welche Vorteile für maskuline Mädchen darstellen, unterstreichen den sozialen Stellenwert von Maskulinität (Feinman, 1981).

In Manuskript 2 untersuchten wir aus einer Peer Perspektive, basierend auf Forschung von Kessels und Heyder (2020) und den paradoxen Auswirkungen von Lob (Meyer et al., 1979; Möller, 2005), wie geschlechtsstereotypisiertes Verhalten im Klassenzimmer Attributionen der Leistungen von Schüler*innen beeinflussen könnte. Studie 2 war eine experimentelle Vignettenstudie, in der leistungsstarke Schülerinnen oder Schüler, die prosoziales oder unauffälliges Verhalten zeigten, beschrieben wurden. $N = 324$ Neuntklässler*innen nahmen teil an der 2 x 2 mixed-Design Studie, in der das Target Geschlecht (Mädchen, Junge) zwischen Teilnehmenden und Vignette (prosozial, unauffällig) innerhalb der Teilnehmenden variiert wurde. Teilnehmende gaben an, (1) ob die beschriebenen Schüler*innen positive Reaktionen von der Lehrkraft erhalten würden, (2) wie die Leistungen der Schüler*innen attribuiert werden könnten und (3) wie sie die fiktiven Schüler*innen in sozialen und geschlechtsbezogenen Variablen wie wahrgenommene Intelligenz, Maskulinität und Femininität einschätzten. Wir stellten die Hypothese auf, Teilnehmende würden angeben, dass prosoziale Schüler*innen im Vergleich zu unauffälligen Schüler*innen mehr positive Reaktionen der Lehrkraft erhalten würden. Darüber hinaus

erwarteten wir für prosoziale Schüler*innen Attributionen von mehr Anstrengung und weniger Fähigkeit, was laut Attributionstheorie (Möller, 2005; Weiner & Kukla, 1970) durch die erwartete Reaktion der Lehrkraft mediiert werden sollte. Die Leistungen von prosozialen Schüler*innen wurden mehr auf Anstrengung und weniger auf Fähigkeit zugeschrieben als die Leistungen von unauffälligen Schüler*innen, dies wurde jedoch nicht durch die erwarteten Reaktionen der Lehrkraft mediiert. Die Ergebnisse zeigen, dass soziales Verhalten im Klassenzimmer, hier feminin-stereotypisiertes, prosoziales Verhalten, die Attributionen akademischer Leistungen durch Peers beeinflussen kann. Die Annahme, dass prosoziale Schüler*innen weniger Fähigkeiten besitzen als unauffällige Mitschüler*innen, kann besonders für Mädchen relevant sein.

Manuskript 3, eine Querschnittsstudie, untersuchte Zusammenhänge zwischen Selbstwertkonstrukten und Leistungsmotivation mit Geschlechtsrollen-Selbstkonzept. N = 355 Neuntklässler*innen nahmen an einer Umfrage zu ihrem Geschlechtsrollen-Selbstkonzept, globalen Selbstwert, akademisch kontingentem Selbstwert, Hoffnung auf Erfolg und Angst vor Misserfolg teil. Wir erwarteten positive Assoziationen zwischen Instrumentalität und globalem Selbstwert, sowie Hoffnung auf Erfolg, während wir eine negative Assoziation mit Instrumentalität und akademisch kontingentem Selbstwert und Angst vor Misserfolg erwarteten. Besonders relevant waren geschlechtsspezifische Unterschiede in den Assoziationen, da bisherige Forschung zeigte, dass der Zusammenhang zwischen einem maskulinen Geschlechtsrollen-Selbstkonzepten und Selbstwert bei weiblichen Personen stärker sein kann (Cate & Sugawara, 1986; Hirokawa & Dohi, 2007; Whitley, 1988). Strukturgleichungsmodelle zeigten die positiven Aspekte einer instrumentellen Orientierung, mit höherem Selbstwert und Hoffnung auf Erfolg und niedrigerem akademisch kontingentem Selbstwert und Angst vor Misserfolg. Der Zusammenhang zwischen Instrumentalität und globalem Selbstwert war bei Mädchen stärker; für Mädchen hing diese Assoziation auch indirekt mit einer geringeren Angst vor Misserfolg

zusammen. Diese Ergebnisse zeigen, wie ein instrumentelles Geschlechtsrollen-Selbstkonzept besonders für Mädchen vorteilhaft sein kann. Unsere Ergebnisse, dass Instrumentalität für Schüler und besonders Schülerinnen positive Auswirkungen haben kann, können die Literatur zur Feminisierung der Schule erweitern, da diese Maskulinität in der Schule meist negativ darstellt (Lyng, 2009).

Die Ergebnisse der drei empirischen Werke dieser Dissertation zeigen die Vorteile von Maskulinität und Nachteile von Femininität aus der Perspektive von pädagogischen Fachkräften, Peers und Schüler*innen. Aus diesem Grund wird in der Diskussion der vorliegenden Dissertation dargestellt, wie Gender als „status characteristic“ verstanden werden kann (J. Berger et al., 1972; Gerber, 2009), wobei Männlichkeit und Maskulinität mit mehr Status verbunden werden als Weiblichkeit und Femininität. Dies könnte als Widerspruch zu früheren Arbeiten zur Feminisierung der Schule verstanden werden, die die bessere Passung zwischen Femininität und Schule (Heyder & Kessels, 2013) darstellen. In der Diskussion werden praktische und theoretische Limitationen dieser Arbeit beleuchtet und neue Forschungsrichtungen dargestellt. Die Implikationen der Arbeit konzentrieren sich auf die Feminisierung der Schule und auf die Frage, ob eine Reduzierung von Gender Salienz in der Schule vorteilhaft sein könnte. Letztlich hebt die Arbeit hervor, wie aktuelle gesellschaftliche Bewegungen sich auf Gender fokussieren.

I. Introduction

Schools are foremost environments intended to foster academic achievement but are also spaces in which such achievements are understood. Social ties between teachers and students and among peers are forged, the achievements gained in school are interpreted, students' self-esteem and motivation are shaped and in turn can shape their future performances. In all of these areas, encompassing likability and perceived competence, attributions for success, as well as self-esteem, and achievement motivation, people may rely on gender and gender stereotyped behaviors and traits to make sense of their experiences. A prominent example is the observation that girls have attained high achievements the academic field over the past decades (O'Dea et al., 2018; OECD, 2023; Stoet & Geary, 2015; Voyer & Voyer, 2014), including in Germany since about the 1980s (Helbig, 2013). In response, researchers have asked whether school has become "feminized" (Brophy & Good, 1973; Budde, 2006; Heyder & Kessels, 2013) and whether boys are being left behind (Hannover & Kessels, 2011).

Enshrined in Germany's education guidelines is the goal of creating a balanced educational system for boys and girls (Kultusministerkonferenz, 2016). We still, however, see instances in which differences between students are created or exacerbated by gender and beliefs surrounding gender. A particular explanatory focus has been placed on the fit between masculine and feminine identities and the demands of school (Kessels et al., 2014). These works find that femininity and feminine stereotyped behaviors and traits are suited to the academic context, while masculinity and masculine behaviors and traits are said to stand in contrast to it (Heyder & Kessels, 2013, 2015, 2017). This dissertation intends to show that gender stereotyped behaviors and traits, in addition to gender, can inform how people make sense of their experiences in a school setting. In line with the gender self-socialization model (GSSM, Tobin et al., 2010), I separate gender identity from gender stereotypes and attribute self-perceptions (gender role self-concepts). Isolating whether an issue is related to a student's gender or perhaps to their adherence to masculine norms, for example, can allow us to

identify problems and opportunities more accurately. We can see the presence of gender stereotypes in the educational setting, such as in the beliefs that exerting effort is feminine (Heyder & Kessels, 2017) or that enactment of masculinity is incompatible with academic engagement (Heyder & Kessels, 2015). Not just gender stereotypes but also gender role self-concepts can play a role in academic life. Gender role self-concepts can relate to help-seeking behaviors (Kessels & Steinmayr, 2013), ability self-concepts (Wolter & Hannover, 2016), and overall school well-being (Korlat et al., 2022). This level of detail has allowed past researchers to show that concepts relating to gender, such as masculinity and femininity, can often be stronger predictors of school-relevant outcomes than gender itself (Cate & Sugawara, 1986; Orlofsky & Stake, 1981; Thomas & Blakemore, 2013; Whitley, 1988). These wide-reaching associations highlight the relevance of gender stereotypes and gender role self-concepts in the academic setting. In a balanced educational system, such as one aspired to by the German government, the influence of gender stereotyped traits and behaviors should be minimal, giving all students a chance to thrive academically and personally.

As stated, the present empirical works thus explore not just gender but also gender stereotypes and gender role self-concept. Furthermore, each paper questions the implications of masculinity and femininity in educational contexts and the broad assumption that behavior in line with gender stereotyped expectations is advantageous. Understanding the influence of gender and gender stereotyped traits and behaviors in school can also provide insight into the future of pupils. This could help to understand why we see girls performing well in school but have not seen this translate into exceptional vocational success (Hadjar et al., 2014; Steinmayr & Kessels, 2013).

This dissertation offers insight from multiple perspectives that are relevant in the school-context, these being the perspectives of educators, peers, and students themselves. The manuscripts feature self-evaluations, peer evaluations, and educators' evaluations of young

children. The three works complement one another, as they focus on different groups and utilize different methods; two of the works are experimental vignette studies, while the other uses self-reported data in a cross-sectional design. Each paper focuses on educational factors outside of academic achievement but relevant in educational settings, such as likability and perceived competence, attributions for success as well as self-esteem and achievement motivation. These outcome variables all have important relations to personal and social functioning (Fairlamb, 2020; Ferkany, 2008; Grow et al., 2016; Lavy & Naama-Ghanayim, 2020; A. J. Martin & Marsh, 2003; Moore & Smith, 2018; Pulles & Hartman, 2017; Steinmayr et al., 2019; Steinmayr & Spinath, 2009a). This research allows us to speak to how gender and gendered characteristics influence one's experience in the world – and whether it aligns with the goals of a more gender balanced educational system and society. While the works complement one another and offer common insight into the educational system in regard to gender and associated traits and behaviors, the present works were not conceptualized as a single work, each manuscript instead addressing a specific theoretical and empirical question.

This dissertation consists of multiple chapters, the first of which ("Theoretical Background") will outline some key definitions. I will first define the concepts of sex and gender before establishing the term "gendered characteristics" using terminology drawn from the GSSM (Tobin et al., 2010) to introduce and clearly delineate these key constructs. After these sections, I provide a theoretical lens in the section "Gender-Schema Theory," which stresses that information is organized and processed in gendered ways (Bem, 1981b). The subsequent section outlines how gender congruity and incongruity have been examined, focusing firstly on gender typicality, secondly on adherence and violations of gender stereotypes, and lastly on masculine, feminine, and androgynous gender role self-concepts. In the next chapter, "Feminization of School," I focus on the educational setting, highlighting the feminization of school hypothesis along the key perspectives of educators, peers, and the self,

showing how past research has used gender and gendered characteristics to organize information about academic functioning. The work then turns to constructs outside of academic achievement in my chapter "Gender and Gendered Characteristics in School," with sections on judgments of likability and competence, causal attributions, and self-esteem and achievement motivation. Within these sections, I focus on how each construct has been explored in regard to gender and gendered characteristics before highlighting potential gaps in the literature. In the last chapter of this introduction, I summarize the main manuscripts.

Theoretical Background

The present work defines gender and sex in line with past literature, which has drawn a distinction between the two terms. Sex is defined as the broad biological categories of males and females, encompassing a range of genetic, hormonal, and chromosomal differences (Unger, 1979). Gender is seen as the social category of male and female (Unger, 1979) or "differences between women and men that are produced socioculturally" (Hyde et al., 2019, p. 172). The complexity associated with the two concepts, sex and gender, is that gender is most often based on sex (Rudman & Glick, 2010), and their influence on people is difficult to disentangle from one another. This dissertation acknowledges that some of the social consequences of gender may be the result of differences in biological sex (Eagly, 1987; Eagly & Wood, 2012; Wood & Eagly, 2012). While acknowledging the possible role of biology in the origin of gendered beliefs, I do not address biological explanations for possible differences between boys and girls. The present work does not focus on sex but rather self-reported gender and its associated characteristics. For the purposes of this dissertation, gendered characteristics are constructs about the self and others that are related to gender, specifically referring to gender stereotypes and gender role self-concept.

In the first of the next three sections, I will introduce the concepts of gender and gendered characteristics using the language of the GSSM (Tobin et al., 2010), a cognitive

model exploring gender development. The model offers a clear separation between constructs that have been used interchangeably in literature (Tobin et al., 2010). Despite the solid base provided by the terms (gender identity, gender stereotypes, and attribute self-perceptions), I will at times depart from their definitions and instead outline gender self-categorization, gender stereotypes, and gender role self-concepts, giving an adapted definition of each construct.

Gender and Gendered Characteristics

The GSSM is a model intended to combine multiple theories into one comprehensive framework (Tobin et al., 2010, p. 601) in order to explore gendered cognition in childhood and beyond. In this model, the researchers follow theoretical outlines by Greenwald et al. (2002) and introduce three concepts: self, attribute, and group (in the case of the GSSM, the latter is the gender category). These concepts interact with one another and thus give rise to key associations, or *constructs*. The GSSM can boast many advantages as a model for childhood gender cognition, perhaps first and foremost being a stress on each individual's "unique perspective" of gender (p. 601). The authors deliberately stress how gender can and should be defined from the viewpoint of the individual (p.605). This highly individualized perspective can make measurements of the constructs complex. Each of the three constructs of the GSSM, gender identity, gender stereotypes, and attribute self-perception, is outlined below.

Gender Identity and Gender Self-Categorization

According to the GSSM, gender identity is the result of the connection between "the self and a gender category (e.g., "I am a girl")" (Tobin et al., 2010, p. 607). In traditional views, gender identity referred to knowledge of one's gender category (Kohlberg, 1966). Tobin and colleagues consider membership knowledge to be only one of five dimensions of gender identity. Further dimensions seen as central to gender identity are gender

contentedness, felt pressure for gender conformity, gender typicality, and gender centrality (Tobin et al., 2010, p. 608). In the GSSM, gender identity is ascertained through a rather complex method, which recognizes the above dimensions and the individual perspective of children in constructing gender. While this lets children express their gender identity according to their personal criteria, it is a rather complex measure to use and interpret.

Both Tobin and colleagues (2010) as well as others (Egan & Perry, 2001; Perry et al., 2019) have written about the multidimensionality of gender identity. In these works, the primary dimension of gender identity is the knowledge of one's gender membership, or gender self-categorization (Perry et al., 2019). Gender self-categorization indicates the group to which people understand they belong. This is similar to what the American Psychological Association (APA) refers to as gender identity. According to APA guidelines, gender identity relates to a person's "psychological sense of their gender," which is expressed by all individuals and may not reflect biological sex (American Psychological Association, 2020b, p. 138). Gender self-categorization as a dimension of gender identity is assessed rather simply by asking participants to indicate their own gender.

Overall, the construct of gender identity in the GSSM captures multiple dimensions, ranging from knowledge of one's membership in a gender category to one's feelings about this group and pressure to conform to expectations placed on the group (Tobin et al., 2010). In the present work, I focus on self-categorization as a measure of gender identity, referring to "gender" or "self-reported gender."

Gender Stereotypes

Within the GSSM, gender stereotypes are the result of the association between a gender category and attributes, such as "boys are strong" (Tobin et al., 2010, p. 604). The researchers outline that stereotypes pertaining to gender capture how the sexes *do* or *should*

differ (Tobin et al., 2010, p. 609). This distinction between *descriptions* and *prescriptions* is a key feature in the theoretical literature on stereotypes (Eckes, 2008; Rudman & Glick, 2010).

In the GSSM, it is recommended that gender stereotypes be regarded in domain and context specific manners. For example, believing boys should be more adept at math and science might not necessarily relate to also believing that boys are more aggressive and prone to violence (Tobin et al., 2010, p. 611). This distinction again highlights how the GSSM acknowledges individual perspectives in these constructs but also highlights its difficulty in implementation, as it requires the collection of complex, personalized data.

Descriptive, Prescriptive, and Proscriptive Stereotypes. In general, stereotypes "are beliefs about the characteristics, attributes, and behaviors of members of certain groups" (Hilton & Hippel, 1996, p. 240). Stereotypes thus group people into categories based on certain characteristics, like socioeconomic status, ethnicity, or gender (Moriizumi, 2018), and shape expectations about group members (Ellemers, 2018). Gender stereotypes, specifically, describe the "behaviors considered normal and appropriate for a male or female" (American Psychological Association, 2020a), a definition capturing an important feature outlined above: the typicality and desirability of characteristics. Stereotypes may thus be descriptive (typical) or prescriptive (desirable) (Eckes, 2008; Koenig, 2018; Nauts, 2015; Rudman & Glick, 2010). Gender stereotypes can also be proscriptive, a subgroup of prescriptive stereotypes, outlining behaviors which *should not* be engaged in (Prentice & Carranza, 2002; Rudman & Glick, 2010).

These categories of stereotypes are theorized to have originated for different purposes in social life. Descriptive stereotypes preserve cognitive resources when interacting with stereotyped individuals by automatically activating certain expectations (Eckes, 2008; Hilton & Hippel, 1996; Rudman & Glick, 2010). Prescriptive and proscriptive stereotypes can guide behavior of group members by dictating which behaviors should and should not be enacted,

preserving social hierarchies and justifying the existing system (Eckes, 2008; Hilton & Hippel, 1996; Rudman & Glick, 2010).

Content of Gender Stereotypes. Synthesizing the findings from multiple investigations on gender stereotypes showed that two fundamental dimensions underlie gender stereotypes, specifically for adults (Koenig, 2018; Prentice & Carranza, 2002; Rudman et al., 2012b). Stereotypes about women related to communality, and stereotypes about men related to agency (Koenig, 2018). Social, developmental, and personality psychologists have identified multiple dimensions that align closely with these feminine and masculine constructs (A. E. Martin & Slepian, 2021). The masculine and feminine dimensions (Bem, 1974, 1981a) are echoed in the constructs of instrumentality and expressiveness (Spence & Helmreich, 1980), competence and warmth (Cuddy et al., 2008; Eckes, 2002), and agency and communion (Abele et al., 2008; Bakan, 1966), respectively. Broadly speaking, the masculine dimensions (instrumentality, agency, competence) are oriented towards the self, independence, and competition, while the feminine dimensions (expressiveness, communion, warmth), relate to orientation towards others, nurture, and sensitivity (Abele & Wojciszke, 2007; A. E. Martin & Slepian, 2021; Oswald & Lindstedt, 2006).

These agentic and communal categories have been called the “big two” (Paulhus & Trapnell, 2008) of social cognition (Abele & Wojciszke, 2013) and are stated to underlie our social judgment. It has been posited that the relevance of these “universal” categories (Cuddy et al., 2008; Fiske et al., 2002, 2007) is so pronounced precisely because they relate so closely to gender, a highly salient social category that is also relevant for the self (A. E. Martin & Slepian, 2021). These categories emerge consistently across disciplines because they readily map onto gender, and we are primed to view our world through a gendered lens and interpret incoming information accordingly (A. E. Martin & Slepian, 2021). This point will be expanded upon in the section “Gender Schema Theory.”

Gender stereotypes have been measured for decades (Liben & Bigler, 2002; Zucker, 1977). One method of measuring descriptive, prescriptive, and proscriptive stereotypes was outlined by Prentice and Carranza (2002): They identified traits that differed significantly in ratings of desirability and typicality for men and women. Other researchers have measured stereotypes according to this method, with researchers introducing numerical cut-offs for stereotype ratings (Rudman et al., 2012b), which are currently used in practice (Koenig, 2018; Nauts, 2015; Sullivan et al., 2018, 2022). The field has established gender stereotypes for infants, toddlers, children, adolescents, adults, and the elderly according to these specific methods (Koenig, 2018; Sullivan et al., 2022). Masculine and feminine stereotypes for these different age groups have been captured in identical measures: by asking adult participants to rate the desirability and typicality of behaviors and traits for male or female infants, toddlers, children, etc. (Koenig, 2018; Sullivan et al., 2022) and applying the established numerical cut-offs (Rudman et al., 2012b). Interestingly, while the fundamental dimensions of agency and communion were found in research assessing stereotypes for adults, the dimensions did not seem to apply to gender stereotypes about children to the same extent (Koenig, 2018; Sullivan et al., 2022): gender stereotypes about children related more to appearance and preferences (toys, activities) than traits (Sullivan et al., 2022).

While this empirical assessment of stereotypes is useful for comparing groups (Koenig, 2018; Sullivan et al., 2022), I do not argue that characteristics need to be measured in this specific manner in order to be considered a stereotype. Traits that are associated with one gender more than the other constitute a gender stereotype, despite not being assessed according to this precise methodology. A key word here is “associated,” as I do not imply that characteristics need to actually be displayed gender differentially; they only need to be viewed as being gendered.

In summary, gender stereotypes in the GSSM are defined as "beliefs about how the sexes differ […] or should differ" (Tobin et al., 2010, p. 609), although the GSSM would stress the need for a more individualized method of collection than what is currently common in empirical practice. According to our present definition, gender stereotypes describe, prescribe, and proscribe behaviors and traits associated with male and female individuals.

Attribute Self-Perceptions

Within the GSSM, attribute self-perceptions are the result of the association between the self and attributes, such as "I am strong" (Tobin et al., 2010, p. 604). Specifically relating to gender, this association refers to the connection between one's perception of oneself to attributes seen as characteristic of the male and female gender groups (Tobin et al., 2010, p. 601). An issue that Tobin and colleagues describe is that past research has inferred gender identity from such attribute self-perceptions, which the GSSM consciously separate from one another.

In the traditional research vein on attribute self-perceptions, "a wide range of normatively gender-typed attributes" (Tobin et al., 2010, p. 612) is compiled by the researcher, which, when rated by participants, may offer insight into their instrumentality and expressiveness, for example. Given the strong focus on individual perspectives on gender, Tobin et al. (2010) caution against this method, arguing that such composite measures are reductive.

Gender Role Self-Concepts. I depart from the above definition, taking a traditional approach to attribute self-perceptions. I will focus on gender role self-concepts, which are understood as the application of gender stereotyped attributes to oneself (Wolter & Hannover, 2016). Gender role self-concepts capture the essence of Tobin et al.'s construct of attribute self-perception (cf. manuscript 3), but I prefer the former term, as it makes the focus on gender more explicit and because the term is more common among German-speaking

researchers and in German research institutions (Athenstaedt et al., 2009; A. Berger & Krahé, 2013; Kessels & Steinmayr, 2013; Wolfram et al., 2009).

Gender role self-concept refers to constructs that have also been labeled sex or gender role orientation (Bem, 1974; Starr & Zurbriggen, 2017) and capture one's sense of masculinity and femininity (Kessels & Steinmayr, 2013). Since masculinity and femininity closely align with other constructs (A. E. Martin & Slepian, 2021), gender role self-concept may also be said to capture instrumentality and expressiveness or agency and communion.

Measuring Gender Role Self-Concepts. Gender role self-concept has been measured in a variety of manners. One of the best-known measures of gender role self-concept, or sex typing, was originated by Sandra Bem (Bem, 1974, 1981a), whose seminal works informed many psychologists working on gender (Liben & Bigler, 2017). The Bem Sex Role Inventory (BSRI) was constructed by gathering desirable traits that participants associated with men or women. This process is similar to how gender stereotypes are gathered, and indeed, researchers succinctly summarized that "gender stereotypes are used not only to characterize others but also to characterize oneself" (Hentschel et al., 2019, p. 2), highlighting the similarity between gender stereotypes and gender role self-concept.

For the construction of the BSRI, Bem selected socially desirable items to represent masculinity and femininity ("ambitious" and "gentle," respectively) and further items which were seen as gender neutral ("happy") (Bem, 1974). Respondents on the BSRI can receive scores placing them into four categories: feminine (high femininity, low masculinity), masculine (low femininity, high masculinity), androgynous (high femininity, high masculinity) or undifferentiated (low femininity, low masculinity) (Bem, 1977; Spence et al., 1975).

Another important measure in this field is the Personal Attributes Questionnaire (PAQ) (Spence et al., 1974), which measures instrumentality and expressiveness. Like the

BSRI, the PAQ does not use a unidimensional scale but rather conceptualizes two scales, aligning with characteristics associated with masculinity and femininity (Fernández & Coello, 2010; Spence, 1991). The PAQ was constructed by gathering stereotypical characteristics (Spence, 1991) and asking participants to rate the typicality and desirability of these items for men and women. Spence explicitly used the terms descriptive and prescriptive stereotypes in outlining the procedure used to construct the scale (Spence, 1991), thus once again highlighting the close relation between gender stereotypes and gender role self-concepts. Characteristics showing large differences in typicality for women and men were selected, and ratings of desirability were used to ensure that items were considered desirable (Spence, 1991).

As beliefs about gender are not only historically (Eagly et al., 2020) but also culturally specific (Bosson et al., 2022), it is beneficial to use culturally specific measures where possible. Recognizing that ideas of gender may differ across cultures, for the purpose of this dissertation specifically between the United States and Germany (Wilde & Diekman, 2005), highlights the need for locally validated measures of gender role self-concept. Most important for the present work is a scale of gender-related traits by Kessels (2005), which was developed for use with adolescents (Kessels & Hannover, 2008). The method used to construct this scale is similar to the BSRI and PAQ, as traits were gathered and adolescent participants were asked to indicate how typical these are for male or female individuals. Using this data, a scale was constructed that views masculinity and femininity as separate dimensions, enabling respondents to receive a score reflecting their masculine and feminine gender role self-concepts. Importantly, for our purpose, this scale is in German and validated for use with an adolescent sample.

An interesting variation of self-concept scales was put forth by Berger and Krahé (2013), who specifically focused on including socially undesirable aspects of gender

stereotypical traits, yielding positive and negative masculinity and femininity scores. Negative masculinity encompasses traits like “inconsiderate” and “arrogant,” and negative femininity reflects traits like “naïve” and “anxious” (A. Berger & Krahé, 2013, pp. 523–524). I mention these German scales to highlight two aspects. The first is that culturally dependent scales of gender role self-concepts exist, and our present research recognizes their relevance. The second stresses the relation between gender stereotypes and gender role self-concepts.

In summary, attribute self-perceptions in the GSSM are defined as self-perceptions of characteristics associated with men and women. While the GSSM recommends focusing on unique perspectives to assess this construct, “traditionalist” researchers compile gender stereotypical characteristics and ask participants to indicate the degree of identification with the attributes. In the present work, I refer to this as gender role self-concept, which stresses the application of the attributes to the self and does not conflate gender role self-concepts and gender identity, a distinction Tobin et al. (2010) recommend.

In the following section, I will introduce gender schema theory as an overarching theoretical framework through which this dissertation may be viewed. This theory was not the subject of the three works that make up the empirical section of this dissertation but rather provides a context through which the three empirical works can relate to one another.

Gender Schema Theory

Following social learning theory (Bandura, 1977) and cognitive development theory (Kohlberg, 1966), in the early 1980s multiple researchers concurrently constructed gender schema theories as relevant for child development (Bem, 1981b; Liben & Bigler, 2017; Liben & Signorella, 1980; C. L. Martin, 2002; C. L. Martin & Halverson, 1981). These theories account for active construction of gender ideas by children instead of presenting children as absorbers of information (Bem, 1983; Canevello, 2020), as had been insinuated by previous theories (Bem, 1983; Liben & Signorella, 1980; C. L. Martin & Halverson, 1981). I will very

briefly outline Martin and Halverson's (1981) and Liben and Signorella's (1980) gender schema theories but focus primarily on the works of Sandra Bem (1981b, 1983, 1993). I will outline the cognitive view of gender schema theory but then also explicate how this theory has found wider, more social resonance and how this view in particular is useful for this dissertation.

Gender schema theory, as the name suggests, builds on the concept of schemata (C. L. Martin & Halverson, 1981; Piaget, 1952), which are defined as cognitive structures used to organize and process information (Bem, 1981b, 1983; C. L. Martin & Halverson, 1981). In Martin and Halverson's gender schema theory, the authors outline how "sex stereotypes" are internalized by children and how these stereotypes subsequently shape the child's perception, cognition, and actions (1981, p. 1119). In the work on gender schemata by Liben and Signorella (1980), the impact of stereotyping on cognition, specifically memory, is outlined. The authors state that through socialization (C. L. Martin, 2002), children learn to associate a wide variety of attributes with gender, using these associations to process new information (Liben & Bigler, 2017).

According to Bem, a gender schema is understood as a cognitive structure in which information is organized "according to the culture's definitions of maleness and femaleness" (1983, p. 603). Gender thus becomes a cognitive organizing principle (Bem, 1983), which guides memory, attention, and behavior for children and adults (Bem, 1981b; Leung, 2020). Bem states that children, highly attuned to the world around them, receive incoming information, organize it according to a gender schema, and then use it to present gender appropriate behavior (1983). The information received from the environment can be related to children's understanding of their social environment ("Boys should not wear hair clips") or to their understanding of themselves ("I, a boy, should not wear hair clips"); the gender schema will include associations relevant to personal and social life (Bem, 1981; Canevello, 2020).

Not only is gender *a* cognitive organizing schema, it is a *primary* schema (Bem, 1983), having far-reaching associations into many aspects of life. The fact that gender is a more powerful schema than religion or eye color, for example (Bem, 1983), is stated to be due to the biological differences between men and women. However, the gender schema network of most people reaches far beyond physiological markers of "anatomy [and] reproductive function" (Bem, 1981b, p. 354). People may form associations between masculinity and femininity and categories (Bem, 1981b) such as animals (eagle vs. nightingale; Bem, 1983), colors (blue vs. pink; Del Giudice, 2017), and school subjects (physics vs. music; Kessels, 2005).

Evidence for gender schematic processing was brought forth by Bem in cognitive studies, showing that people who scored highly in femininity and masculinity, indicating stronger sex-typing, were better at recalling words that linked to gender stereotypes (Bem, 1981b). While Bem's gender schema theory was presented as a theory of gender development for children, the theory finds uses outside of childhood. An investigation finds adults, not children or adolescents, as the participants in most research utilizing gender schema theory (Starr & Zurbriggen, 2017). Presently, this theory is also rarely used as an object of study itself, meaning researchers are not currently testing the veracity of gender schema theory (Koivula, 1995; Starr & Zurbriggen, 2017). Instead, researchers have removed gender schema from the specific cognitive account and use Bem's gender schema theory more broadly. Bem herself became less interested in the individualized view of gender schematicity, later broadening her focus to the wider (American) culture (Bem, 1993). More modern research uses the theory as an overarching construct in a variety of domains, including organizational, social, and developmental psychology (Starr & Zurbriggen, 2017). This is in line with how Bem herself has viewed her work, stating that her "research on gender schematicity is a strategy for exposing an invisible cultural lens," rather than being an empirical account of individual differences (Bem, 1993, p. 131) or dissecting the exact details of gender schematic

processing. This is also how the present dissertation will use gender schema theory, as I do not aim to test hypotheses derived from the theory. I focus on gender schema theory, understanding gender as a lens through which individuals view, interpret, and interact with their world (Bem, 1993; C. L. Martin, 2002; Nielson et al., 2020; Perry et al., 2019). This is in line with other work, in which it serves as an overarching framework. Pellegrini (2011), for instance, theorized that an emphasis on gender could lead adults to view male children as more aggressive, as focusing on gender could engage the gender schema, in which aggression is linked with masculinity. Recognizing gender as a fundamental category that structures our cognitions and behaviors makes gender schema theory, especially its broadest understanding, relevant for this present dissertation.

A variety of behaviors in the educational setting can be and have been linked with gender and gendered characteristics in the eyes of educators and students. Empirical evidence shows that we view actions like putting effort into attainment (Heyder & Kessels, 2017; Jackson & Dempster, 2009), behaving prosocially in class (Bouchard et al., 2015; Eisenberg et al., 2006; Quenneville et al., 2022), and engaging with certain toys (C. L. Martin et al., 1995; Todd et al., 2017; Weisgram et al., 2014) through a gendered lens, which, in turn, could impact how we view and behave towards the acting individual. The three manuscripts presented in this dissertation will investigate how self-reported gender, gender stereotypes, and gender role self-concepts can shape student's experiences in the academic context. Within the works presented here, it is stated that masculine stereotypes and gender role self-concepts should relate to masculine patterns of self-esteem and competence, for example. Feminine stereotyped behaviors, on the other hand, should relate to feminine patterns of causal attributions and prosociality, for example. In other words, masculine and feminine orientations and behaviors should activate associated constructs within the social gender schema. In her work, Bem attempted to "expose, not just the gender schematicity of sex-typed

people in particular but the gender polarization of American culture in general" (Bem, 1993, p. 126). This dissertation similarly focuses on not just the gender schemas of the individual but rather attempts to show the impact of gender and gendered characteristics within educational settings.

Gender In/congruence and Associated Outcomes

As established above, gender has a wide-reaching impact on our cognitions, behaviors, and attitudes towards ourselves and others. Historically and culturally, these views can shift (Bosson et al., 2022; Eagly et al., 2020) but there is a social incentive to adhere to existing gender rules, which is attained through implicit and explicit norm regulation. In this section, I focus on gender congruence, which refers to instances in which one's gender aligns with one's gender identity, gender stereotyped behaviors, and gender role self-concept. A person who behaves in ways that are seen as typical and desirable for their gender (gender stereotypes) and has an orientation in line with their gender (gender role self-concept) would be congruent in their gender stereotypes and gender role self-concept. I will briefly outline typical gender identity (gender typicality) and violations thereof, but focus more closely on adherence to gender stereotypes and violations thereof, and adherence to gender role self-concepts and androgyny, which represents a break with traditional gender role self-concepts.

Gender Typicality and Atypicality

Gender typicality is a term from the multidimensional view of gender identity (Egan & Perry, 2001) and does not refer to incongruity with one's gender self-categorization. Gender typicality refers to one's view of oneself as similar to same-gender others (Hoffman et al., 2019). Individuals who believed themselves to be gender typical reported better outcomes in terms of mental well-being (Vantieghem & Van Houtte, 2018), self-esteem (Skinner et al., 2018), acceptance by peers (Egan & Perry, 2001), and popularity (Jewell & Brown, 2014).

Gender typicality can also be associated with felt pressure for conformity, though this association is inconsistent (Carver et al., 2003; Egan & Perry, 2001; Skinner et al., 2018).

As gender typicality related to well-being (Vantieghem & Van Houtte, 2018), it is not surprising that gender *atypicality* is associated with detrimental outcomes. Adolescents low in gender typicality reported greater scores on measures of depression (Menon et al., 2013), social anxiety, and peer victimization (Smith & Juvonen, 2017). Low gender typicality was empirically linked to lower self-esteem (Egan & Perry, 2001; Menon et al., 2013; Skinner et al., 2018), this association being especially strong in girls (Corby et al., 2007; Skinner et al., 2018). Seeing oneself as similar to others in one's gender group emerges as an important part of mental well-being and social functioning.

Adhering to or Violating Gender Stereotypes

Adhering to gender stereotypes refers to female and male individuals behaving in ways that are stereotyped as feminine and masculine, respectively. Reactions to violations of gender stereotypes have been speculated to vary according to whether descriptive, prescriptive, or proscriptive stereotypes are being eschewed. Violations of descriptive stereotypes are theorized to be met neutrally or even positively (Eckes, 2008; Koenig, 2018), while violations of pre- and proscriptive stereotypes are theorized to be perceived as a threat to social order and thus to be reacted to with negativity or punishment (Eckes, 2008; Koenig, 2018; Rudman & Glick, 2010). This, according to Eckes (2008), is a unique feature of gender stereotypes, as other stereotyped identities (such as nationality, ethnicity, religion) do not elicit the punitive nature of pre-/proscriptive stereotypes. The empirical evidence for the differential reactions to violations based on stereotype category is weak (Nauts, 2015), with more evidence showing the strong correlation between descriptive and pre-/proscriptive stereotypes (Koenig, 2018; Prentice & Carranza, 2002; Rudman et al., 2012a; Sullivan et al., 2018), making a theoretical distinction practically impossible. Consequently, the empirical

investigations outlined here do not distinguish between violations of descriptive or pre- and proscriptive stereotypes.

Regarding adherence to stereotypes, adults were more likely to express positive emotions towards and give opportunities to a gender stereotypical other (Li & Wei, 2022). Adolescents described hypothetical popular students with more gender stereotypical terms than atypical terms, and students rated more gender typical peers, especially boys, as more popular (Jewell & Brown, 2014). Pre-adolescent students reported better attitudes towards a gender-conforming boy but indicated worse ratings of gender-conforming girls, highlighting a positive impact of masculinity and a negative effect of femininity (Braun & Davidson, 2017).

Studies show a contrasting pattern regarding violations of gender stereotypes. Adults expressed negative attitudes towards gender nonstereotypical individuals (Hernandez Bark et al., 2022; Li & Wei, 2022; Sanborn-Overby & Powlishta, 2020), such as women showing agency or dominance (Rudman & Glick, 2001) and men displaying modesty (Moss-Racusin et al., 2010). A hypothetical socially rejected adolescent was described in more gender nonstereotypical ways by adolescents (Jewell & Brown, 2014), although a more recent investigation showed that counterstereotypical behavior does not always incur negative effects from adolescents (Meimoun et al., 2023). Children reported higher rejection of counterstereotypical peers, especially for boys behaving in feminine ways (Blakemore, 2003; Braun & Davidson, 2017; Kwan et al., 2020).

Overall, violations of gender stereotypes were seen negatively by children and adults (Jewell & Brown, 2014; Kwan et al., 2020; Sanborn-Overby & Powlishta, 2020; Young & Sweeting, 2004). However, “there has been little systematic examination of the attitudes of adults concerning varying degrees of childhood gender nonconformity” (Thomas & Blakemore, 2013, p. 400). In work addressing this issue, adults reacted more negatively towards children who were described as violating gender stereotypes than gender stereotype

conforming children (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013), with results showing that participants reported liking the feminine boys the least (Coyle et al., 2016; Sullivan et al., 2018). The lack of findings from Europe underscores a need for further study.

Masculinity, Femininity, and Androgyny

A congruent gender role self-concept may be understood as male individuals scoring high in masculinity and female individuals scoring high in femininity, or as an individual scoring highly in masculinity or femininity while not scoring highly in the cross-gender orientations. I will first briefly outline the effects of male and female individuals scoring highly in masculinity and femininity, respectively.

Adolescent males high in instrumentality, showing a congruent self-concept, displayed greater self-esteem (H. Choi et al., 2010), fewer internalizing problems, and better social support (Exner-Cortens et al., 2021). In women, femininity, but also masculinity, related to better well-being (Matud et al., 2019) and did not relate to measures of aggression (Tolman et al., 2006). Women scoring highly in femininity showed no relation between this gender role self-concept and their self-estimated intelligence in multiple domains (Rammstedt & Rammsayer, 2002).

Such patterns remain consistent even when disregarding an individual's gender: masculinity is beneficial in men and women, while communion relates to fewer advantageous outcomes in people of both genders. Past research with adults showed that high agency related to career success, while communion, on the other hand, showed no significant correlation with objective or subjective vocational success (Abele, 2003). Masculinity consistently correlated with self-esteem, yet a feminine gender role self-concept only showed small effect sizes (Whitley, 1983) or did not correlate significantly with self-esteem (Whitley, 1988).

More recent work also revealed that self-esteem was more strongly predicted by agency rather than communion in several studies (Wojciszke et al., 2011). A similar pattern was found for adolescents and depressive symptoms, showing beneficial effects of masculinity and no effects of femininity (Barrett & White, 2002). The effects outlined here are by no means exhaustive; they illustrate the beneficial outcome of masculinity and more humble outcomes of femininity for both male and female individuals.

Despite the positive effects of masculinity outlined here, adherence to a subset of masculinities and femininities, specifically traditional masculinity and femininity, shows detrimental effects. Traditional femininity in adolescents can be characterized by attention to one's physical appearance and a norm of thinness and domesticity (Yu et al., 2020). This type of femininity was not associated with beneficial outcomes in school, and students endorsing this type of femininity were said to be academically at risk, showing poorer patterns of engagement and motivation (Yu et al., 2020). Similarly, adherence to traditional masculinity, defined by emotional stoicism and physical toughness, for example, also showed poor outcomes (Yu et al., 2020). This specific type of masculinity was associated with lower academic engagement in boys and girls (Rogers et al., 2017). The most adaptive patterns of behavior were displayed by students who rejected traditional gender roles of masculinity and femininity (Yu et al., 2020).

In the 1970s, Bem posited that the most beneficial gender role self-concept was not masculinity but androgyny, a disposition characterized by high scores in masculinity and femininity (Bem, 1974). Such a pattern was said to be most adaptive, as androgynous individuals possessed greater flexibility, managing to thrive in a variety of situational contexts (C. L. Martin et al., 2017; Pauletti et al., 2017; Woodhill & Samuels, 2003). The androgyny hypothesis found support specifically in research with children, adolescents, and young adults (N. Choi, 2004; Korlat et al., 2022; Pauletti et al., 2017) but has yielded mixed results for

adults (Antill & Cunningham, 1979; Signorella & Jamison, 1986; Whitley, 1983). Children benefited from androgyny, reporting better self-esteem and fewer internalizing tendencies than children of other gender role self-concepts (Pauletti et al., 2017). Androgynous adolescent students, compared to masculine, feminine, and undifferentiated students, had better outcomes in terms of happiness, optimism, engagement, and perseverance (Korlat et al., 2022). Notably, adolescents seemed to benefit more from the presence of femininity (as part of androgyny) than adults, in whom a masculine but not a feminine orientation associated with advantageous outcomes for well-being and success (Abele, 2003; Abele & Candova, 2007; Antill & Cunningham, 1979; Korlat et al., 2022; Whitley, 1984). Masculinity has typically been a correlate of positive outcomes for male but especially female individuals (Antill & Cunningham, 1979; Matud et al., 2019; Signorella & Jamison, 1986; Whitley, 1988).

In this section, I outlined how gender typicality and adherence to stereotypes display advantageous outcomes in the literature regarding well-being and interactions with others. An incongruous gender role self-concept, such as having an androgynous orientation, has also been posited as beneficial in the literature. Disregard for gender stereotypes is associated with negative outcomes, but a non-traditional gender role self-concept may be adaptive, especially for adolescents. In the following chapter, I will more closely examine adolescents and children, specifically focusing on the academic domain.

Feminization of School

One important metric in education is academic achievement, which is often understood as grades or school leaving certificates obtained by students but can also encompass results on standardized tests designed to measure competencies in specific subjects (Steinmayr et al., 2014). An influential meta-analysis of over 350 studies outlined that scholastic achievement of girls was higher than that of boys, as indicated by better grades

(Voyer & Voyer, 2014). These results are echoed in a further meta-analysis, which highlighted that STEM subjects showed smaller differences than non-STEM subjects but that overall female students attained higher grades than male students (O'Dea et al., 2018). Large-scale assessment of PISA data highlighted that girls outperformed boys academically internationally (Stoet & Geary, 2015) as well as in Germany (OECD, 2023).

With these developments, we have seen a strong focus on gender in educational settings from educational scientists, sociologists, and psychologists. The *feminization of school hypothesis* (Brophy & Good, 1973; Budde, 2006) has been presented as possibly accounting for the female advantage in terms of academic achievement. This hypothesis encompasses two main arguments (Budde, 2006; Heyder & Kessels, 2013): Firstly, it refers to the overrepresentation of female teachers, which has been investigated as having a possible negative impact on male achievement (Brophy & Good, 1973; Carrington & McPhee, 2008; Helbig, 2010). Secondly, the feminization of school references the better fit of female students to the demands and characteristics of school (Heyder & Kessels, 2013; Kessels et al., 2014; Steinmayr & Kessels, 2017). This present work will focus exclusively on the latter definition.

In the next section, I focus on three unique perspectives that are relevant in the educational context: that of educators, peers, and the self. These three perspectives capture key roles in academic life and are the focus of the empirical work of this present dissertation, with each manuscript focusing on one specific perspective. In each of the following sections, I will outline empirical findings showing how gender and gendered characteristics play key roles in academic functioning.

Educators

When it comes to the demands of school, teachers and high-achieving girls were said to be "on the same wavelength" (Myhill, 2002, p. 350). Interviews with teachers highlighted

that high-achieving girls were seen positively, as more gender typical, compliant (S. Jones & Myhill, 2004), and well-suited to the demands of school (Sáinz et al., 2021; Younger et al., 1999). Boys, compared to girls, were perceived more negatively in the academic context by teachers (Heyder & Kessels, 2015), who were more deficit-oriented towards boys, focusing on their issues, such as being more troublesome (S. Jones & Myhill, 2004) and disruptive (Sáinz et al., 2021).

Compliance and effort, factors strongly related to succeeding in academic environments (Anaya & Zamarro, 2023; Park & Kim, 2023) and valued by teachers (A. Orr, 2011), were associated more with femininity by teachers (S. Jones & Myhill, 2004). Teachers perceived students described as exhibiting effort as more feminine and typical to a girl than a student who did not exhibit effort (Heyder & Kessels, 2017). The stereotyping of effort as feminine also has an impact on boys: displaying low effort was associated with increased perceived similarity to a typical boy by teachers, although this did not extend to higher ratings of masculinity (Heyder & Kessels, 2017).

Positive classroom behavior, such as prosocial behavior, was associated more with female students by teachers (Quenneville et al., 2022). This association was also indicated by early childhood educators, even when no gender differences in actual prosocial behavior were measured (Bouchard et al., 2015). Other positive classroom behaviors, such as being respectful and disciplined, were associated more strongly with female than with male students, using an implicit association test completed by pre-service teachers (Glock & Kleen, 2017). The results from this IAT indicated an implicit association between male students and negative behaviors, such as being naughty and disrespectful (Glock & Kleen, 2017). Overall, it seems girls and feminine stereotyped behaviors are perceived in overwhelmingly positive manners in the academic context by educators, while boys and masculinity are linked with misbehavior and low achievement.

Peers

Academic achievement is seen as feminine to such an extent that boys succeeding in this domain reported higher rates of bullying than high-achieving girls and average-achieving boys (Bergold et al., 2020). The authors theorized that this association may be due to the lack of fit between scholastic achievement and the masculine gender role (Bergold et al., 2020). Male students, specifically, have spoken about the pressures of their peer groups and their attempts to avoid negative social consequences, such as bullying, which result from exhibiting academic diligence (Jackson & Dempster, 2009). Indeed, displays of effort were perceived to be linked with unpopularity and ridicule for boys, a finding accounting for some of the gender differences in displaying effort in school and, via this reduction of effort, subsequent worse academic achievement for boys (Workman & Heyder, 2020). Adolescents associated not displaying effort with masculinity and viewed it as more typical of male students in a vignette study (Heyder & Kessels, 2017). These students associated exhibitions of effort with femininity (Heyder & Kessels, 2017), highlighting that displaying effort aligns with the feminine gender role more than the masculine role in the eyes of peers. Overall, peers in school can have a strong influence over how academic achievement and school-related behaviors, such as showing effort, are experienced.

Self

Aside from the better achievements attained by girls (O'Dea et al., 2018; Voyer & Voyer, 2014), girls also reported better engagement and school belonging than boys (Huyge et al., 2015; Van Houtte, 2023). In an implicit test, adolescents associated "school" with "female" more strongly than with "male" (Heyder & Kessels, 2013). This presented a problem for boys specifically, as it, in conjunction with the endorsement of negative masculinity, related to lower grades (Heyder & Kessels, 2013). Students indicating higher levels of traditional gender roles showed more disruptive behavior (Van Houtte, 2023) and had lower school belonging than those with less traditional beliefs, with boys showing higher

levels of traditional gender role ideas than girls (Huyge et al., 2015). These findings show not only how boys and students with certain gender role beliefs fare in school but also highlight the interaction of gender and gender role self-concepts.

These outcomes can also relate to findings on gender congruity and achievement. Boys who displayed some gender incongruous behaviors, such as parting with traditional attitudes, attained better GPAs than boys who were more traditional (Yavorsky & Buchmann, 2019). Some feminine qualities thus seem to be associated with academic benefits for boys. Researchers have investigated exactly which personality traits are associated with academic success. Unsurprisingly, intelligence was positively related to academic success, as were conscientiousness, need for achievement, and agreeableness, while need for aggression was negatively related to achievement (Steinmayr & Kessels, 2017). In all of these traits, except for the need for aggression, in which they scored lower, girls scored higher than boys. Gender differences in academic achievement were mediated by these traits, supporting the fit between feminine personality traits and success in school (Steinmayr & Kessels, 2017).

Findings like this were called into question somewhat by research in which students indicated the degree to which they believed certain traits relating to academic achievement to be typical of girls or boys (Verniers et al., 2016). While traits like compliance and effort/intelligence (a single factor) were associated more with girls, the factor assertiveness, also deemed key for success in school, was associated more with boys (Verniers et al., 2016). This view is endorsed by an ethnographic study in which a wider variety of masculine and feminine identities were explored and negative attitudes towards school were exhibited by not just the “laddish” or macho male students but also highly feminine girls and more masculine girls (Lyng, 2009).

A few studies call the feminization of school into question (Lyng, 2009; Verniers et al., 2016). However, traits more pronounced in girls relate to academic success (Steinmayr &

Kessels, 2017), and adolescents associate femininity with school (Heyder & Kessels, 2013). The better fit between girls and femininity to school is thus empirically and theoretically supported, but gaps in this research remain.

Gender and Gendered Characteristics in School

A lot of research in the academic field has a focus on achievement, but this is a restricted view of the wide-reaching impact of school. Educational development relates to many factors aside from performance. How we understand and explain achievement can impact students' experience in school and beyond (Grow et al., 2016; Harvey et al., 2014). School also aims to foster students' well-being, developing a healthy sense of self-worth and motivation (Ferkany, 2008; Meece et al., 2006), setting students up for future success. The following section will thus outline factors relevant to academic life, beginning with likability and competence, key variables explored in manuscript 1. This section is followed by an overview of causal attributions, a construct explored in manuscript 2. The section concludes by describing research relating to self-esteem and motivational constructs, which are studied in manuscript 3. In each section, I will briefly define the examined construct before outlining research findings relating to gender differences, followed by research regarding gender stereotypes and gender role self-concepts.

Likability & Competence

The following section will introduce both likability and competence, which are key areas of focus in the literature on gender stereotype violations in children, which is the focus of manuscript 1 of this dissertation. As these constructs primarily, though not exclusively, relate to perceptions by others, I will highlight the views of educators and peers more closely.

Likability

In the educational and developmental literature, likable individuals are viewed favorably (Pulles & Hartman, 2017) and are, simply put, well-liked by others (Heyder &

Kessels, 2017; van der Linden et al., 2010). Likability has been associated with prosociality (Lu et al., 2018; Sandstrom & Cillessen, 2006; van der Linden et al., 2010), extraversion (van der Linden et al., 2010), agreeableness (Ciarrochi & Heaven, 2009; van der Linden et al., 2010), and friendliness (Denham & Holt, 1993).

In academic contexts, high-achieving female students were rated as more likable than high-achieving male students in a vignette study with adolescent respondents (Quatman et al., 2000). A peer nomination study with young adolescents found a gender difference in favor of girls, who were considered more likable than boys (Lu et al., 2018). Other studies conducted with children and adolescents, however, show no significant difference in likability between boys and girls (Engels et al., 2016; Jewell & Brown, 2014; Witvliet et al., 2010). Overall, there may be a female advantage in terms of likability, though not all investigations find an effect of gender.

Likability does show associations with gender stereotypes and gender role self-concept. Likability is highly social and related to positive feelings from others. This is similar to communion, which is characterized as an orientation towards other people and caring (A. E. Martin & Slepian, 2021). The connection between the two constructs has been laid out explicitly, with research showing that likability is influenced more by communion than agency (Wojciszke et al., 2009). The feminine dimension of warmth is so closely related to likability that early investigations initially measured warmth as likability (Fiske et al., 1999), and warmth is assessed with items like “likable” and “good-natured” (Eckes, 2002, p. 103).

Gender nonstereotypical behavior has been explored with likability, finding that hypothetical socially rejected adolescents were described as more gender nonstereotypical compared to popular targets, an effect driven by atypical boys. Furthermore, gender stereotypicality correlated positively with liking by peers; this correlation was also significantly stronger for boys compared to girls (Jewell & Brown, 2014). When confronted

with gender stereotypical and atypical targets in a vignette, adolescent students perceived feminine targets as more likable, regardless of gender (Meimoun et al., 2023). This effect was unexpected; in line with previous research, the researchers hypothesized backlash towards gender atypical targets, specifically for feminine boys.

Backlash towards gender nonstereotypical individuals has been found in adults (Hernandez Bark et al., 2022; Li & Wei, 2022; Rudman et al., 2012b) and adult perceptions of young children (Coyle et al., 2016; Sullivan et al., 2018). The latter studies revealed gender atypical children to be rated lower in likability than their gender typical counterparts (Sullivan et al., 2018). Feminine girls were perceived as the most likable (Coyle et al., 2016; Sullivan et al., 2018), and it was specifically feminine boys who received strong backlash (Sullivan et al., 2018), despite likability being associated with feminine orientations (Wojciszke et al., 2009) and stereotypes (Asbrock, 2010; Eckes, 2002; Fiske et al., 1999). These findings, particularly that femininity is punished in male individuals, are stated to be due to the social status of masculinity in society (Feinman, 1981) and the maintenance of gender hierarchy (Eckes, 2008; Hilton & Hippel, 1996; Rudman & Glick, 2010). It has not been established whether similar reactions to gender stereotype violations by children can be found in populations of adults who regularly interact with children. The present dissertation investigates this specific sample in manuscript 1, which aims to establish differences in perceived likability in gender typical and atypical targets in a vignette study with pre-service pedagogical educators from Germany, thus expanding on the primarily US-American literature.

Competence

Competence has been defined broadly as the ability to take charge of one's life and enact change in one's environment but also more specifically as one's skill in specific areas (American Psychological Association, 2018). Competence encompasses skill, efficacy, creativity, and intelligence (Cuddy et al., 2008; Fiske et al., 2007, p. 77). Along with warmth,

it is a universal dimension in social cognition, making it one of the first characteristics we note about people (Dupree & Fiske, 2017; Fiske et al., 2007). It is thus something one possesses and a dimension noted by others about individuals and groups (Fiske et al., 2007).

Perception of competence was not related to gender in a sample of young children (Jambunathan & Hurlbut, 2000), nor did teachers of pre-adolescents note gender differences in competence in their students (Granleese et al., 1989). Differences in specific academic competencies highlight that girls outperformed boys on reading competencies, while boys scored better in mathematics, although the gap in the latter domain was smaller than in reading competencies (Hadjar et al., 2014; OECD, 2023). Overall, specific competencies may display gender differences, although this dissertation is more concerned with a broad understanding of competence, which does not currently show strong gender differences.

Compared to differences in actual competence, ascriptions of competence have, however, been linked with the male gender. Competence-related beliefs in math and sports, masculine domains, showed gender differences, with boys believing themselves to have higher competence compared to girls (Fredricks & Eccles, 2002). This mirrors results showing boys to have stronger self-concepts in multiple domains of intelligence than girls, even when controlling for actual intelligence (Steinmayr & Spinath, 2009b). More broadly, "men have been thought to be more competent than women" (Hamilton & Lordan, 2023, p. 2), and men were ascribed greater competence than women in past work (Asbrock, 2010). Ascriptions of competence have shifted with time, and more contemporary work found women were being ascribed equal (Hentschel et al., 2019) or greater competence (Eagly et al., 2020) than men. Ascribed competence is not (necessarily) reflective of actual competence, indeed this ascribed competence more accurately reflects gender stereotypes than actual gender differences in competence.

Since competence is strongly associated with masculinity (A. E. Martin & Slepian, 2021), masculine people are generally seen as more competent. Some gendered patterns are found in adolescents; for instance, masculine participants rated themselves higher in social and physical competence, an effect stronger for female adolescents (Cate & Sugawara, 1986). Androgynous adolescents, who are high in masculinity as well as femininity, showed the highest levels of perceived scholastic competence (Rose & Montemayor, 1994).

Not just gender role self-concept but also gender stereotyped behaviors have been examined in relation to competence in adolescents. Gender nonstereotypical targets were rated as more competent than stereotypical targets by adolescents, although this was only significant in participants with high SES (Meimoun et al., 2023). In younger children, evidence on perceived competence and gendered characteristics has not been not straightforward. One study manipulated positive and negative masculinity and femininity, constructs previously outlined in the section "Measuring Gender Role Self-Concepts," in vignettes and found that adults rated girls with positive femininity as most competent, while girls enacting positive masculinity and boys showing positive femininity were seen as rather competent, more so than a boy showing positive masculinity and a boy displaying negative femininity (Coyle et al., 2016). These findings suggest a degree of acceptance towards nonconformity but are not clear in regard to ratings of perceived competence in targets of young children, which should be explored further. Notably, these studies on competence and gender stereotyped behaviors were conducted with adults, not educators. To the best of our knowledge, early childhood educators and their attitudes towards competence and child gender stereotyped behavior has not been explored, a gap addressed in the present dissertation. In manuscript 1, we investigate whether pre-service pedagogical educators rate vignettes describing gender stereotypical and nonstereotypical children differently in terms of likability and competence, as well as expanding the literature to other variables of interest.

Causal Attributions

I will now outline research on causal attributions, as this is a key focus in manuscript 2, in which the effect of gender stereotyped behavior on causal attributions is examined empirically. Attributions are "perceptions of causality" for certain outcomes (Weiner, 1972, p. 203). People ascribe different causes to successes and failures, which can have an impact on how individuals understand these events (Heider, 1958; Weiner, 1986). Attributions are relevant in school, as they reflect how students make sense of their own achievement, shaping how responsible they feel for their performance (Grow et al., 2016). For our purpose, the most relevant attributions are ability and effort attributions. Ability and effort, while both conducive to scholastic success, are often presented in opposition to one another. They are believed, implicitly, to work in a compensatory manner: someone low in ability needs to exert great effort to complete a task, while a highly able person would be able to complete the task without much effort (Heider, 1958; Möller, 2005).

Despite causal attributions being presented as relating primarily to the self, it is also insightful to regard attributions through others (Weiner, 2000). Some researchers have used both educator and peer perspectives to investigate how attributions may be made in school (Butler, 1994; Schoneveld & Brummelman, 2023). Causal attributions have emotional correlates, or indirect attributional cues (Graham & Williams, 2009), meaning that a teacher attributing a student's failure to lack of effort would likely display anger while showing pity to a student whose failure they believe to be due to a lack of ability (Hareli & Weiner, 2002; Weiner, 2000). This process is social, encompassing a recipient (a student), a person showing an emotional correlate of an attribution (educator), and witnesses to the exchange, also passing attributional judgments (peers). Educators and peers thus become important parts of the attributional system.

Psychologists have investigated gender in relation to attributions with great interest (Weiner, 1972). Studies on gender differences in attributions revealed a distinct pattern of "male-favoring," "female-derogating" attributions in adults (Hill & Augoustinos, 1997). While the occasional study finds no gender effects on how boys and girls attribute their successes and failures (Yailagh et al., 2009), many studies show gender differences in attributional patterns (Assouline et al., 2021; Bornholt & Möller, 2003; Lohbeck et al., 2017; see Meece et al., 2006; Mok et al., 2011). Overall, girls reported attributing their successes to higher effort and their failures to a lack of ability. Boys, on the other hand, attributed successes to their greater ability, while failures were due to a lack of effort (Assouline et al., 2021; Bornholt & Möller, 2003; Lohbeck et al., 2017; Meece et al., 2006; Mok et al., 2011).

The denigrating attributional pattern relating girls' achievements to effort rather than innate ability is further ingrained by findings showing that teachers made similar attributions about their students. Especially in mathematics, teachers were susceptible to the bias that girls' achievements were more likely to be due to effort rather than ability (Espinoza et al., 2014; Fennema et al., 1990; Tiedemann, 2000), while boys' achievements were believed to be a result of ability rather than effort (Espinoza et al., 2014; Fennema et al., 1990).

Peers are an integral part of the classroom and witnesses to the attributional process (Grow et al., 2016). Interviews with students reveal that boys claimed girls earn good grades through their effortful learning, while they attempted to downplay their effort to increase their ascribed ability (Jackson & Dempster, 2009). Quantitative evidence in peer attributions along gendered lines is more rare, but existing studies have found no differences in perceived attributions between girls and boys, contrary to the researchers hypotheses (Grow et al., 2016; Kisfalusi et al., 2019).

Causal attributions offer insight into how scholastic outcomes are interpreted and are thus an interesting factor to investigate with gendered characteristics, specifically gender

stereotypes, which also shape social cognitions and expectations for groups (Reyna, 2000). By adolescence, students are already aware of the compensatory nature of attributions (Barker & Graham, 1987; Folmer et al., 2008; Möller, 2005), believing that students with high effort are likely low in ability (Möller, 2005). In interviews, boys spoke about how not displaying effort serves as an alternate explanation for possible failure rather than suggesting a lack of ability (Jackson, 2003; Jackson & Dempster, 2009). In other words, not displaying effort not only enhanced one's masculinity but also enabled a self-serving causal attribution.

Students could thus be engaging in certain behaviors in order to enable advantageous attributional patterns in the eyes of others. Thus far, empirical investigations of this theoretical outline are rare. There is one study in which enactment of masculine stereotyped behavior was related to advantageous attributional outcomes (Kessels & Heyder, 2020). In this experimental vignette study, failing students behaving unobtrusively or disruptively were presented, and participating adolescents were asked to evaluate how teachers would react to the target students, as well as to indicate the causal attribution for the poor performance. The theoretical outline of this study proposed that misbehavior, stereotyped as masculine, would elicit an expectation of reprimands from the teacher, which is a reaction related to a causal attribution of low effort (Hareli & Weiner, 2002; Weiner, 2000), making a lack-of-effort attribution more likely than a lack-of-ability attribution (Kessels & Heyder, 2020). Notably, the expected reaction by the teacher mediated the effect of the stereotyped behavior on the attributions of lack-of-effort. Behaving in this masculine manner made ascriptions of low effort rather than low ability, a more masculine attributional pattern, more likely. Whether feminine behavior in classrooms would make more feminine attributional patterns of greater effort but lower ability more likely is explored for the first time in this dissertation. Research in adults supports the claim that feminine, warm behavior can lead to ascriptions of lower competence (Kervyn et al., 2012, 2016). This, however, is not the only theoretical link

between positive classroom behavior and lower ascribed competence. An area of research has highlighted how *praise* can lead to *negative* effects on peer evaluations. The "paradox of praise" describes the negative outcomes that receiving praise can have on an individual (Brummelman et al., 2016; Möller, 2005), as praise is an indirect attributional cue of high effort (Covington & Omelich, 1979; Weiner & Kukla, 1970) rather than high ability. We investigate this possible pattern of causal attributions in manuscript 2. We will examine whether desirable, prosocial classroom behavior, stereotyped as feminine (Eisenberg et al., 2006; Piché & Plante, 1991; Quenneville et al., 2022), will relate to more ascriptions of effort compared to ability, following teacher praise. Finding an effect could partly explain the female-derogating attributions (Hill & Augoustinos, 1997), seen more often in girls in investigations (Assouline et al., 2021; Bornholt & Möller, 2003; Lohbeck et al., 2017; Meece et al., 2006; Mok et al., 2011).

Self-Esteem, Academic Contingent Self-Esteem, and Achievement Motivation

I now turn to self-esteem, academic contingent self-esteem, and achievement motivation, which are key variables explored in study 3 of this dissertation, which cross-sectionally examines gender role self-concept with these constructs. In the following, I will highlight research primarily relating to the self.

Self-Esteem

Self-esteem is defined as a positive view of the self or sense of self-worth (Rosenberg, 1965). Individuals with high self-esteem are said to have positive self-regard for themselves across multiple contexts and are generally satisfied with their being (Schöne & Stiensmeier-Pelster, 2016, p. 10).

While some studies find null results (Freudenthaler et al., 2008) or girls scoring higher than boys on self-esteem (Szcześniak et al., 2022), most findings display that boys reported higher self-esteem than girls, especially in adolescence (Bachman et al., 2011; Bleidorn et al.,

2016; Kling et al., 1999; Schöne et al., 2015; Schöne & Stiensmeier-Pelster, 2016; Twenge & Campbell, 2001; see Zeigler-Hill & Myers, 2012; Zuckerman et al., 2016). Meta-analyses of multiple age ranges showed significant gender differences in adolescence, with boys scoring higher in self-esteem (Kling et al., 1999; Twenge & Campbell, 2001).

Investigations taking not just gender but gendered characteristics, such as gender role self-concept, into account found higher global self-esteem in adolescents scoring higher in masculine self-concepts (H. Choi et al., 2010; Holland & Andre, 1994). Instrumentality and masculinity related to greater self-esteem for male (H. Choi et al., 2010; B. Orr & Ben-Eliahu, 1993) and female adolescents (Cate & Sugawara, 1986; Holland & Andre, 1994), which can be stronger than the relationship between gender and self-esteem (Holland & Andre, 1994). It is theorized that this relation is in part due to the high social standing of instrumental traits, which increase self-esteem by being desirable and held in high regard by others (H. Choi et al., 2010; B. Orr & Ben-Eliahu, 1993). Indeed, the link between masculine gender role self-concepts and self-esteem is so pronounced it has also been theorized that they capture the same underlying factor (Marsh et al., 1987; Whitley, 1988).

Academic Contingent Self-Esteem

The construct of contingent self-esteem describes the extent to which self-esteem is drawn from certain domains such as familial relations, physical appearance, and/or academic achievement (Moore & Smith, 2018). Our interest is on contingencies of self-worth relating to academic life, thus I will focus on investigations specifically targeting academic contingent self-esteem, which is typically assessed using self-report measures. Having high academic contingent self-esteem indicates that one's sense of self-worth is dependent on one's academic performance (Crocker & Wolfe, 2001).

While some findings on academic contingent self-esteem found no differences between male and female adolescents (Chen-Bouck & Patterson, 2016; Wouters et al., 2013),

others show that girls scored higher than boys (Schöne et al., 2015; Schöne & Stiensmeier-Pelster, 2016; van der Kaap-Deeder et al., 2016). The finding that girls were more likely to attach their sense of self-worth to their academic performance is concerning, despite the better achievement attained by girls on average. The association of academic contingent self-esteem with certain detrimental outcomes (lower self-esteem, greater depressive symptoms) were stronger for female students compared to male students (Crocker et al., 2003; Schöne et al., 2015), highlighting the relevance of gender in this domain.

While the evidence of gendered effects for academic contingent self-esteem showed a rather clear pattern to the detriment of girls, there is a lack of research on the role of gendered characteristics, specifically gender role self-concepts. In fact, studies in which academic contingent self-esteem and gender role self-concepts are evaluated in conjunction are very limited, and indeed to my knowledge, no studies with adolescents have been conducted. One study in which these constructs were explored jointly was conducted with young women in university, finding no relation between masculinity or femininity and a sense of self-worth contingent on academic performance. (Mandal & Moroń, 2019). I will thus instead offer a theoretical link between academic contingent self-esteem and specifically gender role orientation (cf. manuscript 3). Masculinity could be expected to relate negatively to contingent self-esteem, because masculinity relates positively to global self-esteem (Hirokawa & Dohi, 2007; Sharpe et al., 1995; Whitley, 1983, 1988), which in turn relates negatively to contingent self-esteem (Moore & Smith, 2018). Furthermore, the link between femininity and academia has been established above (Heyder & Kessels, 2013) and would suggest that performing well in this domain should be less important to people with a masculine orientation.

Achievement Motivation

A person's achievement motivation is defined as having two key components (McClelland et al., 1953): fear of failure and hope for success. The former may be characterized by more negative emotions and the need to avoid situations highlighting potential inadequacies, while the latter encompasses more positive emotions and a desire to achieve favorable outcomes (Steinmayr & Spinath, 2009a).

Research on achievement motivation characterized by hope for success and fear of failure has shown interesting gender differences. While women generally scored higher on fear of failure than men (Abdi Zarrin et al., 2020; Severiens & Dam, 1998), the gender difference on hope for success was not as strongly empirically supported (Engeser, 2005; Lang & Fries, 2006; Stoeber & Rambow, 2007). School students showed similar patterns, with nonsignificant gender differences in the domain of hope for success but significantly higher fear of failure in girls compared to boys (Steinmayr & Spinath, 2008; Wach et al., 2015).

Achievement motive has been investigated in conjunction with gendered characteristics, specifically gender role self-concepts. Masculinity has been examined as a related factor: as it is associated with agency, encompassing ambition and self-determination, masculinity should relate to beliefs in one's ability to attain successful outcomes, suggesting greater hope for success and lower fear of failure. One investigation revealed a strong negative association between masculinity and fear of failure, finding that adults' gender role self-concepts were more strongly related to achievement motivation than gender (Orlofsky & Stake, 1981).

Achievement motivation was also related to gender role attitudes held by Chinese students, finding traditional attitudes in male students associated with a positive, hopeful pattern on achievement motivation. The same could not be said for female students, whose

traditional attitudes were associated with negative achievement motivation (Yang & Gao, 2021). One investigation on adolescents found mastery orientation, an adaptive desire to learn new skills and information in challenging situations in a similar vein to hope for success, to be associated with both communion and agency; the authors note, however, that agency is a more adaptive orientation (Strage, 1997). Overall, it seems that more masculine orientations related to better achievement motivational patterns, specifically lower fear of failure (Orlofsky & Stake, 1981). Feminine orientations were rarely related to beneficial achievement motivation (for an exception, see Strage, 1997).

In summary, studies relating to self-esteem, academic contingent self-esteem, and achievement motive have explored gender differences thoroughly, but there is a lack of current research on the effect of gendered characteristics, specifically gender role self-concepts on these domains. The third manuscript in this dissertation will address this gap, investigating the relation between gender role self-concept, specifically masculinity, with self-esteem, academic contingent self-esteem, and achievement motivation.

Summary and Study Overview

In the previous sections, I outlined the effects of gender, gender stereotypes, and gender role self-concepts regarding the constructs of likability and competence, as well as attributional patterns, specifically focusing on effort and ability, in addition to self-esteem, academic contingent self-esteem, and achievement motivation. These sections highlight that these constructs have been largely well-examined in terms of gender, but the research on gendered characteristics and these constructs is often lacking or outdated. The previous section also shows how wide-reaching the gender schema network can be. Effort is seen as feminine (Heyder & Kessels, 2017; Workman & Heyder, 2020), while competence and an adaptive achievement motivation are linked with masculinity (A. E. Martin & Slepian, 2021; Orlofsky & Stake, 1981; Rose & Montemayor, 1994; Yang & Gao, 2021). The consequences

for behavior, cognitions, and attitudes of an active gender schema (Bem, 1981b) can be seen in the educational context in educators, peers, and students themselves in the section "Feminization of School". I also point to an interesting feature of the literature: femininity has been presented as a good fit for school (Heyder & Kessels, 2013; S. Jones & Myhill, 2004; Steinmayr & Kessels, 2017), while masculinity is seen as burdensome in this environment (Glock & Kleen, 2017; S. Jones & Myhill, 2004). However, masculinity is associated with tremendously beneficial outcomes, such as greater perceived competence (Koenig et al., 2011; A. E. Martin & Slepian, 2021), greater self-esteem (H. Choi et al., 2010; Holland & Andre, 1994; Whitley, 1983, 1988), and advantageous achievement motivation (Orlofsky & Stake, 1981; Yang & Gao, 2021), while femininity was not associated with benefits for some academically relevant constructs, such as self-esteem (Cate & Sugawara, 1986; Holland & Andre, 1994; Whitley, 1988) and achievement motivation (Orlofsky & Stake, 1981; Yang & Gao, 2021). This dissertation is critical of the way femininity and masculinity have been presented in school, questioning the overarching assumption that, violations of gender stereotypes are necessarily negative (manuscript 1), feminine stereotyped behaviors are overwhelmingly beneficial (manuscript 2), and masculinity is detrimental in the academic context (manuscript 3).

Previous sections on gender in/congruity have indicated that, in general, adhering to gender stereotypes shows advantageous outcomes, while violating gender stereotypes incurred backlash. This was especially true for male individuals engaging in feminine behavior, who received the strongest negative reactions (Blakemore, 2003; Braun & Davidson, 2017; Coyle et al., 2016; Kwan et al., 2020; Sullivan et al., 2018), stated to be because violations of stereotypes by men and boys undermine the gender hierarchy (Feinman, 1981). Key variables of interest in this field are likability and competence: the former showing strong effects, with gender nonstereotypical children receiving backlash (Coyle et al., 2016; Sullivan et al., 2018), the latter showing less straightforward or inconclusive results

(Coyle et al., 2016; Sullivan et al., 2018). This research field is further complicated by the fact that a majority of work was conducted with US American samples (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013), which showed backlash towards gender atypical children. More recent European research, however, although investigating a different sample, indicated some acceptance of gender stereotype violations (Bochicchio et al., 2019; Meimoun et al., 2023). This dissertation aims to clarify and expand these findings. We are particularly interested in how adults view violations of gender stereotypes in children. Adults have a guiding function for children's development of gendered behavior in both social cognitive theory (Bussey & Bandura, 1999) and gender schema theories (Bem, 1981b; C. L. Martin & Halverson, 1981), and this dynamic has been overlooked (Thomas & Blakemore, 2013). The literature typically draws a random sample of adults (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013), some of whom may never have interacted with children in a meaningful way and may not intend on doing so. In our first manuscript, we draw upon a more relevant sample and investigate attitudes of future educators who will play a major role in children's lives. We first collected gender stereotypes held about young children in Germany, assessing descriptive, prescriptive, and proscriptive stereotypes in line with established procedures (Koenig, 2018; Nauts, 2015; Rudman et al., 2012b; Sullivan et al., 2018, 2022) with a sample of $N = 397$ university students. Using insight from this first study, we constructed vignettes depicting feminine and masculine girls and boys. Using a 2 x 2 between-participants design, we recruited $N = 414$ pre-service pedagogical educators to take part in our experimental vignette study. We assessed likability of the target child as well as perceived competence and further variables like creativity and self-esteem. Our hypotheses stated that we would find backlash in liking towards gender nonstereotypical targets, especially for a feminine boy. We further expected to find a difference in perceived competence between stereotypical and nonstereotypical children but did not specify a direction. We also hypothesized that gender nonstereotypical children would be perceived as

more creative and as having higher self-esteem, broadening the research to include the possibility of positive associations with counterstereotypical behavior. The works in manuscript 1 expand the current literature by establishing stereotypes in a non-US American sample and examining whether relevant adult socializers display backlash towards gender (non)stereotypical children.

While *violations* of stereotypes can incur backlash, the following work questions whether behavior *in line with* gender stereotypes can lead to backlash in the educational setting. Specifically, we question whether prosocial behavior, stereotyped as feminine, can lead to a denigrated attributional pattern. It has been established above that there is a lack of evidence connecting gender stereotyped behavior to attributions, specifically attributions by peers. One paper exploring gendered behavior from an attributional perspective using an experimental vignette study has revealed positive attributions resulting from expected teacher reactions to problematic behavior (Kessels & Heyder, 2020); behavior that is observed more in boys and which has even been described as "an enactment of masculine gender identity" (Heyder et al., 2021, p. 61). Whether positive, feminine behavior in school could result in a disadvantageous pattern of attributions via expected teacher reactions is the subject of manuscript 2. Based on research on praise and attributions, study 2 sought to investigate whether prosocial behavior of academically successful students leads to greater attributions of effort and lower ability, mediated by expected teacher reaction. The design, procedure, and analytical plan of study 2 was based on work by Kessels and Heyder (2020). We constructed vignettes of academically successful students, who were presented as being either prosocial or nondescript. We utilized a 2 x 2 design, with behavior of student (prosocial, nondescript) varying within participants and gender of portrayed student (boy, girl) varying between participants. In this study, ninth grade students (N = 324) were asked to indicate whether they expect a teacher to react to the students in the vignettes with praise and happiness, as well as the degree to which they believe the target students' grades to be a result of effort and ability.

Participants further reported the perceived femininity and masculinity of the students in the vignette, as well as providing other information about perceived intelligence and similarity to and frequency of same gender peers. We hypothesized that prosocial students would elicit expected praise and happiness from teachers, which would mediate the association between target student behavior and attributions. It was hypothesized that prosocial students would be perceived as more effortful and less able compared to nondescript students. The prosocial student was expected to be seen as less intelligent and masculine but more feminine than the nondescript student. This research extends existing literature by examining how peers interpret social classroom behavior and teacher reactions and whether behaving in ways that are stereotyped as feminine may have negative attributional consequences. Study 2 adds to the literature by exploring whether praise for social classroom behavior can induce attributional backlash, complementing work on the paradox of praise and negative classroom behavior's positive effect on attributions (Kessels & Heyder, 2020).

Whereas the previous studies focused on gender stereotyped behaviors from the perspective of educators and peers, study 3 focuses on gender role self-concepts and students themselves. Specifically, we turn to gender role self-concepts and the relation to factors associated with academic well-being. Evidence on gender role self-concepts showed that children and adolescents benefited from some cross-gender orientation, specifically androgyny (Korlat et al., 2022; Pauletti et al., 2017). Of course, it has been stated that the root of the positive outcomes of androgyny is not femininity but rather the presence of high masculinity (W. H. Jones et al., 1978; Korlat et al., 2022; Whitley, 1984). Self-ascribed masculinity, and by extension agency and instrumentality, was associated with many positive outcomes outside of the school context. Higher masculinity, for instance, has shown a relation with greater self-esteem (Cate & Sugawara, 1986; H. Choi et al., 2010; Holland & Andre, 1994; Whitley, 1988), and increased instrumentality was associated with lower anxiety and higher self-esteem (Sharpe et al., 1995). Interestingly, the associations between masculinity,

instrumentality, and agency to positive outcomes were often stronger in women than in men (Hirokawa & Dohi, 2007; Whitley, 1988). It seems that describing oneself with traits deemed more appropriate for the opposite gender can be more predictive than responding in line with gender stereotypes.

Given the many positive outcomes of a more instrumental masculinity, it is worth investigating this gender role self-concept in school more closely, but in the literature on school the focus has largely been on negative masculinity (Lyng, 2009). While many studies have examined gender differences in constructs relating to self-esteem (Schöne & Stiensmeier-Pelster, 2016) and motivation (see Butler, 2014), a focus on gender role self-concepts is not as common. To address this, the third manuscript investigated constructs of self-esteem, academic contingent self-esteem, and achievement motivation, encompassing hope for success and fear of failure, with gender role self-concept in $N = 355$ ninth grade students. In this cross-sectional study, we hypothesized that high instrumentality would relate to greater global self-esteem, lower academic contingent self-esteem, higher hope for success, and lower fear of failure. We expected results to vary by gender, with stronger effects to be shown by girls compared to boys. Relations between self-esteem constructs and achievement motivation were also investigated. Study 3 thus expands the literature on gender role self-concepts in school by examining how specifically instrumental masculinity can be an asset for students, particularly for adolescent girls.

Overall, the studies will show a range of dependent variables, samples, and methods, but they will all focus on gender and gendered characteristics, these being either gender stereotypes or gender role self-concepts. With this collection of studies, the present work argues that regarding the impact of gender and gendered characteristics is valuable in the academic context, as gender schemas in educators, peers and students themselves can influence judgments of likability, competence, effort, ability, self-esteem constructs, and achievement motive.

References I.

Abdi Zarrin, S., Gracia, E., & Paixão, M. P. (2020). Prediction of academic procrastination by fear of failure and self-regulation. *Educational Sciences: Theory and Practice*, *20*(3), 34–43. https://doi.org/10.12738/jestp.2020.3.003

Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology*, *85*(4), 768–776. https://doi.org/10.1037/0022-3514.85.4.768

Abele, A. E., & Candova, A. (2007). Prädiktoren des Belastungserlebens im Lehrerberuf: Befunde einer 4-jährigen Längsschnittstudie [Predicting teachers' stress experience: Findings from a 4-year longitudinal study]. *Zeitschrift Für Pädagogische Psychologie*, *21*(2), 107–118. https://doi.org/10.1024/1010-0652.21.2.107

Abele, A. E., Uchronski, M., Suitner, C., & Wojciszke, B. (2008). Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. *European Journal of Social Psychology*, *38*(7), 1202–1217. https://doi.org/10.1002/ejsp.575

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Abele, A. E., & Wojciszke, B. (2013). The big two in social judgment and behavior. *Social Psychology*, *44*(2), 61–62. https://doi.org/10.1027/1864-9335/a000137

American Psychological Association. (2018). Competence. In *American Psychological Association Dictionary*. https://dictionary.apa.org/competence

American Psychological Association. (2020a). Gender stereotype. In *American Psychological Association Dictionary*. https://dictionary.apa.org/gender-stereotype

American Psychological Association. (2020b). *Publication manual of the American Psychological Association: The official guide to APA style* (Seventh ed.). American Psychological Association

Anaya, L. M., & Zamarro, G. (2023). The role of student effort on performance in PISA: revisiting the gender gap in achievement. *Oxford Economic Papers*, *76*(2), 533–560. https://doi.org/10.1093/oep/gpad018

Antill, J. K., & Cunningham, J. D. (1979). Self-esteem as a function of masculinity in both sexes. *Journal of Consulting and Clinical Psychology*, *47*(4), 783–785. https://doi.org/10.1037/0022-006X.47.4.783

Asbrock, F. (2010). Stereotypes of social groups in Germany in terms of warmth and competence. *Social Psychology*, *41*(2), 76–81. https://doi.org/10.1027/1864-9335/a000011

Assouline, S. G., Mahatmya, D., Ihrig, L., & Lane, E. (2021). High-achieving rural middle-school students' academic self-efficacy and attributions in relationship to gender. *High Ability Studies*, *32*(2), 143–169. https://doi.org/10.1080/13598139.2020.1740582

Athenstaedt, U., Mikula, G., & Bredt, C. (2009). Gender role self-concept and leisure activities of adolescents. *Sex Roles*, *60*(5–6), 399–409. https://doi.org/10.1007/s11199-008-9543-y

Bachman, J. G., O'Malley, P. M., Freedman-Doan, P., Trzesniewski, K. H., & Donnellan, M. B. (2011). Adolescent self-esteem: Differences by race/ethnicity, gender, and age. *Self and Identity*, *10*(4), 445–473. https://doi.org/10.1080/15298861003794538

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion.* Rand McNally.

Bandura, A. (1977). *Social learning theory*. Prentice-Hall.

Barker, G. P., & Graham, S. (1987). Developmental study of praise and blame as attributional cues. *Journal of Educational Psychology*, *79*(1), 62–66. https://doi.org/10.1037/0022-0663.79.1.62

Barrett, A. E., & White, H. R. (2002). Trajectories of gender role orientations in adolescence and early adulthood: A prospective study of the mental health effects of masculinity and femininity. *Journal of Health and Social Behavior, 43*(4), 451-468.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155–162.

Bem, S. L. (1977). On the utility of alternative procedures for assessing psychological androgyny. *Journal of Consulting and Clinical Psychology, 45*(2), 196–205. https://doi.org/10.1037/0022-006X.45.2.196

Bem, S. L. (1981a). *Bem Sex-Role Inventory: A professional manual*. Consulting Psychologist Press.

Bem, S. L. (1981b). Gender schema theory: a cognitive account of sex typing. *Psychological Review*, *88*(4), 354–364.

Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs: Journal of Women in Culture and Society*, *8*(4), 598–616. https://doi.org/10.1086/494131

Bem, S. L. (1993). *The lenses of gender: Transforming the debate on sexual inequality*. Yale University Press.

Berger, A., & Krahé, B. (2013). Negative attributes are gendered too: Conceptualizing and measuring positive and negative facets of sex-role identity. *European Journal of Social Psychology*, *43*(6), 516–531. https://doi.org/10.1002/ejsp.1970

Berger, J., Cohen, B. P., & Zelditch, M. (1972). status characteristics and social interaction. *American Sociological Review*, *37*(3), 241–255. https://www-jstor-org.ezp2.lib.umn.edu/stable/2093465

Bergold, S., Kasper, D., Wendt, H., & Steinmayr, R. (2020). Being bullied at school: the case of high-achieving boys. *Social Psychology of Education*, *23*(2), 315–338. https://doi.org/10.1007/s11218-019-09539-w

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles*, *48*(9–10), 411–419. https://doi.org/10.1023/A:1023574427720

Bleidorn, W., Arslan, R. C., Denissen, J. J. A., Rentfrow, P. J., Gebauer, J. E., Potter, J., & Gosling, S. D. (2016). Age and gender differences in self-esteem: A cross-cultural window. *Journal of Personality and Social Psychology*, *111*(3), 396–410. http://dx.doi.org/10.1037/pspp0000078.supp

Bochicchio, V., Perillo, P., Valenti, A., Chello, F., Amodeo, A. L., Valerio, P., & Scandurra, C. (2019). Pre-service teachers' approaches to gender-nonconforming children in preschool and primary school: Clinical and educational implications. *Journal of Gay and Lesbian Mental Health*, *23*(2), 117–144. https://doi.org/10.1080/19359705.2019.1565791

Bornholt, L., & Möller, J. (2003). Attributions about achievement and intentions about further study in social context. *Social Psychology of Education*, *6*(3), 217–231. https://doi.org/10.1023/A:1024715609124

Bosson, J. K., Wilkerson, M., Kosakowska-Berezecka, N., Jurek, P., & Olech, M. (2022). Harder won and easier lost? testing the double standard in gender rules in 62 countries. *Sex Roles*, *87*(1–2), 1–19. https://doi.org/10.1007/s11199-022-01297-y

Bouchard, C., Coutu, S., Bigras, N., Lemay, L., Cantin, G., Bouchard, M. C., & Duval, S. (2015). Perceived, expressed and observed prosociality among four-year-old girls and boys in childcare centres. *Early Child Development and Care*, *185*(1), 44–65. https://doi.org/10.1080/03004430.2014.903940

Braun, S. S., & Davidson, A. J. (2017). Gender (non)conformity in middle childhood: A mixed methods approach to understanding gender-typed behavior, friendship, and peer preference. *Sex Roles*, *77*(1–2), 16–29. https://doi.org/10.1007/s11199-016-0693-z

Brophy, J. E., & Good, T. L. (1973). Feminization of American elementary schools. *The Phi Delta Kappan*, *54*(8), 564–566.

Brummelman, E., Crocker, J., & Bushman, B. J. (2016). The praise paradox: when and why praise backfires in children with low self-esteem. *Child Development Perspectives*, *10*(2), 111–115. https://doi.org/10.1111/cdep.12171

Budde, J. (2006). Jungen als Verlierer? Anmerkungen zum Topos der "Feminisierung von Schule." [Boys as loser? Remarks to the topic of the 'feminism of school'.] *Die Deutsche Schule*, *98*(4), 488–500. https://doi.org/10.25656/01:27352

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*(4), 676–713. https://doi.org/10.1037/0033-295X.106.4.676

Butler, R. (1994). Teacher communications and student interpretations: effects of teacher responses to failing students on attributional inferences in two age groups. *British Journal of Educational Psychology*, *64*(2), 277–294. https://doi.org/10.1111/j.2044-8279.1994.tb01102.x

Butler, R. (2014). Motivation in educational contexts: Does gender matter? In L. S. Liben, & R. S. Bigler (Eds.), *Advances in child development and behavior: The role of gender in educational contexts and outcomes* (Vol. 47, pp. 1–41). https://doi.org/10.1016/bs.acdb.2014.05.001

Canevello, A. (2020). Gender schema theory. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1741–1743). Springer Nature. https://doi.org/10.4324/9781315664095

Carrington, B., & McPhee, A. (2008). Boys' "underachievement" and the feminization of teaching. *Journal of Education for Teaching*, *34*(2), 109–120. https://doi.org/10.1080/02607470801979558

Carver, P. R., Yunger, J. L., & Perry, D. G. (2003). Gender identity and adjustment in middle childhood. *Sex Roles*, *49*(3–4), 95–109. https://doi.org/10.1023/A:1024423012063

Cate, R., & Sugawara, A. I. (1986). Sex role orientation and dimensions of self-esteem among middle adolescents. *Sex Roles*, *15*(3–4), 145–158. https://doi.org/10.1007/BF00287480

Chen-Bouck, L., & Patterson, M. M. (2016). Contingent self-worth in Chinese adolescents and young adults: Relations with global self-esteem and depressive symptoms. *Social Development*, *25*(4), 846–865. https://doi.org/10.1111/sode.12173

Choi, H., Kim, J. H., Hwang, M. H., & Heppner, M. J. (2010). Self-esteem as a mediator between instrumentality, gender role conflict and depression in male Korean high school students. *Sex Roles*, *63*(5), 361–372. https://doi.org/10.1007/s11199-010-9801-7

Choi, N. (2004). Sex role group differences in specific, academic, and general self-efficacy. *Journal of Psychology: Interdisciplinary and Applied*, *138*(2), 149–159. https://doi.org/10.3200/JRLP.138.2.149-159

Ciarrochi, J., & Heaven, P. C. L. (2009). A longitudinal study into the link between adolescent personality and peer-rated likeability and adjustment: Evidence of gender differences. *Journal of Research in Personality*, *43*(6), 978–986. https://doi.org/10.1016/j.jrp.2009.08.006

Corby, B. C., Hodges, E. V. E., & Perry, D. G. (2007). Gender identity and adjustment in Black, Hispanic, and White preadolescents. *Developmental Psychology*, *43*(1), 261–266. https://doi.org/10.1037/0012-1649.43.1.261

Covington, M. V., & Omelich, C. L. (1979). Effort: The double-edged sword in school achievement. *Journal of Educational Psychology*, *71*(2), 169–182. https://doi.org/10.1037/0022-0663.71.2.169

Coyle, E. F., Fulcher, M., & Trübutschek, D. (2016). Sissies, mama's boys, and tomboys: Is children's gender nonconformity more acceptable when nonconforming traits are positive? *Archives of Sexual Behavior*, *45*, 1827–1838. https://doi.org/10.1007/s10508-016-0695-5

Crocker, J., Karpinski, A., Quinn, D. M., & Chase, S. K. (2003). When grades determine self-worth: Consequences of contingent self-worth for male and female engineering and psychology majors. *Journal of Personality and Social Psychology*, *85*(3), 507–516. https://doi.org/10.1037/0022-3514.85.3.507

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*(3), 593–623. https://doi.org/10.1037/0033-295X.108.3.593

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, *40*, 61–149.

Del Giudice, M. (2017). Pink, blue, and gender: An update. *Archives of Sexual Behavior*, *46*(6), 1555–1563. https://doi.org/10.1007/s10508-017-1024-3

Denham, S. A., & Holt, R. W. (1993). Preschoolers' likability as cause or consequence of their social behavior. *Developmental Psychology*, *29*(2), 271–275. https://doi.org/10.1037/0012-1649.29.2.271

Dupree, C. H., & Fiske, S. T. (2017). Universal dimensions of social signals: Warmth and competence. In J. K. Burgoon, N. Magnenat-Thalmann, M. Pantic, & A. Vinciarelli (Eds.), *Social signal processing* (pp. 23–33). Cambridge University Press. https://doi.org/10.1017/9781316676202.003

Eagly, A. H. (1987). *Sex differences in social behavior: a social-role interpretation*. Erlbaum.

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Eagly, A. H., & Wood, W. (2012). Social role theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 458–476). SAGE Publications Ltd. https://doi.org/10.4135/9781446249222.n49

Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles*, *47*(3–4), 99–114. https://doi.org/10.1023/A:1021020920715

Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen [Gender stereotypes: Roles, identities and prejudice]. In R. Becker & B. Kortendiek (Eds.), *Handbuch Frauen- und Geschlechterforschung: Theorie, Methoden, Empirie* (pp. 178–189). Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91972-0_20

Egan, S. K., & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*, *37*(4), 451–463. https://doi.org/10.1037/0012-1649.37.4.451

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial Development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 646–718). John Wiley & Sons, Inc. https://doi.org/10.4324/9781315560984-2

Ellemers, N. (2018). Gender stereotypes. *Annual Reviews*, *69*, 275–298. https://doi.org/10.4135/9781412956246.n190

Engels, M. C., Colpin, H., Van Leeuwen, K., Bijttebier, P., Van Den Noortgate, W., Claes, S., Goossens, L., & Verschueren, K. (2016). Behavioral engagement, peer status, and teacher–student relationships in adolescence: A longitudinal study on reciprocal influences. *Journal of Youth and Adolescence*, *45*(6), 1192–1207. https://doi.org/10.1007/s10964-016-0414-5

Engeser, S. (2005). *Messung des expliziten Leistungsmotivs: Kurzform der Achievement Motives Scale [Measuring the explicit achievement motive: A short version of the Achievement Motives Scale]*.

Espinoza, P., Arêas da Luz Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: Teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, *17*(1), 105–126. https://doi.org/10.1007/s11218-013-9226-6

Exner-Cortens, D., Wright, A., Claussen, C., & Truscott, E. (2021). A systematic review of adolescent masculinities and associations with internalizing behavior problems and social support. *American Journal of Community Psychology, 68*(1–2), 215–231. https://doi.org/10.1002/ajcp.12492

Fairlamb, S. (2020). We need to talk about self-esteem: The effect of contingent self-worth on student achievement and well-being. *Scholarship of Teaching and Learning in Psychology*. *8*(1), 45–57. https://doi.org/10.1037/stl0000205

Feinman, S. (1981). Why is cross-sex-role behavior more approved for girls than for boys? A status characteristic approach. *Sex Roles*, *7*(3), 289–300. https://doi.org/10.1007/BF00287543

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69. https://doi.org/10.1007/BF00311015

Ferkany, M. (2008). The educational importance of self-esteem. *Journal of Philosophy of Education*, *42*(1), 119–132. https://doi.org/10.1111/j.1467-9752.2008.00610.x

Fernández, J., & Coello, M. T. (2010). Do the BSRI and PAQ really measure masculinity and femininity? *Spanish Journal of Psychology*, *13*(2), 1000–1009. https://doi.org/10.1017/S113874160000264X

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.4324/9781315187280

Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis)respecting versus (dis)liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, *55*(3), 473–489. https://doi.org/10.1111/0022-4537.00128

Folmer, A. S., Cole, D. A., Sigal, A. B., Benbow, L. D., Satterwhite, L. F., Swygert, K. E., & Ciesla, J. A. (2008). Age-related changes in children's understanding of effort and ability: Implications for attribution theory and motivation. *Journal of Experimental Child Psychology*, *99*(2), 114–134. https://doi.org/10.1016/j.jecp.2007.09.003

Fredricks, J. A., & Eccles, J. S. (2002). Children's competence and value beliefs from childhood through adolescence: growth trajectories in two male-sex-typed domains. *Developmental Psychology*, *38*(4), 519–533. https://doi.org/10.1037/0012-1649.38.4.519

Freudenthaler, H. H., Spinath, B., & Neubauer, A. C. (2008). Predicting school achievement in boys and girls. *European Journal of Personality*, *22*, 231–245. https://doi.org/10.1002/per

Gerber, G. L. (2009). Status and the gender stereotyped personality traits: Toward an integration. *Sex Roles*, *61*(5–6), 297–316. https://doi.org/10.1007/s11199-008-9529-9

Glock, S., & Kleen, H. (2017). Gender and student misbehavior: Evidence from implicit and explicit measures. *Teaching and Teacher Education*, *67*, 93–103. https://doi.org/10.1016/j.tate.2017.05.015

Graham, S., & Williams, C. (2009). An attributional approach to motivation in school. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 11–35). Routledge.

Granleese, J., Turner, I., & Trew, K. (1989). Teachers' and boys' and girls' perceptions of competence in the primary school: The importance of physical competence. *British Journal of Educational Psychology*, *59*(1), 31–37. https://doi.org/10.1111/j.2044-8279.1989.tb03073.x

Greenwald, A. G., Rudman, L. A., Nosek, B. A., Banaji, M. R., Farnham, S. D., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*(1), 3–25. https://doi.org/10.1037/0033-295X.109.1.3

Grow, A., Takács, K., & Pál, J. (2016). Status characteristics and ability attributions in Hungarian school classes: An exponential random graph approach. *Social Psychology Quarterly*, *79*(2), 156–167. https://doi.org/10.1177/0190272516643052

Hadjar, A., Krolak-Schwerdt, S., Priem, K., & Glock, S. (2014). Gender and educational achievement. *Educational Research*, *56*(2), 117–125. https://doi.org/10.1080/00131881.2014.898908

Hamilton, O. S., & Lordan, G. (2023). Ability or luck: A systematic review of interpersonal attributions of success. *Frontiers in Psychology*, *13*, 1035012. https://doi.org/10.3389/fpsyg.2022.1035012

Hannover, B., & Kessels, U. (2011). Sind Jungen die neuen Bildungsverlierer? Empirische Evidenz für Geschlechterdisparitäten zuungunsten von Jungen und Erklärungsansätze [Are boys left behind at school? Reviewing and explaining education-related gender disparities]. *Zeitschrift Für Pädagogische Psychologie*, *25*(2), 89–103. https://doi.org/10.1024/1010-0652/a000039

Hareli, S., & Weiner, B. (2002). Social emotions and personality inferences: A scaffold for a new direction in the study of achievement motivation. *Educational Psychologist*, *37*(3), 183–193. https://doi.org/10.1207/S15326985EP3703_4

Harvey, P., Madison, K., Martinko, M., Crook, T. R., & Crook, T. A. (2014). Attribution theory in the organizational sciences: The road traveled and the path ahead. *Academy of Management Perspectives*, *28*(2), 128–146. https://doi.org/10.5465/amp.2012.0175

Heider, F. (1958). *The psychology of interpersonal relations*. John Wiley & Sons Inc.

Helbig, M. (2010). Sind Lehrerinnen für den geringeren Schulerfolg von Jungen verantwortlich? [Are female teachers responsible for the school performance gap between boys and girls?]. *Kölner Zeitschrift Für Soziologie Und Sozialpsychologie*, *62*(1), 93–111. https://doi.org/10.1007/s11577-010-0095-0

Helbig, M. (2013). Geschlecht und Bildungserfolg: Historische Perspektiven auf die "Krise der Jungen" [Gender and educational achievement: Historical perspectives on the „boy crisis"]. *WZBrief Bildung, 23*(Januar), 2-6. https://nbn-resolving.de/urn:nbn:de:101:1-201302125341

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The multiple dimensions of gender stereotypes: A current look at men's and women's characterizations of others and themselves. *Frontiers in Psychology*, *10*, 376558. https://doi.org/10.3389/fpsyg.2019.00011

Hernandez Bark, A. S., Seliverstova, K., & Ohly, S. (2022). Getting credit for proactivity? The effects of gender. *Journal of Applied Social Psychology*, *52*(8), 660–675. https://doi.org/10.1111/jasp.12833

Heyder, A., & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles*, *69*(11–12), 605–617. https://doi.org/10.1007/s11199-013-0309-9

Heyder, A., & Kessels, U. (2015). Do teachers equate male and masculine with lower academic engagement? How students' gender enactment triggers gender stereotypes at school. *Social Psychology of Education*, *18*(3), 467–485. https://doi.org/10.1007/s11218-015-9303-0

Heyder, A., & Kessels, U. (2017). Boys don't work? On the psychological benefits of showing low effort in high school. *Sex Roles*, *77*(1–2), 72–85. https://doi.org/10.1007/s11199-016-0683-1

Heyder, A., van Hek, M., & van Houtte, M. (2021). When gender stereotypes get male adolescents into trouble: A longitudinal study on gender conformity pressure as a predictor of school misconduct. *Sex Roles*, *84*, 61–75. https://doi.org/https://doi.org/10.1007/s11199-020-01147-9

Hill, M. E., & Augoustinos, M. (1997). Re-examining gender bias in achievement attributions. *Australian Journal of Psychology*, *49*(2), 85–90. https://doi.org/10.1080/00049539708259858

Hilton, J. L., & Hippel, W. Von. (1996). Stereotypes. *Annual Review of Psychology*, *47*, 237–271.

Hirokawa, K., & Dohi, I. (2007). Agency and communion related to mental health in Japanese young adults. *Sex Roles*, *56*(7–8), 517–524. https://doi.org/10.1007/s11199-007-9190-8

Hoffman, A. J., Dumas, F., Loose, F., Smeding, A., Kurtz-Costes, B., & Régner, I. (2019). Development of gender typicality and felt pressure in European French and North African French adolescents. *Child Development*, *90*(3), e306–e321. https://doi.org/10.1111/cdev.12959

Holland, A., & Andre, T. (1994). The relationship of self-esteem to selected personal and environmental resources of adolescents. *Adolescence*, *29*(114), 345–360.

Huyge, E., Van Maele, D., & Van Houtte, M. (2015). Does students' machismo fit in school? Clarifying the implications of traditional gender role ideology for school belonging. *Gender and Education*, *27*(1), 1–18. https://doi.org/10.1080/09540253.2014.972921

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, *74*(2), 171–193. https://doi.org/10.1037/amp0000307

Jackson, C. (2003). Motives for "Laddishness" at school: Fear of failure and fear of the "feminine." *British Educational Research Journal*, *29*(4), 583–598. https://doi.org/10.1080/01411920301847

Jackson, C., & Dempster, S. (2009). 'I sat back on my computer … with a bottle of whisky next to me': Constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *Journal of Gender Studies*, *18*, 341–356. https://doi.org/doi.org/10.1080/09589 23090 32600 19.

Jambunathan, S., & Hurlbut, N. L. (2000). Gender comparisons in the perception of self-competence among four-year-old children. *Journal of Genetic Psychology*, *161*(4), 469–477. https://doi.org/10.1080/00221320009596725

Jewell, J. A., & Brown, C. S. (2014). Relations among gender typicality, peer relations, and mental health during early adolescence. *Social Development*, *23*(1), 137–156. https://doi.org/10.1111/sode.12042

Jones, S., & Myhill, D. (2004). 'Troublesome boys' and 'compliant girls': Gender identity and perceptions of achievement and underachievement. *British Journal of Sociology of Education*, *25*(5), 547–561.

Jones, W. H., Chernovetz, M. E., & Hansson, R. O. (1978). The enigma of androgyny: Differential implications for males and females? *Journal of Consulting and Clinical Psychology*, *46*(2), 298–313. https://doi.org/10.1037/0022-006X.46.2.298

Kervyn, N., Bergsieker, H. B., & Fiske, S. T. (2012). The innuendo effect: Hearing the positive but inferring the negative. *Journal of Experimental Social Psychology*, *48*(1), 77–85. https://doi.org/10.1016/j.jesp.2011.08.001

Kervyn, N., Bergsieker, H. B., Grignard, F., & Yzerbyt, V. Y. (2016). An advantage of appearing mean or lazy: Amplified impressions of competence or warmth after mixed descriptions. *Journal of Experimental Social Psychology*, *62*, 17–23. https://doi.org/10.1016/j.jesp.2015.09.004

Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, *20*(3), 309–323. https://doi.org/10.1007/BF03173559

Kessels, U., & Hannover, B. (2008). When being a girl matters less: Accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students'

physics-related self-concept of ability. *British Journal of Educational Psychology*, *78*(2), 273–289. https://doi.org/10.1348/000709907X215938

Kessels, U., & Heyder, A. (2020). Not stupid, but lazy: Psychological benefits of disruptive classroom behavior from an attributional perspective. *Social Psychology of Education*, *23*, 583–613. https://doi.org/10.1007/s11218-020-09550-6

Kessels, U., Heyder, A., Latsch, M., & Hannover, B. (2014). How gender differences in academic engagement relate to students' gender identity. *Educational Research*, *56*(2), 220–229. https://doi.org/10.1080/00131881.2014.898916

Kessels, U., & Steinmayr, R. (2013). Macho-man in school: Toward the role of gender role self-concepts and help seeking in school performance. *Learning and Individual Differences*, *23*(1), 234–240. https://doi.org/10.1016/j.lindif.2012.09.013

Kisfalusi, D., Janky, B., & Takács, K. (2019). Double standards or social identity? The role of gender and ethnicity in ability perceptions in the classroom. *Journal of Early Adolescence*, *39*(5), 745–780. https://doi.org/10.1177/0272431618791278

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin*, *125*(4), 470–500.

Koenig, A. M. (2018). Comparing prescriptive and descriptive gender stereotypes about children, adults, and the elderly. *Frontiers in Psychology*, *9*, 381468. https://doi.org/10.3389/fpsyg.2018.01086

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*(4), 616–642. https://doi.org/10.1037/a0023557

Kohlberg, L. (1966). A cognitive-developmental analysis of children's sex-role concepts and attitudes. In E. E. Maccoby (Ed.), *The development of sex differences* (pp. 82–173). Stanford University Press.

Koivula, N. (1995). Ratings of gender appropriateness of sports participation: Effects of gender-based schematic processing. *Sex Roles*, *33*(7/8), 543–557.

Korlat, S., Holzer, J., Schultes, M.-T., Buerger, S., Schober, B., Spiel, C., & Kollmayer, M. (2022). Benefits of psychological androgyny in adolescence: The role of gender role self-concept in school-related well-being. *Frontiers in Psychology*, *13*, 856758. https://doi.org/10.3389/fpsyg.2022.856758

Kultusministerkonferenz. (2016). *Leitlinien zur Sicherung der Chancengleichheit durch geschlechtersensible schulische Bildung und Erziehung [Guidelines for ensuring equal opportunities through gender-sensitive school education and upbringing]*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2016/2016_10_06-Geschlechtersensible-schulische_Bildung.pdf

Kwan, K. M. W., Shi, S. Y., Nabbijohn, A. N., MacMullin, L. N., VanderLaan, D. P., & Wong, W. I. (2020). Children's appraisals of gender nonconformity: Developmental pattern and intervention. *Child Development*, *91*(4), e780–e798. https://doi.org/10.1111/cdev.13316

Lang, J. W. B., & Fries, S. (2006). A revised 10-item version of the achievement motives scale: Psychometric properties in German-speaking samples. *European Journal of Psychological Assessment*, *22*(3), 216–224. https://doi.org/10.1027/1015-5759.22.3.216

Lavy, S., & Naama-Ghanayim, E. (2020). Why care about caring? Linking teachers' caring and sense of meaning at work with students' self-esteem, well-being, and school engagement. *Teaching and Teacher Education*, *91*, 103046. https://doi.org/10.1016/j.tate.2020.103046

Leung, E. (2020). Gender roles. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1738–1740). Springer Nature.

Li, Y., & Wei, X. (2022). Examining the effects of male candidates' gender nonconformity on employment decisions. *Asia Pacific Journal of Human Resources*, *62*(1), e12363. https://doi.org/10.1111/1744-7941.12363

Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*, *67*(2), i–183.

Liben, L. S., & Bigler, R. S. (2017). Understanding and undermining the development of gender dichotomies: The legacy of Sandra Lipsitz Bem. *Sex Roles*, *76*(9–10), 544–555. https://doi.org/10.1007/s11199-015-0519-4

Liben, L. S., & Signorella, M. L. (1980). Gender-related schemata and constructive memory in children. *Child Development*, *51*(1), 11–18. https://www.jstor.org/stable/1129584?seq=1&cid=pdf-

Lohbeck, A., Grube, D., & Moschner, B. (2017). Academic self-concept and causal attributions for success and failure amongst elementary school children. *International Journal of Early Years Education*, *25*(2), 190–203. https://doi.org/10.1080/09669760.2017.1301806

Lu, T., Li, L., Niu, L., Jin, S., & French, D. C. (2018). Relations between popularity and prosocial behavior in middle school and high school Chinese adolescents. *International Journal of Behavioral Development*, *42*(2), 175–181. https://doi.org/10.1177/0165025416687411

Lyng, S. T. (2009). Is there more to "Antischoolishness" than masculinity? On multiple student styles, gender, and educational self-exclusion in secondary school. *Men and Masculinities*, *11*(4), 462–487. https://doi.org/10.1177/1097184X06298780

Mandal, E., & Moroń, M. (2019). Contingencies of self-worth and global self-esteem among college women: The role of masculine and feminine traits endorsement. *Social Psychological Bulletin*, *14*(3), 1-30. https://doi.org/10.32872/spb.v14i1.33507

Marsh, H. W., Antill, J. K., & Cunningham, J. D. (1987). Masculinity, femininity, and androgyny: Relations to self-esteem and social desirability. *Journal of Personality*, *55*(4), 661–685. https://doi.org/10.1111/j.1467-6494.1987.tb00457.x

Martin, A. E., & Slepian, M. L. (2021). The primacy of gender: gendered cognition underlies the big two dimensions of social cognition. *Perspectives on Psychological Science*, *16*(6), 1143–1158. https://doi.org/10.1177/1745691620904961

Martin, A. J., & Marsh, H. W. (2003). Fear of failure: Friend or foe? *Australian Psychologist*, *38*(1), 31–38. https://doi.org/10.1080/00050060310001706997

Martin, C. L. (2002). Cognitive theories of gender development. In T. Eckes & H. Trautner (Eds.), *The developmental social psychology of gender*. Routledge.

Martin, C. L., Cook, R. E., & Andrews, N. C. Z. (2017). Reviving androgyny: A modern day perspective on flexibility of gender identity and behavior. *Sex Roles*, *76*(9–10), 592–603. https://doi.org/10.1007/s11199-016-0602-5

Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development*, *66*(5), 1453–1471. https://doi.org/10.1111/j.1467-8624.1995.tb00945.x

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, *52*(4), 1119–1134. https://www.jstor.org/stable/1129498

Matud, M. P., López-Curbelo, M., & Fortes, D. (2019). Gender and psychological well-being. *International Journal of Environmental Research and Public Health*, *16*(19), 1–11. https://doi.org/10.3390/ijerph16193531

McClelland, D. C., Atkinson, J., Clark, R., & Lowell, E. (1953). *The achievement motive*. NY: Appleton-Century-Croft.

Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. *Journal of School Psychology*, *44*(5), 351–373. https://doi.org/10.1016/j.jsp.2006.04.004

Meimoun, E., Bonnot, V., Berenguer, J., & Aelenei, C. (2023). How are gender counter-stereotypical adolescents evaluated by their peers? Investigating the backlash effect in French schools. *Social Psychology of Education*, 1–25. https://doi.org/10.1007/s11218-023-09807-w

Menon, M., Schellhorn, K., & Lowe, C. A. (2013). Interactive influences of gender identity and gender typing on early adolescents' well-being. *Psychological Studies*, *58*(1), 1–9. https://doi.org/10.1007/s12646-012-0159-5

Mok, M. M. C., Kennedy, K. J., & Moore, P. J. (2011). Academic attribution of secondary students: Gender, year level and achievement level. *Educational Psychology*, *31*(1), 87–104. https://doi.org/10.1080/01443410.2010.518596

Möller, J. (2005). Paradoxical effects of praise and criticism: Social, dimensional and temporal comparisons. *British Journal of Educational Psychology*, *75*(2), 275–295. https://doi.org/10.1348/000709904X24744

Moore, J. S. B., & Smith, M. (2018). Children's levels of contingent self-esteem and social and emotional outcomes. *Educational Psychology in Practice*, *34*(2), 113–130. https://doi.org/10.1080/02667363.2017.1411786

Moriizumi, S. (2018). Stereotypes. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 5217–5219). Springer Nature.

Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules: Status incongruity backlash against modest men. *Psychology of Men & Masculinity*, *11*(2), 140–151. https://doi.org/10.1037/a0018093

Muntoni, F., & Retelsdorf, J. (2019). At their children's expense: How parents' gender stereotypes affect their children's reading outcomes. *Learning and Instruction*, *60*, 95–103. https://doi.org/10.1016/j.learninstruc.2018.12.002

Myhill, D. (2002). Bad boys and good girls? Patterns of interaction and response in whole class teaching. *British Educational Research Journal*, *28*(3), 339–352.

Nauts, S. (2015). *Backlash for gender atypicality*. Radboud University Nijmegen.

Nielson, M. G., Schroeder, K. M., Martin, C. L., & Cook, R. E. (2020). Investigating the relation between gender typicality and pressure to conform to gender norms. *Sex Roles*, *83*(9–10), 523–535. https://doi.org/10.1007/s11199-020-01136-y

O'Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, *9*(1), 3777. https://doi.org/10.1038/s41467-018-06292-0

OECD. (2023). *PISA 2022 Assessment and Analytical Framework*. https://oecdch.art/a40de1dbaf/C074.

Orlofsky, J. L., & Stake, J. E. (1981). Psychological masculinity and femininity: Relationship to striving and self-concept in the achievement and interpersonal domains. *Psychology of Women Quarterly*, *6*(2), 218–233. https://doi.org/10.1111/j.1471-6402.1981.tb00409.x

Orr, A. J. (2011). Gendered capital: Childhood socialization and the "boy crisis" in education. *Sex Roles*, *65*(3–4), 271–284.

Orr, E., & Ben-Eliahu, E. (1993). Gender differences in idiosyncratic sex-typed self-images and self-esteem. *Sex Roles*, *29*(3–4), 271–296. https://doi.org/10.1007/BF00289939

Oswald, D. L., & Lindstedt, K. (2006). The content and function of gender self-stereotypes: An exploratory investigation. *Sex Roles*, *54*(7–8), 447–458. https://doi.org/10.1007/s11199-006-9026-y

Park, G. R., & Kim, J. (2023). Parental involvement and academic achievement of adolescents: The mediating roles of school adjustment. *School Psychology International*. Online First https://doi.org/10.1177/01430343231207134

Pauletti, R. E., Menon, M., Cooper, P. J., Aults, C. D., & Perry, D. G. (2017). Psychological androgyny and children's mental health: A new look with new measures. *Sex Roles*, *76*(11–12), 705–718. https://doi.org/10.1007/s11199-016-0627-9

Paulhus, D. L., & Trapnell, P. D. (2008). Self-presentation of personality: An agency-communion framework. In O. P. John, R.W. Robins, & L. A. Pervin (Eds.), *Handbook of personality psychology* (pp. 492–517). Guilford Press.

Pellegrini, A. D. (2011). “In the eye of the beholder”: Sex bias in observations and ratings of children’s aggression. *Educational Researcher*, *40*(6), 281–286. https://doi.org/10.3102/0013189X11421983

Perry, D. G., Pauletti, R. E., & Cooper, P. J. (2019). Gender identity in childhood: A review of the literature. *International Journal of Behavioral Development*, *43*(4), 289–304. https://doi.org/10.1177/0165025418811129

Piaget, J. (1952). *The origins of intelligence in children*. International Universities Press.

Piché, C., & Plante, C. (1991). Perceived masculinity, femininity and androgyny among primary school boys. *European Journal of Psychology of Education*, *6*(4), 423–435. https://doi.org/10.1007/BF03172775

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn’t be, are allowed to be, and don’t have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, *26*(4), 269–281. https://doi.org/10.1111/1471-6402.t01-1-00066

Pulles, N. J., & Hartman, P. (2017). Likeability and its effect on outcomes of interpersonal interaction. *Industrial Marketing Management*, *66*, 56–63. https://doi.org/10.1016/j.indmarman.2017.06.008

Quatman, T., Sokolik, E., & Smith, K. (2000). Adolescent perception of peer success: A gendered perspective over time. *Sex Roles*, *43*(1–2), 61–84.

Quenneville, S., Talwar, V., & Bosacki, S. (2022). Teacher ratings and adolescent students’ perceived social behaviours and gender-role orientations. *Journal of Gender Studies*, *31*(4), 444–456. https://doi.org/10.1080/09589236.2021.1988530

Rammstedt, B., & Rammsayer, T. H. (2002). Gender differences in self-estimated intelligence and their relation to gender-role orientation. *European Journal of Personality*, *16*(5), 369–382. https://doi.org/10.1002/per.454

Reyna, C. (2000). Lazy, dumb, or industrious: When stereotypes convey attribution information in the classroom. *Educational Psychology Review*, *12*(1), 85–110. https://doi.org/10.1023/A:1009037101170

Rogers, A., Updegraff, K., Santos, C., & Martin, C. L. (2017). Masculinity and school engagement in middle school. *Psychology of Men & Masculinity*, *18*(1), 50–61.

Rose, A. J., & Montemayor, R. (1994). The relationship between gender role orientation and perceived self-competency in male and female adolescents. *Sex Roles*, *31*(9–10), 579–595. https://doi.org/10.1007/BF01544281

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, *57*(4), 743–762. https://doi.org/10.1111/0022-4537.00239

Rudman, L. A., & Glick, P. (2010). *The social psychology of gender: How power and intimacy shape gender relations*. Guilford Press.

Rudman, L. A., Moss-Racusin, C. A., Glick, P., & Phelan, J. E. (2012a). Reactions to vanguards: Advances in backlash theory. In *Advances in experimental social psychology* (1st ed., Vol. 45, pp. 167–227). Elsevier Inc. https://doi.org/10.1016/B978-0-12-394286-9.00004-4

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012b). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Sáinz, M., Solé, J., Fàbregues, S., & García-Cuesta, S. (2021). Secondary school teachers' views of gender differences in school achievement and study choices in Spain. *SAGE Open*, *11*(3), 21582440211047573. https://doi.org/10.1177/21582440211047573

Sanborn-Overby, U. A., & Powlishta, K. K. (2020). Differential reactions to male and female gender-role violations: Testing the sexual orientation hypothesis. *Archives of Sexual Behavior*, *49*(7), 2531–2546. https://doi.org/10.1007/s10508-020-01803-x

Sandstrom, M. J., & Cillessen, A. H. N. (2006). Likeable versus popular: Distinct implications for adolescent adjustment. *International Journal of Behavioral Development*, *30*(4), 305–314. https://doi.org/10.1177/0165025406072789

Schöne, C., & Stiensmeier-Pelster, J. (2016). *Selbstwertinventar für Kinder und Jugendliche – SEKJ [Self-Esteem Inventory for Children and Adolescents]*. Hogrefe. https://doi.org/10.13140/RG.2.1.3583.5607.

Schöne, C., Tandler, S. S., & Stiensmeier-Pelster, J. (2015). Contingent self-esteem and vulnerability to depression: Academic contingent self-esteem predicts depressive symptoms in students. *Frontiers in Psychology*, *6*, 1573. https://doi.org/10.3389/fpsyg.2015.01573

Schoneveld, E., & Brummelman, E. (2023). "You did incredibly well!": teachers' inflated praise can make children from low-SES backgrounds seem less smart (but more hardworking). *Science of Learning*, *31*, 1–12. https://doi.org/10.1038/s41539-023-00183-w

Severiens, S., & Dam, G. T. (1998). A multilevel meta-analysis of gender differences in learning orientations. *British Journal of Educational Psychology*, *68*(4), 595–608. https://doi.org/10.1111/j.2044-8279.1998.tb01315.x

Sharpe, M. J., Heppner, P. P., & Dixon, W. A. (1995). Gender role conflict, instrumentality, expressiveness, and well-being in adult men. *Sex Roles*, *33*(1–2), 1–18. https://doi.org/10.1007/BF01547932

Signorella, M. L., & Jamison, W. (1986). Psychological bulletin masculinity, femininity, androgyny, and cognitive performance: A meta-analysis. *Psychological Bulletin*, *100*(2), 207–228.

Skinner, O. D., Kurtz-Costes, B., Wood, D., & Rowley, S. J. (2018). Gender typicality, felt pressure for gender conformity, racial centrality, and self-esteem in African American adolescents. *Journal of Black Psychology*, *44*(3), 195–218. https://doi.org/10.1177/0095798418764244

Smith, D. S., & Juvonen, J. (2017). Do I fit in? Psychosocial ramifications of low gender typicality in early adolescence. *Journal of Adolescence*, *60*, 161–170. https://doi.org/10.1016/j.adolescence.2017.07.014

Spence, J. T. (1991). Do the BSRI and PAQ measure the same or different concepts? *Psychology of Women Quarterly*, *15*, 141–165.

Spence, J. T., & Helmreich, R. L. (1980). Masculine instrumentality and feminine expressiveness: Their relationships with sex role attitudes and behaviors. *Psychology of Women Quarterly*, *5*(2), 147–163. https://doi.org/10.1111/j.1471-6402.1980.tb00951.x

Spence, J. T., Helmreich, R. L., & Stapp, J. (1974). *The personal attributes questionnaire: A measure of sex role stereotypes and masculinity-femininity*. University of Texas.

Spence, J. T., Helmreich, R., & Stapp, J. (1975). Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. In *Journal of Personality and Social Psychology, 32*(1), 29-39. https://doi.org/10.1037/h0076857

Starr, C. R., & Zurbriggen, E. L. (2017). Sandra Bem's gender schema theory after 34 years: A review of its reach and impact. *Sex Roles*, *76*(9–10), 566–578. https://doi.org/10.1007/s11199-016-0591-4

Steinmayr, R., & Kessels, U. (2017). Good at school = successful on the job? Explaining gender differences in scholastic and vocational success. *Personality and Individual Differences*, *105*, 107–115. https://doi.org/10.1016/j.paid.2016.09.032

Steinmayr, R., Meißner, A., Weidinger, A. F., & Wirthwein, L. (2014). Academic achievement. In L. H. Meyer (Ed.), *Oxford bibliographies online: Education.* Oxford University Press. https://doi.org/10.1093/OBO/9780199756810

Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, *22*, 185–209. https://doi.org/10.1002/per

Steinmayr, R., & Spinath, B. (2009a). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80–90. https://doi.org/10.1016/j.lindif.2008.05.004

Steinmayr, R., & Spinath, B. (2009b). What explains boys' stronger confidence in their intelligence? *Sex Roles*, *61*(9–10), 736–749. https://doi.org/10.1007/s11199-009-9675-8

Steinmayr, R., Weidinger, A. F., Schwinger, M., & Spinath, B. (2019). The importance of students' motivation for their academic achievement-replicating and extending previous findings. *Frontiers in Psychology*, *10*, 1730. https://doi.org/10.3389/fpsyg.2019.01730

Stoeber, J., & Rambow, A. (2007). Perfectionism in adolescent school students: Relations with motivation, achievement, and well-being. *Personality and Individual Differences*, *42*(7), 1379–1389. https://doi.org/10.1016/j.paid.2006.10.015

Stoet, G., & Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence*, *48*, 137–151. https://doi.org/10.1016/j.intell.2014.11.006

Strage, A. (1997). Agency, communion, and achievement motivation. *Adolescence*, *32*(126), 299–312.

Sullivan, J., Ciociolo, A., & Moss-Racusin, C. A. (2022). Establishing the content of gender stereotypes across development. *PLoS ONE*, *17*(7), e0263217. https://doi.org/10.1371/journal.pone.0263217

Sullivan, J., Moss-Racusin, C., Lopez, M., & Williams, K. (2018). Backlash against gender stereotype-violating preschool children. *PLoS ONE*, *13*(4), e0195503. https://doi.org/10.1371/journal.pone.0195503

Szcześniak, M., Bajkowska, I., Czaprowska, A., & Sileńska, A. (2022). Adolescents' self-esteem and life satisfaction: Communication with peers as a mediator. *International Journal of Environmental Research and Public Health*, *19*, 3777. https://doi.org/10.3390/ijerph19073777

Thomas, R. N., & Blakemore, J. E. O. (2013). Adults' attitudes about gender nonconformity in childhood. *Archives of Sexual Behavior*, *42*(3), 399–412. https://doi.org/10.1007/s10508-012-0023-7

Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics*, *41*, 191–207.

Tobin, D. D., Menon, M., Menon, M., Spatta, B. C., Hodges, E. V., & Perry, D. G. (2010). The intrapsychics of gender: a model of self-socialization. *Psychological Review*, *117*(2), 601–622.

Todd, B. K., Barry, J. A., & Thommessen, S. A. O. (2017). Preferences for 'gender-typed' toys in boys and girls aged 9 to 32 months. *Infant and Child Development*, *26(3)*, e1986. https://doi.org/10.1002/icd.1986

Tolman, D. L., Impett, E. A., Tracy, A. J., & Michael, A. (2006). Looking good, sounding good: Femininity ideology and adolescent girls' mental health. *Psychology of Women Quarterly*, *35*, 85–95.

Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Review*, *5*(4), 321–344. https://doi.org/10.1207/S15327957PSPR0504_3

Unger, R. K. (1979). Toward a redefinition of sex and gender. *American Psychologist*, *34*(11), 1085–1094. https://doi.org/10.1037//0003-066x.34.11.1085

van der Kaap-Deeder, J., Wouters, S., Verschueren, K., Briers, V., Deeren, B., & Vansteenkiste, M. (2016). The pursuit of self-esteem and its motivational implications. *Psychologica Belgica*, *56*(2), 143–168. https://doi.org/10.5334/pb.277

van der Linden, D., Scholte, R. H. J., Cillessen, A. H. N., te Nijenhuis, J., & Segers, E. (2010). Classroom ratings of likeability and popularity are related to the Big Five and the general factor of personality. *Journal of Research in Personality*, *44*(5), 669–672. https://doi.org/10.1016/j.jrp.2010.08.007

van Houtte, M. (2023). Understanding the gender gap in school (dis)engagement from three gender dimensions: the individual, the interactional and the institutional. *Educational Studies*, *49*(2), 260–278. https://doi.org/10.1080/03055698.2020.1842722

Vantieghem, W., & Van Houtte, M. (2018). Differences in study motivation within and between genders: An examination by gender typicality among early adolescents. *Youth and Society*, *50*(3), 377–404. https://doi.org/10.1177/0044118X15602268

Verniers, C., Martinot, D., & Dompnier, B. (2016). The feminization of school hypothesis called into question among junior and high school students. *The British Journal of Educational Psychology*, *86*(3), 369–381. https://doi.org/10.1111/bjep.12111

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174– 1204. https://doi.org/10.1037/a0036620

Wach, F. S., Spengler, M., Gottschling, J., & Spinath, F. M. (2015). Sex differences in secondary school achievement: The contribution of self-perceived abilities and fear of failure. *Learning and Instruction*, *36*, 104–112. https://doi.org/10.1016/j.learninstruc.2015.01.005

Weiner, B. (1972). Attribution theory, achievement motivation, and the educational process. *Review of Educational Research*, *42*(2), 203–215.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. Springer.

Weiner, B. (2000). Intrapersonal and interpersonal theories of motivation from an attributional perspective. *Educational Psychology Review*, *12*(1), 1–14. https://doi.org/10.1023/A:1009017532121

Weiner, B., & Kukla, A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, *15*(1), 1–20. https://doi.org/10.1037/h0029211

Weisgram, E. S., Fulcher, M., & Dinella, L. M. (2014). Pink gives girls permission: Exploring the roles of explicit gender labels and gender-typed colors on preschool children's toy preferences. *Journal of Applied Developmental Psychology*, *35*(5), 401–409. https://doi.org/10.1016/j.appdev.2014.06.004

Whitley, B. E. (1983). Sex role orientation and self-esteem: A critical meta-analytic review. *Journal of Personality and Social Psychology*, *44*(4), 765–778. https://doi.org/10.1037/0022-3514.44.4.765

Whitley, B. E. (1984). Sex-role orientation and psychological well-being: Two meta-analyses. *Sex Roles*, *12*(1–2), 207–225. https://doi.org/10.1007/BF00288048

Whitley, B. E. (1988). Masculinity, femininity, and self-esteem: A multitrait-multimethod analysis. *Sex Roles*, *18*(7–8), 419–431. https://doi.org/10.1007/BF00288393

Wilde, A., & Diekman, A. B. (2005). Cross-cultural similarities and differences in dynamic stereotypes: A comparison between Germany and the United States. *Psychology of Women Quarterly*, *29*(2), 188–196. https://doi.org/10.1111/j.1471-6402.2005.00181.x

Witvliet, M., Olthof, T., Hoeksma, J. B., Goossens, F. A., Smits, M. S. I., & Koot, H. M. (2010). Peer group affiliation of children: The role of perceived popularity, likeability, and behavioral similarity in bullying. *Social Development*, *19*(2), 285–303. https://doi.org/10.1111/j.1467-9507.2009.00544.x

Wojciszke, B., Abele, A. E., & Wiesław, B. (2009). Two dimensions of interpersonal attitudes: Liking depends on communion, respect depends on agency. *European Journal of Social Psychology*, *39*, 973–990. https://doi.org/10.1002/ejsp

Wojciszke, B., Baryla, W., Parzuchowski, M., Szymkow, A., & Abele, A. E. (2011). Self-esteem is dominated by agentic over communal information. *European Journal of Social Psychology*, *41*(5), 617–627. https://doi.org/10.1002/ejsp.791

Wolfram, H. J., Mohr, G., & Borchert, J. (2009). Gender role self-concept, gender-role conflict, and well-being in male primary school teachers. *Sex Roles*, *60*(1–2), 114–127. https://doi.org/10.1007/s11199-008-9493-4

Wolter, I. B., & Hannover, B. (2016). Gender role self-concept at school start and its impact on academic self-concept and performance in mathematics and reading. *European Journal of Developmental Psychology*, *13*(6), 681–703. https://doi.org/10.1080/17405629.2016.1175343

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In J. M. Olson & M. P. Zanna (Eds.), *Advances in experimental social psychology* (1st ed., Vol. 46, pp. 55–123). Elsevier Inc. https://doi.org/10.1016/B978-0-12-394281-4.00002-7

Woodhill, B. M., & Samuels, C. A. (2003). Positive and negative androgyny and their relationship with psychological health and well-being. *Sex Roles*, *48*(11–12), 555–565. https://doi.org/10.1023/A:1023531530272

Workman, J., & Heyder, A. (2020). Gender achievement gaps: The role of social costs to trying hard in high school. *Social Psychology of Education*, *23*(6), 1407–1427. https://doi.org/10.1007/s11218-020-09588-6

Wouters, S., Doumen, S., Germeijs, V., Colpin, H., & Verschueren, K. (2013). Contingencies of self-worth in early adolescence: The antecedent role of perceived parenting. *Social Development*, *22*(2), 242–258. https://doi.org/10.1111/sode.12010

Yailagh, M., Lloyd, J., & Walsh, J. (2009). The causal relationships between attribution styles, mathematics self-efficacy beliefs, gender differences, goal setting, and math achievement of school children. *Journal of Education & Psychology*, *3*(2), 95–114.

Yang, X., & Gao, C. (2021). Missing women in STEM in China: An empirical study from the viewpoint of achievement motivation and gender socialization. *Research in Science Education*, *51*, 1705–1723. https://doi.org/10.1007/s11165-019-9833-0

Yavorsky, J. E., & Buchmann, C. (2019). Gender typicality and academic achievement among American high school students. *Sociological Science*, *6*, 661–683. https://doi.org/10.15195/v6.a25

Young, R., & Sweeting, H. (2004). Adolescent bullying, relationships, psychological well-being, and gender-atypical behavior: A gender diagnosticity approach. *Sex Roles*, *50*(7–8), 525–537.

Younger, M., Warrington, M., & Williams, J. (1999). The gender gap and classroom interactions: Reality and rhetoric? *British Journal of Sociology of Education*, *20*(3), 325–341.

Yu, J., McLellan, R., & Winter, L. (2020). Which boys and which girls are falling behind? Linking adolescents' gender role profiles to motivation, engagement, and achievement. *Journal of Youth and Adolescence*, *50*, 336–352. https://doi.org/10.1007/s10964-020-01293-z

Zeigler-Hill, V., & Myers, E. M. (2012). A review of gender differences in self-esteem. In S. P. McGeown (Ed.), *Psychology of gender differences* (pp. 131–143). Nova.

Zucker, K. J. (1977). Sex-typing of children's toys by college students and fourth-graders. *Catalogue of Selected Documents in Psychology*, *7*, 6–7.

Zuckerman, M., Li, C., & Hall, J. A. (2016). When men and women differ in self-esteem and when they don't: A meta-analysis. *Journal of Research in Personality*, *64*, 34–51. https://doi.org/10.1016/j.jrp.2016.07.007

II. Empirical Work

Manuscript 1

Linda's Cars vs. Dominik's Dolls: How Do Pedagogical Educators in Training React to Children's Violations of Gender Stereotypes?

Streck, H. & Kessels, U. (Accepted, pending revision). Linda's Cars vs. Dominik's Dolls: How Do Pedagogical Educators in Training React to Children's Violations of Gender Stereotypes?. *Archives of Sexual Behavior*.

This manuscript has since been published in Archives of Sexual Behavior:

Streck, H., & Kessels, U. (2024). Linda's cars versus Dominik's dolls: How do pedagogical educators in training react to children's violations of gender stereotypes? *Archives of Sexual Behavior.* https://doi.org/10.1007/s10508-024-02958-7

Note: This is the first-author version of the paper submitted to Archives of Sexual Behavior. Formatting that diverges from APA 7 follows journal-specific guidelines.

ABSTRACT

The gender stereotypes adults hold can influence whether they approve or disapprove of behavior shown by children, depending on whether this behavior is in line with stereotypes. Adults report negative evaluations towards children whose behavior does not adhere to gender stereotypes, particularly towards feminine boys. Whether pedagogical educators in training show negative reactions towards children who violate gender stereotypes has not been examined. We investigate this question by firstly assessing what gender stereotypes adults hold about children in Germany. In study 1, we assessed descriptive, prescriptive, and proscriptive gender stereotypes identified by adults for children in German society. Stereotypes gathered from this first study were used to construct four vignettes of stereotypical and nonstereotypical boys and girls in order to examine how pedagogical educators in training (N = 414) evaluate these children in study 2. We investigated ratings of one of these vignettes (2 X 2 between-participants design) regarding liking, perceived competence, creativity, self-esteem, prosocial behavior, as well as internalizing and externalizing problems. A series of ANOVAs revealed that girls displaying masculine behavior received advantageous ratings on competence, creativity, and self-esteem, while boys showing femininity were perceived as the most prosocial. More than gender nonconformity, masculinity and femininity strongly related to externalizing and internalizing problems, respectively. We review how our results in Germany differ from the literature originating in the US, as we did not find backlash for feminine boys. Possible bias against femininity and towards masculinity within society and cultural and sampling factors are discussed.

Keywords: Gender Nonconformity, Gender Stereotypes, Gender Socialization, Children, Toys

INTRODUCTION

Social Cognitive Theory: Relevance of Stereotypes for Children

As they develop over childhood, children become more understanding of the complex phenomenon of gender. In theoretical frameworks the role of social interactions and feedback is prominently featured (see Coyle & Fulcher, 2022). Social cognitive theory (Bussey & Bandura, 1999) highlights that an individual's environment plays a major role in how gender is understood and enacted. From birth, children are exposed to three social modes of influence: modeling, enactive experience, and direct tuition (Bussey & Bandura, 1999). The relevance of gender stereotypes held by adult socializers is particularly pronounced in enactive experience and direct tuition: Enactive experience refers to how children display certain behavior and use cues of approval or disapproval by others to guide their future behavior. Direct tuition encompasses direct, verbal feedback from others about the appropriateness of one's gendered behavior. A boy playing with dolls may note heads shaking, or frowning and interpret that this behavior, which, when enacted by his sister, was met with encouragement, is not appropriate for himself on the basis of his gender (Bussey, 2011).

Gender Stereotypes

Gender stereotypes have been defined as culturally shared characteristics ascribed to men and women based on their gender (Myers et al., 2010, p. 467). Stereotypes about men and women fall into the categories agency and communion (Bakan, 1966), respectively. Agency refers to ambition, independence, and self-determination (Abele & Wojciszke, 2013) and is closely related to competence (Cuddy et al., 2008), instrumentality (Spence et al., 1974), and masculinity (Bem, 1981a). Communion, on the other hand, refers to an orientation towards other people, caring, and kindness (Abele & Wojciszke, 2013) and relates to warmth (Cuddy

et al., 2008), expressiveness (Spence et al., 1974), and femininity (Bem, 1981a). Stereotypes, however, not only reflect traits but also appearance and behaviors deemed gender-appropriate.

Descriptive, Prescriptive, and Proscriptive Gender Stereotypes

Researchers have posited that all stereotypes could be considered "descriptive" stereotypes, meaning that they describe how members of stereotyped groups *typically* behave (Koenig, 2018; Prentice & Carranza, 2002; Rudman & Glick, 2010; Sczesny et al., 2019). A subset of these stereotypes, "prescriptive" stereotypes, contain information about how each gender *should* behave (Prentice & Carranza, 2002; Rudman & Glick, 2010). "Proscriptive" gender stereotypes, a further subset of prescriptive stereotypes, contain information about how each gender *should not* behave (Rudman & Glick, 2010). Descriptive stereotypes serve as a cognitive simplification tool, allowing us to save time and energy when confronted with stereotyped group members (Eckes, 2008; Rudman & Glick, 2010). Prescriptive and proscriptive stereotypes, however, originate in system justification, where their function as guidance for behavior ensures that group members adhere to acceptable norms (Eckes, 2008; Rudman & Glick, 2010).

Gender Stereotypes About Children

Investigations into the content of stereotypes have focused mainly on adults, though there are findings showing gendered stereotypes of children (Koenig, 2018; Martin, 1995; Sullivan et al., 2018, 2022; Zucker, 1977). Generally, stereotypes about children relate to their appearance and activities rather than their traits. Girls were described as (enjoying) playing with dolls and boys as (enjoying) playing with trucks (Koenig, 2018; Sullivan et al., 2018, 2022). Appearing masculine and feminine and playing with masculine and feminine toys were also prescriptive stereotypes for boys and girls, respectively (Koenig, 2018; Sullivan et al., 2018). Proscriptive stereotypes for boys related to appearing feminine and engaging with

feminine activities (Koenig, 2018) and being dirty or appearing masculine for girls (Sullivan et al., 2022).

Reactions to Violations of Gender Stereotypes

A wide variety of investigations showed stereotype violations are punished by others, with most research focusing on adults (Moss-Racusin et al., 2010; Nauts, 2015; see Rudman, Moss-Racusin, Glick, et al., 2012; Rudman, Moss-Racusin, Phelan, et al., 2012; Rudman & Glick, 2001; Sanborn-Overby & Powlishta, 2020). Adolescents (Braun & Davidson, 2017; Young & Sweeting, 2004) and children (Blakemore, 2003; Kwan et al., 2020; Xiao et al., 2019; Zucker et al., 1995) also experience negative consequences for counterstereotypical behavior from their peers.

Reactions to Violations of Stereotypes by Children

A number of studies showed that adults reacted more negatively to children who defied gender stereotypes than to children who adhered to stereotypes (Cahill & Adams, 1997; Coyle et al., 2016; Fagot, 1977; Feinman, 1981; Martin, 1990; Sandnabba & Ahlberg, 1999; Sullivan et al., 2018; Thomas & Blakemore, 2013). Some of these studies were conducted in the 1980s and 90s and social values can shift over these timeframes (Eagly et al., 2020). In the following sections we will thus focus more heavily on more recent research.

Likability

The perceived likability of children was claimed to be a "primary measure of interest" (Sullivan et al., 2018, p. 13) in one study from the United States, which found children described as atypical to be liked less than their typical peers by adults (Sullivan et al., 2018). Girls described as behaving stereotypically feminine were rated as most likable, significantly more so than masculine girls (Sullivan et al., 2018). A different vignette study also found the typical girl, characterized as a girl showing positive femininity, to be the most likable.

"Tomboys", girls enacting positive masculinity, were ambivalently evaluated (Coyle et al., 2016). Ratings of likability of typical and nontypical boys revealed boys enacting femininity to be rated as the least likable (Sullivan et al., 2018). "Sissies", boys enacting negative femininity, were similarly not well-liked (Coyle et al., 2016). Boys showing positive femininity ("mama's boys"), however, were more tolerated by adults.

A more specific sample of adults, preschool teachers, had more lenient attitudes towards girls showing nonconformity but reacted more negatively towards nonconforming boys (Cahill & Adams, 1997). Boys engaging in feminine play also received more criticism, from teachers (Fagot, 1977). However, in some work teachers reported more positive evaluations towards feminine boys compared to masculine boys (Piché & Plante, 1991).

Competence

Past research from the United States has investigated the perceived competence of gender-typical and atypical children (Coyle et al., 2016). Adults rated a hypothetical typical girl as the most competent but also indicated finding a girl described as enacting positive masculinity rather competent. Similarly, target boys enacting positive femininity were rated as rather competent, more so than the typical boy, but "sissies" were found to be the least competent (Coyle et al., 2016). The boys and girls described as showing positive femininity and masculinity, respectively, were not penalized in terms of competence, but a boy displaying negative femininity was. These findings support linking perceived competence with gender typicality but also suggest that some gender atypicality can have positive associations with competence ratings. Furthermore, a study with adolescents showed targets described as gender atypical to be perceived as more competent than gender stereotypical targets, this effect being significant in participants with high SES (Meimoun et al, 2023).

Experiencing Problems

Studies have measured the extent to which adults believed gender nonconforming children experience problems in their happiness, relationships with others (Coyle et al., 2016), and psychological adjustment (Thomas & Blakemore, 2013). Internalizing tendencies can encompass emotional problems and social isolation from peers according to the Strength and Difficulties Questionnaire, an established measure of adjustment (Goodman, 2001; Goodman et al., 2010). Externalizing problems may relate to hyperactivity or behavioral conduct issues (Goodman, 2001; Goodman et al., 2010). Past researchers constructed scales of externalizing and internalizing tendencies, encompassing aggression, misconduct and self-esteem, worrying, respectively. They revealed it was not nonconformity itself but the typing of behavior as masculine and feminine that predicted these latter ascribed pathologies (Thomas & Blakemore, 2013). Masculine behavior related to whether children were believed to display externalizing tendencies, while feminine behavior related to internalizing tendencies, irrespective of child gender (Thomas & Blakemore, 2013).

Differential Reactions to Children's Stereotype Violations by Gender

A persistent theme in this literature is the comparatively more positive attitudes towards nonconforming girls, when compared to nonconforming boys (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013). Greater disapproval for boys displaying counterstereotypical behavior is common in adult (Feinman, 1981; Martin, 1990) and even teacher (Fagot, 1977) samples. It has been theorized that the harsher reaction to male violations of stereotypes is due to society's patriarchal values, which place masculinity above femininity in terms of status and desirability (Berger et al., 1972; Feinman, 1981; Nauts, 2015; Rudman & Glick, 2010).

Possibility of Positive Effects for Gender Nonconformity?

Much research has focused on the pressing issue of negative backlash towards gender atypical children (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013), but gender atypicality can be evaluated positively by others (Bochicchio et al., 2019; Meimoun et al., 2023). Empirical and theoretical associations between perceived prosocial behavior, self-esteem, and creativity are explored in the present study.

Prosocial Behavior

Prosocial behavior is defined as "voluntary actions intended to benefit another" (Skoe et al., 2002, p. 296) and tends to be stereotyped as feminine (Eisenberg et al., 2006; Quenneville et al., 2022). Evidence from educational settings showed feminine boys to be perceived as more prosocial than masculine boys by teachers (Piché & Plante, 1991). The link between femininity and prosociality has also been cited in research conducted in daycare. Young children showed no significant gender differences in observed prosocial behavior, but their childhood educators perceived girls as significantly more prosocial than boys (Bouchard et al., 2015).

Creativity

Creativity has been linked to artistic pursuits (Dumas & Dunbar, 2016) but can also capture the ability to think and behave in unique ways that "diverge from the normative" (Proudfoot et al., 2015, p. 1751). Individuals showing greater gender atypicality could be characterized as displaying eccentric behavior and behaving in ways that deviate from the norm. Empirically, emphasizing an individual's eccentricity has been related to an increase in perceived creativity of said individual (Van Tilburg & Igou, 2014). Such associations are not surprising given cultural stereotypes of eccentric geniuses or peculiar artists; defying social norms could be seen as thinking "outside the box," a persistent belief about creative people

(Proudfoot et al., 2015; Weisberg, 2010). These results suggest that gender atypicality could be associated with being evaluated as being more creative.

Self-Esteem

Self-esteem refers to a person's evaluation of themselves, with high self-esteem being a positive sense of self-regard (Zeigler-Hill & Myers, 2012). Actual rates of self-esteem in gender atypical people can be low (Egan & Perry, 2001), due in part to factors such bullying (Hu et al., 2023) and restrictive societal norms (Zentner & Von Aufsess, 2022). However, actual and perceived self-esteem can (Zeigler-Hill et al., 2013), but do not necessarily correlate (Kilianski, 2008; Watson et al., 2002). *Perceived* self-esteem may thus offer different insight and is worth investigating.

Greater adherence to gender typical behavior has been associated with greater felt pressure to conform to gender norms (Cook et al., 2019, p. 1913), which in turn was associated with low self-esteem (Egan & Perry, 2001; Good & Sanchez, 2010; Skinner et al., 2018). It could follow that those who display counterstereotypical behavior might have freed themselves from pressure to conform, benefiting their self-esteem. Individuals who display some gender atypicality are perhaps more likely to be ascribed greater self-esteem by others due to their disengagement from societal expectations around gender, despite facing possible backlash.

Cultural Considerations

The available investigations into gender stereotypes of young children and reactions to violations of stereotypes by adults have been primarily conducted in the United States. It would be beneficial to establish the content of gender stereotypes for this age group in a European sample, as evidence showed that gender stereotypes can differ between German and American cultures (Wilde & Diekman, 2005). Investigations into reactions towards gender atypical children and adolescents from European countries showed differences to findings

from the United States (Sullivan et al., 2018; Thomas & Blakemore, 2013), with the former finding some positive opinions regarding gender atypical individuals (Bochicchio et al., 2019; Meimoun et al., 2023). It is not established whether it is possible to extrapolate findings on gender differences from American to German or even European society.

Sampling Considerations

While previous studies offered important insight into gender stereotypes and adult reactions to child development (Sullivan et al., 2018; Thomas & Blakemore, 2013), the adults sampled were not chosen for their familiarity with children. Parents and teachers represent relevant socializing agents from which children receive subtle or explicit feedback about gender appropriate and inappropriate behavior, according to social cognitive theory (Bussey & Bandura, 1999). Investigating how pedagogical faculty engages with gender nonconformity is especially relevant, as they will experience interactions with a wide variety of children as compared to most adults and even parents. Specifically, we will focus on faculty in training, as they represent a new generation in the educational workforce and can give insight into whether current training is preparing teachers for fostering gender sensitivity in educational settings (Koordinationsstelle: Chance Quereinstieg/Männer in Kitas, 2019; Kultusministerkonferenz, 2016).

Study Overview

The overall purpose of this research is thus twofold: firstly, to examine gender stereotypes about children in German society. Secondly, to investigate the impact that violations of these stereotypes, as well as masculinity and femininity, have on evaluations of children by educational professionals in training.

In order to investigate these questions, our current study was influenced by the empirical work of Sullivan, Moss-Racusin, Lopez, and Williams (2018), whose investigation

ascertained descriptive, prescriptive, and proscriptive stereotypes about children in American society and tested whether violations of these stereotypes related to backlash from adults. We follow the outline of their experiments, adapting material as necessary in order to investigate stereotypes in German society and with a more specific sample of adults, namely pedagogical educators in training.

We will first investigate stereotypes about 3-year-old children in German society. Descriptive, prescriptive, and proscriptive stereotypes will be gathered in study 1 in order to collect material for vignettes, which will be used in study 2, a procedure used in Sullivan et al. (2018). Following established procedure (Koenig, 2018; Nauts, 2015; Rudman, Moss-Racusin, Phelan, et al., 2012; Sullivan et al., 2018), our first study will ask adult participants to indicate the degree to which certain characteristics are considered typical or desirable for young children by German society. This should reveal descriptive, prescriptive, and proscriptive stereotypes about children in Germany.

Using the stereotypes from study 1, we will investigate whether children adhering to or violating these stereotypes will be regarded differently by pedagogical educators in training. With our investigation, we hope to expand upon findings in past literature, using a new cultural and sampling context but keeping the established methods used previously (Sullivan et al., 2018). The past literature outlines an effect of gender atypicality on liking, finding that particularly feminine boys were more disliked than other children, and that girls tended to be rated higher in terms of liking (Sullivan et al., 2018). We hypothesize that children behaving in a nonstereotypical manner will be liked less than typical children (H1a) and that the atypical boy will receive the greatest backlash in terms of liking (H1b). We also expect girls to receive higher ratings in terms of liking than boys (H1c).

The relation between competence and gender atypicality is not straightforward. For children, findings showed that behaving atypically, enacting positive masculinity and

femininity, can be perceived as rather competent, although typical girls were rated as most and "sissies" as the least competent (Coyle et al., 2016), suggesting additional research is necessary. We expect that ratings of competence will differ between children behaving in line with and in violation of gender stereotypes (H2).

In terms of internalizing and externalizing problems, gender atypicality was not a relevant factor in past research; instead, gender typing of behavior as masculine or feminine predicted externalizing and internalizing tendencies, although this past study did not use established measures (Thomas & Blakemore, 2013). We hypothesize that perceptions of internalizing and externalizing tendencies will differ between masculine and feminine children (H3a). Specifically, perceptions of internalizing problems should be higher in feminine children (H3b). Masculine children are expected to be perceived as showing greater externalizing tendencies (H3c).

Since gender atypicality may be associated with further constructs, we will also investigate prosociality, creativity, and self-esteem. Prosociality has been related to atypicality, showing atypical boys were perceived as more prosocial (Piché & Plante, 1991), likely due to the stereotyping of helping behaviors as feminine (Eisenberg et al., 2006; Quenneville et al., 2022). We hypothesize that atypical boys (H4a) and feminine children (H4b) will receive higher ratings of prosociality.

A theoretical link between gender atypicality and being evaluated as being creative and as having high self-esteem is possible due to the associations that "rejecting the norm" has in society. We hypothesize that stereotype violating children will be perceived as more creative than those who conform to gender stereotypes (H5). Similarly, children who behave in gender nonconforming ways are hypothesized to be perceived as having higher self-esteem than children who conform to gender stereotypes (H6).

STUDY 1 METHODS

Participants

An a-priori power analysis was conducted to estimate the sample size needed to find effect sizes of at least $|d| = 0.4$ (Koenig, 2018; Nauts, 2015; Rudman, Moss-Racusin, Phelan, et al., 2012; Sullivan et al., 2018) with power of 80%. The power analysis was completed in G*Power 3 (Faul et al., 2007) and revealed that a between-participants design would require 100 participants per cell. Just shy of this goal, we recruited 397 participants from a university in a large German city to take part in this research. Participants who indicated that they did not understand the instructions or were missing more than 25% of the data were excluded from this analysis ($n = 7$). Of the remaining 390 participants, $n = 250$ identified as female, while $n = 104$ identified as male and a further $n = 12$ indicated identifying as a different gender and $n = 24$ did not indicate a gender. The mean age of the sample was 23.01 ($SD = 3.97$, range = 18 - 43). A total of 71% of participants came from what would be classed as humanities, while 24% studied natural sciences, with roughly 5% of participants not indicating their studied subject.

Materials and Procedure

A list of 72 characteristics was compiled for this study. The items were based on items from past research (Kessels, 2005; Koenig, 2018; Liben & Bigler, 2002; Sullivan et al., 2018; Tobin et al., 2010) and in certain cases modified to suit a German sample (e.g., changing "baseball" to "soccer"). Items included: "plays soccer" and "dances ballet," as well as "brave" and "helpful."

The research took place during regular class time and was completed using paper and pencil questionnaires under supervision of the researcher. After reading a short introductory text outlining consent and the nature of the experiment, participants moved onto the experimental phase. Participants were made aware that their judgments should not be based

on their own opinion of what is typical or desirable but instead should reflect what they believed to be typical or desirable "in German society." Participants were instructed to indicate "how typical [or desirable, depending on condition] it is in German society for three-year-old boys [or girls, depending on condition] to display the following traits and behaviors." The prompt was based on past research (Koenig, 2018; Prentice & Carranza, 2002; Rudman, Moss-Racusin, Phelan, et al., 2012; Sullivan et al., 2018) and modified for a German speaking sample. Participants were asked to rate these items in terms of their typicality or desirability on a 9-point Likert scale ranging from 1 (not at all [typical/desirable] for a [boy/girl]) to 9 (very [typical/desirable] for a [boy/girl]). Our 2 (target gender: male, female) X 2 (stereotype rating: descriptive, prescriptive) design was conducted between-participants.

STUDY 1 RESULTS

Analytical Plan

The method for determining descriptive and pre-/proscriptive stereotypes for girls and boys (or women and men, in the case of other research) was outlined by Rudman, Moss-Racusin, Phelan, et al. (2012) and adapted from Prentice and Carranza (2002) and has since been used in multiple studies (Koenig, 2018; Nauts, 2015; Sullivan et al., 2018, 2022). To attain descriptive stereotypes, in a first step, the means for each item are calculated separately for boys and girls. Independent samples *t*-tests are conducted for each item, comparing the typicality rating for boys and girls. A trait is deemed a descriptive stereotype if (1) its mean is above a "6," (2) the independent samples *t*-test for ratings between girls and boys is significant, and (3) the effect size for the difference is larger than $\left| d \right| > 0.4$. All three criteria must be satisfied for a characteristic to be labeled a descriptive stereotype.

Prescriptive stereotypes are obtained through similar criteria, though these stereotypes are drawn from the scales assessing desirable traits. In order for a characteristic to be labeled a prescriptive stereotype, the mean for items must be rated above a 6 on the desirability scale,

the difference between ratings for boys and girls must reach the significance threshold, and the effect size for the difference needs to be larger than $\left| d \right| > 0.4$.

Proscriptive stereotypes differ slightly in that the mean on the desirability scale must be below 4 and thus considered undesirable. Similar to descriptive and prescriptive stereotypes, the differences between ratings of girls and boys must be significant and have an effect size larger than $\left| d \right| > 0.4$.

Analyses

We analyzed both the typicality and desirability ratings for boys and girls of 72 items. After calculating descriptive statistics, we conducted a series of independent samples *t*-tests, comparing ratings for boys and girls and ascertaining Cohen's *d* effect sizes for each *t*-test. Since items were rated both on typicality and desirability, it is possible for items to be considered both descriptive, prescriptive, and proscriptive (or any combination thereof).

Descriptive, Prescriptive, and Proscriptive Stereotypes

We found a large number of descriptive items for both boys and girls; see Table 1 for items. The items that showed the highest means and strongest effect sizes for descriptive stereotypes pertained to activities and appearance of children rather than more internal traits. We found a large number of prescriptive stereotypes for girls and boys. The items with the largest effect sizes related to appearance and play preferences, a finding also displayed by characteristics rated as desirable for boys. We found fewer proscriptive stereotypes, compared to both descriptive and prescriptive stereotypes. The most "forbidden" characteristics for boys (lowest means) were appearance related items, like wearing dresses and skirts or wearing nail polish. The least desirable traits for girls to have related to behaviors such as fighting, being loud, or playing with a wooden sword. A correlation between items showed that items measuring typicality and desirability for ratings for girls was $r = .88$, while a similar analysis for boys displayed a correlation of $r = .89$.

Table 1

Means, Standard Deviations, and Cohen's d for Items on the Descriptive and Prescriptive Scales

Item	Descriptive Scale			Prescriptive Scale		
	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*
1. Plays rough and tumble games	2.63 (1.30)	7.27 (1.67)	-3.11***	2.31 (1.34)	5.56 (2.01)	-1.90***
2. Plays with Playmobil™	5.52 (1.88)	6.75 (1.61)	-0.71***	5.52 (1.62)	6.65 (1.81)	-0.66***
3. Has a dollhouse	7.69 (1.46)	2.38 (1.63)	3.44***	7.34 (1.67)	2.71 (1.48)	2.94***
4. Builds paper planes	4.38 (1.69)	6.3 (1.71)	-1.13***	4.77 (1.47)	7.13 (1.66)	-1.51***
5. Plays with a wooden sword♂	2.67 (1.35)	7.25 (1.49)	-3.22***	3.1 (1.52)	7.28 (1.52)	-2.75***
6. Fearless	3.96 (1.68)	5.75 (1.60)	-1.09***	4.38 (1.68)	7.04 (1.8)	-1.53***
7. Wild	4.43 (1.68)	6.82 (1.42)	-1.53***	3.17 (1.62)	6.42 (1.68)	-1.97***
8. Plays soccer	3.33 (1.56)	7.58 (1.62)	-2.68***	4.16 (1.9)	7.82 (1.49)	-2.15***
9. Plays cops and robbers♂	3.93 (1.89)	7.26 (1.52)	-1.94***	3.95 (1.9)	7.35 (1.5)	-2.00***
10. Shy	6.42 (1.53)	4.62 (1.16)	1.32***	5.95 (1.7)	3.42 (1.33)	1.66***
11. Plays with tools	2.73 (1.49)	6.52 (1.71)	-2.37***	3.51 (1.68)	7.05 (1.6)	-2.16***
12. Tough	4.98 (1.9)	5.67 (1.65)	-0.39***	4.59 (1.63)	6.7 (1.58)	-1.31***
13. Dominant	4.27 (1.94)	5.64 (1.75)	-0.74**	3.38 (1.75)	6.07 (1.93)	-1.46***
14. Wears pink	7.48 (1.69)	2.01 (1.18)	3.74***	6.86 (1.6)	2.53 (1.38)	2.91***
15. Plays house	7.81 (1.63)	3.48 (1.81)	2.51***	7.16 (1.6)	4.28 (1.96)	1.60***
16. Has short hair	3.49 (1.57)	7.8 (1.54)	-2.77***	3.79 (1.46)	7.29 (1.6)	-2.29***
17. Rebellious	4.31 (1.65)	6.42 (1.48)	-1.35***	3.05 (1.53)	5.64 (1.74)	-1.58***
18. Collects stickers	6.45 (1.71)	4.89 (1.93)	0.85***	6.34 (1.39)	5.1 (1.69)	0.80***
19. Has long hair	7.72 (1.4)	3.16 (1.37)	3.29***	7.48 (1.51)	3.61 (1.54)	2.54***
20. Helpful	6.8 (1.44)	4.86 (1.12)	1.50***	7.94 (1.2)	6.56 (1.88)	0.87***
21. Brave	5.28 (1.51)	6.18 (1.41)	-0.61***	6.08 (1.65)	7.39 (1.48)	-0.83***

Item	Descriptive Scale			Prescriptive Scale		
	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*
22. Wears blue	4.38 (1.53)	7.16 (1.52)	-1.82***	4.67 (1.19)	6.63 (1.47)	-1.46***
23. Persistent	4.77 (1.51)	6.06 (1.28)	-0.93*	3.84 (1.62)	6.02 (1.44)	-1.43***
24. Plays in a toy-kitchen♀	7.48 (1.25)	3.54 (1.72)	2.63***	7.09 (1.51)	3.83 (1.65)	2.06***
25. Plays outside	7.02 (1.41)	7.46 (1.56)	-0.29*	6.94 (1.67)	8.00 (1.33)	-0.71***
26. Plays with cars♂	3.42 (1.67)	8.07 (1.13)	-3.25***	3.94 (1.44)	7.7 (1.35)	-2.70***
27. Plays dress-up	7.56 (1.26)	4.84 (2.05)	1.60***	7.03 (1.59)	4.87 (1.81)	1.27***
28. Considerate	6.55 (1.61)	4.2 (1.38)	1.56***	7.76 (1.37)	6.26 (1.94)	0.89***
29. Has good manners	6.53 (1.49)	4.15 (1.35)	1.68***	8.07 (1.32)	6.96 (1.69)	0.73***
30. Likes horses	7.45 (1.36)	3.26 (1.69)	2.74***	6.68 (1.42)	3.84 (1.64)	1.86***
31. Proud	5.47 (1.28)	6.06 (1.41)	-0.44***	5.41 (1.49)	6.25 (1.62)	-0.54***
32. Likes sports	5.37 (1.62)	7.43 (1.35)	-1.38***	6.29 (1.42)	8.05 (1.00)	-1.44***
33. Plays with blocks	4.51 (1.77)	7.56 (1.29)	-1.96***	5.09 (1.31)	7.54 (1.22)	-1.93***
34. Plays with dolls♀	8.08 (0.98)	2.43 (1.37)	4.77***	7.41 (1.47)	2.83 (1.45)	3.14***
35. Wears nail polish	6.26 (2.25)	1.75 (1.43)	2.39***	5.69 (2.26)	1.98 (1.44)	1.98***
36. Patient	4.96 (1.82)	3.56 (1.46)	0.85***	7.27 (1.37)	5.84 (1.92)	0.86***
37. Sings in a children's choir	5.31 (1.5)	3.63 (1.53)	1.11***	5.88 (1.67)	4.49 (1.47)	0.88***
38. Neat	6.21 (1.75)	3.19 (1.38)	1.92***	7.55 (1.51)	5.66 (1.68)	1.18***
39. Obedient	6.36 (1.55)	4.04 (1.31)	1.62***	7.72 (1.38)	6.01 (1.85)	1.04***
40. Jumps rope	6.79 (1.53)	4.11 (1.59)	1.72***	6.59 (1.42)	4.73 (1.54)	1.26***
41. Sensible	6.18 (1.7)	4.08 (1.22)	1.41***	7.54 (1.34)	6.01 (1.86)	0.94***
42. Active	6.27 (1.44)	7.63 (1.23)	-1.01***	6.47 (1.6)	7.88 (1.14)	-1.02***
43. Draws dresses♀	7.18 (1.3)	2.25 (1.41)	3.63***	6.65 (1.69)	2.77 (1.46)	2.46***
44. Likes princesses♀	8.07 (1.2)	2.04 (1.44)	4.55***	7.31 (1.64)	2.3 (1.27)	3.41***
45. Cuddly	7.04 (1.48)	4.55 (1.7)	1.57***	7.24 (1.46)	4.77 (1.76)	1.53***

Item	Descriptive Scale			Prescriptive Scale		
	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*
46. Talkative	6.43 (1.39)	5.61 (1.76)	0.52***	6.44 (1.64)	6.38 (1.45)	0.04
47. Wears skirts and dresses	7.94 (1.22)	1.55 (1.1)	5.49***	7.34 (1.55)	1.9 (1.14)	4.00***
48. Goes fishing	2.13 (1.34)	5.16 (1.94)	-1.82***	3.53 (1.65)	6.02 (1.58)	-1.55***
49. Strong	3.57 (1.84)	6.33 (1.47)	-1.66***	4.04 (1.78)	7.17 (1.46)	-1.92***
50. Energetic	5.28 (1.72)	7.00 (1.31)	-1.12***	4.4 (1.81)	6.91 (1.28)	-1.61***
51. Emotional	7.38 (1.02)	4.63 (2.01)	1.74***	6.43 (1.82)	4.07 (1.8)	1.30***
52. Dances ballet♀	7.21 (1.54)	1.69 (1.11)	4.09***	6.77 (1.62)	2.24 (1.38)	3.01***
53. Cautious	6.22 (1.52)	3.65 (1.34)	1.80***	7.04 (1.4)	4.61 (1.65)	1.59***
54. Does horseback-riding	6.43 (2.04)	2.66 (1.57)	2.07***	6.41 (1.55)	3.42 (1.72)	1.82***
55. Likes dinosaurs	3.44 (1.67)	8.08 (1.08)	-3.28***	4.24 (1.6)	7.57 (1.2)	-2.36***
56. Effortful	6.52 (1.62)	4.29 (1.38)	1.48***	7.72 (1.33)	6.3 (1.75)	0.91***
57. Likes superheroes	3.98 (1.73)	7.89 (1.3)	-2.55***	4.46 (1.56)	7.34 (1.35)	-1.97***
58. Loving	7.00 (1.39)	4.88 (1.67)	1.38***	7.44 (1.29)	5.64 (1.67)	1.21***
59. Friendly	7.07 (1.29)	5.48 (1.14)	1.31***	8.22 (1.24)	7.01 (1.57)	0.86***
60. Plays hop-scotch	7.12 (1.46)	4.27 (1.73)	1.79***	7.02 (1.53)	4.65 (1.81)	1.42***
61. Climbs trees	4.99 (1.8)	7.43 (1.24)	-1.58***	4.63 (1.73)	7.23 (1.33)	-1.69***
62. Has a chemistry set	2.97 (1.87)	5.74 (1.67)	-1.56***	3.81 (1.82)	6.24 (1.66)	-1.39***
63. Recites poems	5.41 (1.93)	3.15 (1.56)	1.29***	5.99 (1.77)	4.39 (1.69)	0.93***
64. Does gymnastics	6.41 (1.6)	3.3 (1.66)	1.91***	6.6 (1.34)	4.29 (1.7)	1.51***
65. Ice-skates	5.85 (1.89)	3.92 (1.67)	1.08***	6.16 (1.48)	4.55 (1.94)	0.93***
66. Adventurous	4.91 (1.71)	7.48 (1.21)	-1.74***	4.64 (1.66)	7.41 (1.32)	-1.85***
67. Draws rocketships♂	2.98 (1.54)	7.41 (1.4)	-3.00***	3.83 (1.58)	7.15 (1.48)	-2.17***
68. Self-confident	5.27 (1.65)	6.53 (1.3)	-0.85***	5.62 (1.72)	7.27 (1.58)	-1.00***
69. Courageous	5.03 (1.65)	6.96 (1.33)	-1.29***	5.29 (1.65)	7.58 (1.43)	-1.49***

Item	Descriptive Scale			Prescriptive Scale		
	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*	Mean (SD) Girl	Mean (SD) Boy	Cohen's *d*
70. Flies a kite	4.89 (1.78)	6.89 (1.3)	-1.28***	5.03 (1.57)	7.07 (1.31)	-1.42***
71. Loud	5.06 (2.12)	7.5 (1.42)	-1.35***	2.8 (1.83)	5.82 (1.98)	-1.59***
72. Collects soccer-cards♂	2.02 (1.41)	7.51 (1.57)	-3.68***	3.08 (1.65)	7.23 (1.5)	-2.63***

Note. Items with ♀ or ♂ denote these were used to construct feminine or masculine vignettes, respectively. Range 1-9. $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

STUDY 1 DISCUSSION

In many ways, our results are in line with the previous establishment of gender stereotypes for young children in the United States. We found that (1) most stereotypes address appearance and activities rather than traits; (2) there is considerable overlap between descriptive and prescriptive stereotypes; and (3) there is a lower number of proscriptive stereotypes when compared to descriptive and prescriptive stereotypes. Our results differ from the past literature in that we found a very large number of stereotypes and that our effect sizes appear to be larger compared to past research (Sullivan et al., 2018).

Using information gathered in study 1, we were able to construct vignettes for our second study. We chose to focus on items that belong in the category of descriptive and prescriptive for one gender and proscriptive for the other gender, as these showed the largest differences between ratings for girls and boys. We chose 5 items to signify masculinity and femininity, respectively. These items depict behavior relating to play-activities observed in a kindergarten, such as playing with dolls or drawing a rocket ship. We will use these gender stereotyped behaviors to investigate whether children who display masculine or feminine behavior in line with their gender will be evaluated differently than children who defy gender stereotypes.

STUDY 2 METHODS

Participants

We conducted an a-priori power analysis in G*Power (Faul et al., 2007) showing our study to require around 380 participants. We recruited a sample of pedagogical educators in training. In Germany, these pedagogical educators are trained to be able to work with children, adolescents, and young adults in environments such as kindergartens, care homes, and youth centers (Bundesagentur für Arbeit, 2023).The process of becoming a pedagogical educator in this federal system involves practical experience in a pedagogical environment prior to taking classes in vocational school for at least 6 semesters while also working in an educational institution for an extended amount of time to gain further practical experience. We recruited 6 vocational schools from two federal states in and around a large German city and sampled 448 of their pedagogical educators in training. The sample was 75% female, which, while being heavily skewed, is actually more equal in terms of gender representation than actual kindergarten teachers, a job dominated by female staff (93% female) (Autorengruppe Fachkräftebarometer, 2021). The remaining 23% were male, while 1% did not indicate a gender, or indicated a gender outside the male-female binary. The mean age of our sample was 27.5 (*SD* = 8.4), with a range between 18 and 59 years. The average number of years of vocational school completed by our sample was 1.7 (*SD* = 0.68), with 42% of the sample having completed their first year and a further 42% having completed their second year. On a 10-point Likert scale ranging from 0-9, participants rated their mean level of experience with young children as 7.4 (*SD* = 2.0). Ethical approval was received from the federal Senate Department for Education, Youth, and Family.

Materials and Design

We developed four vignettes in which a boy or a girl behaved in a masculine or feminine fashion. Two of our vignettes were thus "typical" (boy showing masculine behavior,

girl showing feminine behavior), while two vignettes showed "nontypical" behavior (boy showing feminine behavior, girl showing masculine behavior). The names used in vignettes (Dominik or Linda) were chosen from a large dataset, which allowed us to match names in terms of average levels of perceived education, attractiveness, intelligence, warmth, and competence (Nett et al., 2020).

This 2 X 2 design was conducted between-participants, with the gender of the target child (male, female) and the behavior of the target child (masculine, feminine), varying between respondents. Participants thus only received one vignette, featuring a description of a masculine or feminine child, that was either a boy (Dominik) or a girl (Linda). The number of participants was roughly equal in each condition.

The vignettes, included in full in the appendix, feature a dialogue between a kindergarten teacher and a pedagogical educator in training (Natalie), who is an apprentice in the kindergarten where she gathered work experience for vocational school. Natalie describes the play activities (e.g., playing with a wooden sword and playing "cops and robbers" or playing with dolls and playing in a toy kitchen) and drawing preferences (e.g., drawing a dress or drawing a rocket ship) of a child (Dominik or Linda) with whom she is not familiar, with the fellow kindergarten teacher offering more information about the child's preferences (e.g., liking cars or princesses) and hobbies (e.g., dances ballet or collects soccer cards) and saying that such activities are very typical of this child.

Dependent Variables

We constructed scales in order to measure perceived liking and competence of the child. Whether the rater liked the child was assessed with a single item measure ("How much do you like this child?" – Personal liking) rated on a 7-point Likert scale, ranging from "not at all" to "very much." A further three items (on the same 7-point Likert scale) measured how much other children and faculty would like this child ("To what extent do you think that other

children would enjoy playing with this child?" – Perceived liking by others), with this scale showing an acceptable reliability (Cronbach's α = .72). Competence was rated using a three-item scale consisting of items such as "Does the child seem competent?", similarly rated on a 7-point Likert scale, which showed acceptable reliability (Cronbach's α = .76). Items were derived from the literature (Sullivan, 2018) and translated and adapted for a German sample.

We further included single items relating to perceived creativity and self-esteem of the child, with each item being rated on a 7-point Likert scale ranging from "not at all creative" to "very creative" and "not a lot of self-esteem" to "a lot of self-esteem," respectively.

Next, we measured prosocial behavior, internalizing, and externalizing tendencies using the German version of the strength and difficulties questionnaire, the SDQ-Deu (Goodman, 1997, 2001; Klasen et al., 2000; Koglin et al., 2007; Petermann et al., 2010), which consists of 25 items and has 5 subscales (5 items each): prosocial behavior, emotional problems, conduct problems, hyperactivity, peer problems. Items are rated on a 3-point Likert scale: "does not apply" (0), "partly applies" (1), "strongly applies" (2).

The scale includes items such as "restless, overactive" (hyperactive subscale, part of externalizing problems), "often fights with other children" (conduct problems subscale, part of externalizing problems), "often unhappy, downhearted" (emotional problems subscale, part of internalizing problems), "rather solitary, tends to play alone" (peer problems subscale, part of internalizing problems), and "kind to younger children" (prosocial behavior subscale). 5 items were reverse coded in line with SDQ scoring procedures (https://www.sdqinfo.org/py/sdqinfo/c0.py) and means for subscales were calculated from relevant items. Prosocial behavior was obtained by summing scores from the relevant subscale and showed acceptable internal reliability α = .77. An indication of internalizing and externalizing behavior was obtained by summing the subscales peer problems and emotional problems (internalizing behavior) and the subscales hyperactivity and conduct problems

(externalizing behavior). Both composite scales showed acceptable reliability (internalizing behavior Cronbach's $\alpha = .77$; externalizing behavior Cronbach's $\alpha = .84$).

Procedure

Data were collected in vocational schools using paper questionnaires distributed by the research team. The procedure took roughly 15 minutes and took place in the classroom. After receiving a form outlining informed consent and data protection, participants gave consent verbally, in line with procedures recommended by ethical guidelines. Participants were given a paper packet containing one of the four possible vignettes (masculine girl, feminine girl, masculine boy, or feminine boy) and the questionnaires. After reading the vignette and the instruction to picture the child portrayed in the vignette, participants answered questions about their impressions of the child. First, participants indicated their personal liking of the child, followed by their impression about whether other children and educators would like the child in the vignette. Secondly, participants indicated the perceived level of competence, followed by the perceived creativity and self-esteem of the child. Participants then moved onto the SDQ-Deu, which gathered information about the ascribed internalizing, externalizing and prosocial tendencies of the target child. Finally, participants provided some demographic details about themselves before returning the questionnaires to the research team.

STUDY 2 RESULTS

Analytical Plan

After investigating the data for exclusionary criteria (such as indicating guessing the purpose of the experiment or ticking boxes in a systematic pattern), we excluded 34 cases, leaving 414 participants. Data were analyzed with SPSS 27 and SPSS 29 (IBM Corp, 2020). We checked for normality using a visual inspection of histograms and QQ Plots, showing acceptable normality to use parametric inferential tests.

We calculated a series of two-way ANOVAs in order to investigate the impact of gender stereotype (masculine/feminine) and target gender (male/female) of our vignette on multiple dependent variables. As we investigated a large number of variables and the main effects of gender, gender stereotyped behavior, and the interactions of these factors, we applied corrections for multiple testing. Given that our study is one of the first studies investigating this question in this specific population, in this specific culture, we opted for a more conservative correction to mitigate the risk of type I errors (Anderson, 2008). We thus applied Bonferroni correction to these results, multiplying each of our *p*-values by the initial number of statistical tests carried out (24) (Anderson, 2008, p. 1485). Please note that in these cases, it is common to receive *p*-values that exceed the value of 1 (Anderson, 2008, p. 1485)[1]. Some of our initial results reached significance levels of $p < .001$, in these cases we used a *p*-value of .0009 in our calculations, since a more precise indication of exact *p*-values was not provided by our statistical program. We report the *original p-values* in italics before the corrected *p*-values written in standard format, using the latter, corrected values to interpret our findings.

Liking

Descriptive statistics can be found in Table 2 and correlations of dependent variables are presented in Table 3. In terms of personal liking for the child in the vignette, there were no significant interactions of gender and gender stereotyped behavior ($F(1, 408) = 6.83$, $p =$ *.009*/.216, *pre*/post correction), or significant main effects of gender ($F(1, 408) = 3.79$, $p =$ *.052*/1.248) or gender stereotype ($F(1, 408) = 0.36$, $p =$ *.547*/13.128).

[1] In our analysis, a pre correction *p*-value of .05 would be multiplied by 24 (the number of tests we carried out), yielding a post-correction *p*-value of: 1.2. Ordinarily, a non-corrected *p*-value could only range between 0 and 1, yet the multiplicative correction makes *p*-values exceeding 1 possible and likely (Anderson, 2008, p. 1485). As the post-correction value (1.2) is far above the accepted significance threshold of .05, we would reject this hypothesis. While some of our pre-correction results may seem significant, our post-correction results show non-significance.

The results showed no significant interactions ($F(1, 410) = 2.71$, $p = .101$/2.424) or main effects of gender ($F(1, 410) = 2.55$, $p = .111$/.888) and gender stereotyped behavior ($F(1, 410) = 4.39$, $p = .037$/2.664) for liking of children by other children or faculty (perceived liking by others). These results are not in line with our hypothesis, which stated that we would find significant differences in regard to liking (H1a), especially strong backlash for feminine boys (H1b), and increased liking for girls (H1c).

Competence

Our two-way ANOVA revealed a significant interaction of gender and gender stereotyped behavior for ratings of competence: ($F(1, 409) = 14.67$, $p < .001$/.0216, partial $\eta^2 = .035$). We investigated simple main effects (with Bonferroni correction) using SPSS syntax, revealing that the masculine girl was perceived as significantly more competent than the feminine girl ($F(1, 409) = 16.26$, $p < .001$), and the masculine boy ($F(1, 409) = 7.21$, $p = .008$). The feminine boy was perceived as significantly more competent than the feminine girl ($F(1, 409) = 7.46$, $p = .007$), while the masculine and feminine boys did not differ significantly from one another ($F(1, 409) = 1.93$, $p = .166$). We expected significant differences between children behaving in stereotypical and nonstereotypical ways in regards to ratings of competence (H2) and found partial support for this hypothesis in the case of the masculine girl ($M = 5.55$, $SD = 1.0$), who was perceived as more competent than the feminine girl ($M = 5.00$, $SD = 1.03$); in contrast, gender stereotyped behavior was not related to differences in competence for the feminine ($M = 5.37$, $SD = 1.0$) and masculine boys ($M = 5.18$, $SD = 0.92$).

Table 2

Means, Standard Deviations for Dependent Variables

Item	Masculine Boy Mean (SD)	Feminine Boy Mean (SD)	Masculine Girl Mean (SD)	Feminine Girl Mean (SD)
Perceived liking by others[a]	4.88 (0.84)	4.53 (1.00)	4.87 (0.94)	4.83 (0.96)
Personal liking[a]	4.64 (1.03)	4.98 (1.21)	4.71 (1.09)	4.50 (0.91)
Competence[a]	5.18 (0.92)	5.37 (1.00)	5.55 (1.00)	5.00 (1.03)
Creativity[a]	5.44 (1.35)	5.74 (1.03)	5.94 (1.06)	5.24 (1.12)
Self-esteem[a]	5.63 (1.03)	5.90 (1.10)	6.21 (0.94)	4.78 (1.24)
Prosocial behavior[b]	5.04 (1.83)	7.69 (1.80)	6.25 (2.20)	6.67 (2.00)
Internalizing problems[c]	4.68 (3.26)	7.33 (3.25)	4.94 (3.27)	6.50 (3.05)
Externalizing problems[c]	9.56 (4.03)	4.68 (2.77)	7.79 (4.00)	4.41 (2.65)

Note. [a] Range from 1-7; [b] Range from 0-10; [c] Range from 0-20

Table 3

Correlations for Dependent Variables

Variable	1	2	3	4	5	6	7	8
1. Liking by others	-							
2. Personal liking	.37**	-						
3. Competence	.49**	.35**	-					
4. Internalizing problems	-.34**	-.06	-.32**	-				
5. Externalizing problems	-.05	-.06	-.13*	.04	-			
6. Prosociality	-.05	-.09	-.13**	-.03	.37**	-		
7. Creativity	.26**	.30**	.47**	-.24**	-.15**	-.15**	-	
8. Self-esteem	.18**	.31**	.41**	-.18**	.08	-.01	.49**	-

*$p < .05$; **$p < .01$; ***$p < .001$

Strength and Difficulties Questionnaire

Internalizing and Externalizing Behavior

Our two-way ANOVA revealed that there was no significant interaction between child gender and gender stereotyped behavior ($F(1, 407) = 2.98$, $p = .085$/2.04) for internalizing behavior. Results showed no significant main effect of child gender ($F(1, 407) = 0.80$, $p = .371$/8.904). However, we saw a significant main effect of stereotyped behavior: $F(1, 407) = 44.16$, $p < .001$/.0216, partial $\eta^2 = .098$, showing that the feminine girl ($M = 6.50$, $SD = 3.05$) and boy ($M = 7.33$, $SD = 3.25$) were given greater internalizing scores than the masculine girl ($M = 4.94$, $SD = 3.27$) and boy ($M = 4.68$, $SD = 3.26$). This result supports our hypothesis H3b, which stated that we would find a difference in internalizing behavior, with feminine children ($M = 6.91$, $SD = 3.17$) being rated as having higher scores than masculine children ($M = 4.81$, $SD = 3.26$).

From inferential analyses, we did not have a significant interaction between child gender and gender stereotyped behavior regarding externalizing problems ($F(1, 407) = 4.91$, $p = .027$/.648). While we had no significant main effect of child gender ($F(1, 407) = 9.04$, $p = .003$/.072), we see significant main effects of stereotyped behavior ($F(1, 407) = 148.17$, $p < .001$/.0216, partial $\eta^2 = .267$). The masculine boy ($M = 9.56$, $SD = 4.03$) and girl ($M = 7.79$, $SD = 4.00$) were perceived as having more externalizing tendencies than the feminine boy ($M = 4.68$, $SD = 2.77$) and girl ($M = 4.41$, $SD = 2.65$), lending support to our hypothesis H3c, which stated that masculine children ($M = 8.68$, $SD = 4.10$) would be ascribed greater externalizing problems than feminine children ($M = 4.54$, $SD = 2.71$). As both internalizing and externalizing subscales differed between masculine and feminine children, our hypothesis H3a was also fully supported.

Prosocial Behavior

Analyses revealed a significant interaction between child gender and gender stereotyped behavior, $F(1, 406) = 33.27$, $p < .001/.0216$, partial $\eta^2 = .076$. Simple main effects analyses were conducted and revealed that the feminine boy was perceived as significantly more prosocial than the masculine boy $F(1, 406) = 93.31$, $p < .001$. The masculine boy, in turn, was seen as less prosocial than the masculine girl, $F(1, 406) = 19.86$, $p < .001$. The feminine boy was also seen as more prosocial than the feminine girl, $F(1, 406) = 13.70$, $p < .001$. For girls, gender stereotyped behavior did not make a significant difference in perceived prosocial behavior, $F(1, 406) = 2.28$, $p = .132$. Our hypotheses stated that we expect feminine boys (H4a) and feminine children (H4b) to have favorable ratings on prosociality. We found a significant interaction between gender and gender stereotyped behavior, showing that the feminine boy was perceived as the most prosocial ($M = 7.69$, $SD = 1.80$), more so than the masculine boy ($M = 5.04$, $SD = 1.83$), supporting our hypothesis H4a, but since the masculine ($M = 6.25$, $SD = 2.20$) and feminine girls ($M = 6.67$, $SD = 2.00$) did not differ significantly from one another, hypothesis H4b was not supported.

Creativity

Inferentially, we saw a statistically significant interaction between gender and stereotyped behavior for creativity: $F(1, 408) = 19.35$, $p < .001/.0216$, partial $\eta^2 = .045$. An analysis of the simple main effects showed significant differences between the children. We saw the masculine girl judged as more creative than the feminine girl ($F(1, 408) = 19.09$, $p < .001$). The masculine girl was rated as significantly more creative than the masculine boy ($F(1, 408) = 9.84$, $p = .002$) and the feminine boy as significantly more creative than the feminine girl ($F(1, 408) = 9.52$, $p = .002$). We noted that there was no difference in creativity for boys, whether they were masculine or feminine ($F(1, 408) = 3.43$, $p = .065$). In our hypothesis H5, we stated that violations of gender stereotyped behavior would lead to

children being ascribed greater creativity compared to children who act in accordance with gender stereotypes. We found partial support for this hypothesis, as the masculine girl (M = 5.94, SD = 1.06) was rated as more creative than her feminine counterpart (M = 5.24, SD = 1.12), while there were no significant differences in perceived creativity between the feminine (M = 5.74, SD = 1.03) and masculine boys (M = 5.44, SD = 1.35).

Self-Esteem

The two-way ANOVA revealed a statistically significant interaction of child gender and gender stereotyped behavior, $F(1, 409) = 64.16$, $p < .001$/.0216, partial $\eta^2 = .136$. Simple main effects analyses revealed that the masculine girl was seen as having the highest self-esteem. She was rated as having significantly higher self-esteem than the feminine girl ($F(1, 409) = 91.38$, $p < .001$). Results showed that the feminine boy had significantly higher self-esteem than the feminine girl ($F(1, 409) = 54.91$, $p < .001$). The masculine girl was rated as having significantly higher self-esteem than the masculine boy ($F(1, 409) = 15.12$, $p < .001$). We did not see significant differences in the masculine or feminine boys in terms of self-esteem ($F(1, 409) = 3.17$, $p = .076$). We hypothesized that children not acting in accordance with gender stereotyped behavior for their gender would receive higher ratings on perceived self-esteem and our findings partially supported this hypothesis (H6), as girls who adhered to stereotypical behavior (M = 4.78, SD = 1.24) were perceived as having lower self-esteem than girls who acted in a masculine gender stereotyped manner (M = 6.21, SD = 0.94). The feminine (M = 5.90, SD = 1.10) and masculine boys (M = 5.63, SD = 1.03), however, did not show significant differences from one another in regards to perceived self-esteem.

STUDY 2 DISCUSSION

The aim of this study was to investigate how pedagogical educators in training evaluated vignettes describing gender stereotypical and non-stereotypical three-year-old

children, so as to understand whether there are systematic differences in how children are evaluated based on their gender and gendered behavior.

Our analyses showed no effect of gender or stereotyped behavior on either measure of liking. We saw an impact of gender stereotype and gender for ratings of competence, creativity, and self-esteem, which all showed that the masculine girl received higher ratings, while the feminine girl received the lowest ratings. Gender stereotyped behavior seemed not to make a large difference for boys for these variables. Interestingly, prosocial behavior showed an interaction of gender and gender stereotype, with the feminine boy perceived as most prosocial, while the masculine boy was rated as the least prosocial. Regarding the SDQ-Deu, we saw masculine children perceived as more externalizing, while feminine children were rated as having more internalizing tendencies. Gender stereotyped behavior was a significant factor here, while child gender, or nonconformity, was not. Overall, we see that femininity in boys is not viewed overwhelmingly negatively in this sample. Instead, we see negative outcomes for the feminine girl and positive outcomes for the masculine girl, specifically for ascribed competence, creativity and self-esteem.

GENERAL DISCUSSION

The present work investigated two main questions: which descriptive, prescriptive, and proscriptive stereotypes exist about three-year-old children in German society, and how do pedagogical educators in training evaluate children who conform to or defy these gender stereotypes? We found that adults in Germany identified gender stereotypes for three-year-old boys and girls and that these stereotypes were comparable to past research from the United States (Koenig, 2018; Sullivan et al., 2018, 2022).

In a similar vein to Koenig (2018), the typical characteristics of boys and girls showed more emphasis on activities rather than traits. Compared to girls, whose play activities heavily revolved around humanoid, caring play (dolls, princesses, playing house), boys' activities

featured more "things" (cars, dinosaurs). In terms of highest means, prescriptive stereotypes for girls were related to traits and behaviors, such as being helpful, friendly, and having good manners. These reflect communal values ascribed to women (Hentschel et al., 2019; Hsu et al., 2021). The active, physical nature of prescriptive stereotypes for boys is in line with the beliefs that "sport" is a male domain (Messner, 2011; Plaza et al., 2017). Whereas girls' prescriptions related to their *being*, prescriptions for boys related to what they are *doing*. However, these results should be interpreted cautiously: while the top prescriptions for girls related to having good manners, the data showed that these traits were also highly desirable for boys.

Noticeably, the most proscriptive or "forbidden" characteristics for boys, in terms of lowest means, were appearance related items that signal femininity (wearing dresses and skirts, wearing nail polish, or dancing ballet). The least desirable characteristic for girls related to aggressive behaviors, such as fighting, or being loud. Negative masculinity is partly captured by violent and aggressive tendencies (Krahé et al., 2007), which seem represented in the items deemed proscriptive for girls. Proscriptive items for boys related to feminine physical appearance and clothing. It thus seems that *femininity*, removed from positive and negative labels, was proscriptive for boys. This finding is in line with the theoretical argument stating that prescriptive and proscriptive stereotypes serve to maintain hierarchical systems (Nauts, 2015; Rudman & Glick, 2010). This positive evaluation of masculinity is found in the results of study 2. Masculine girls received the most flattering ratings in terms of competence, creativity, and self-esteem, while feminine boys were not evaluated differently than masculine boys on these measures.

Liking

Prior research has shown that nontypical children were liked less than their typical peers (Sullivan et al., 2018). Despite some acceptance of gender atypicality for "tomboys" and

"mama's boys," gender typical girls were viewed as most likable in past research, with boys showing negative femininity rated as the least likable (Coyle et al., 2016). Contrary to our hypotheses, we found no effect on liking, which perhaps is linked to our sample: pedagogical faculty are discouraged from showing personal preferences. Overall, our present study found no differential treatment between children adhering to or violating gender stereotypes in terms of "liking".

Competence

Previously, typical girls were rated as the most and "sissies" rated as the least competent, while "mama's boys" and "tomboys" were rather positively evaluated as possessing higher competence than the typical boy (Coyle et al., 2016). For adolescents, a gender atypical target was seen as competent, but only by participants with high SES (Meimoun et al. 2023). Our results stand in contrast with some past findings, showing significant impact of gender nonconformity on ratings of competence for girls but not boys. For ratings of competence, masculine and feminine boys were perceived similarly. As masculinity and androgyny have been associated with competence (Heilbrun, 1981; Korlat et al., 2022; Martin & Slepian, 2021), our finding that masculine girls benefitted from greater masculinity is in line with this literature. We believe that the counterstereotypical nature of masculinity in girls may account for the significant difference we see between masculine girls and boys. As masculinity in girls is unexpected, it might therefore be evaluated as more extreme and given more credence than masculinity displayed by a boy (cf. Streck et al., 2022).

Internalizing and Externalizing Problems

Past research showed that rather than nonconformity, it was the gender typing of the child that predicted internalizing and externalizing tendencies (Thomas & Blakemore, 2013). Our current study expanded upon these findings using the established SDQ, which has

previously been used as a measure of such adjustment problems (Goodman et al., 2010). We found a similar pattern as Thomas and Blakemore (2003), showing that masculine children were ascribed greater externalizing problems, while feminine children were ascribed greater internalizing tendencies. We did not find an effect of gender but only of gender stereotyped behavior, showing that it is not necessarily conformity to gender but expressions of gender that relate to perceived problems.

Prosocial Behavior

In our findings on average, feminine children were rated as more prosocial than masculine children. A closer look revealed that the feminine boy was rated as the most prosocial, significantly more so than the masculine boy and feminine girl. Finding feminine boys to be considered more prosocial than masculine boys has also been shown in children previously (Piché & Plante, 1991). Notably, the feminine and masculine girls did not differ significantly from one another in terms of prosociality. There is evidence showing childhood educators believed girls to be more prosocial than boys, despite no observed differences in prosociality between genders (Bouchard et al., 2015). Perhaps the girlhood of the masculine girl activated the "female" stereotype of prosociality (Eisenberg et al., 2006; Quenneville et al., 2022), enough to outweigh her masculine tendencies.

Creativity

Our analyses showed significant interactions of gender and gender stereotyped behavior in terms of perceived creativity. Overall, the masculine girl received beneficial ratings on creativity, while gender stereotyped behavior was not a significant factor for boys. Eccentricity has been linked to greater perceived creativity (Van Tilburg & Igou, 2014), although this would not account for our gender specific findings. The greater creativity ascribed to masculine girls could be seen as an endorsement of the connection between creativity and masculinity (Proudfoot et al., 2015). The masculine girl may have benefitted

from the association between creativity and both atypicality and masculinity, leading to the finding that she was seen as the most creative.

Self-Esteem

In regard to perceived self-esteem of children adhering to or violating gender stereotypes, we found a similar pattern as for creativity and competence, with the greatest self-esteem ascribed to the masculine girl and the lowest self-esteem ascribed to the feminine girl, while gender stereotyped behavior was not a significant factor for boys. People behaving in ways that are seen as outside the norm are perceived as brave innovators, not bound by social convention – their self-esteem is not impacted by the opinions of others. Our results are in line with literature linking actual, not *perceived*, self-esteem to psychological androgyny (Bem, 1974) and masculinity or instrumentality (Antill & Cunningham, 1979; Heilbrun, 1981; Marsh et al., 1987; Whitley & Gridley, 1993), which showed that specifically women can benefit from more masculine orientations (Heilbrun, 1981; Streck et al., 2022; Whitley, 1988). This would also explain why the feminine boy was not seen as having particularly high self-esteem, despite disregarding social norms: his femininity is not conducive to high self-esteem, unlike high masculinity, which is strongly related to positive self-esteem (Antill & Cunningham, 1979; Whitley & Gridley, 1993). We note, however, that perceived self-esteem does not necessarily relate to actual self-esteem (Kilianski, 2008; Watson et al., 2002).

Implications

Past research has revealed a consistent and strong bias against feminine boys, who received the greatest backlash. The focus on the negative aspect of femininity in boys did not emerge in our findings. Interestingly, femininity in our study was only "punished" in girls, while boys displaying the same behavior did not receive backlash in terms of competence, creativity, and self-esteem. Indeed, the feminine boy was perceived as the most prosocial, suggesting that femininity can be an asset for boys but a liability in girls. Our results support

the notion that femininity is seen as "lesser than" masculinity. Gender can be seen as a status characteristic, with masculinity having higher status than femininity (Berger et al., 1972; Feinman, 1981). Denigration of the feminine begins as early as kindergarten, with feminine activities given less time and space than masculine activities (Prioletta & Davies, 2022). Denigrating patterns (Hill & Augoustinos, 1997) continue in school, where feminine achievement has been attributed to effort, rather than ability (Espinoza et al., 2014; Fennema et al., 1990). Whereas younger girls often preferred "pink, frilly dresses" (Halim et al., 2014, p. 1091), girls in elementary school girls reported greater affinity for a tomboy aesthetic (Halim et al., 2011). This effect was theorized to relate to increased understanding of status of gender by girls as they age. From a status perspective it is not surprising that a girl would want to enact masculinity, this, in fact, supports the existing gender hierarchy, as it shows that masculinity is desirable. A boy deigning to engage with femininity, however, threatens the hierarchy by casting high status masculinity aside and undermining its value (Feinman, 1981). This violation is then punished more harshly in order to sustain the hierarchical gender system. While the lack of masculinity in boys was viewed critically in past studies (Feinman, 1981; Sullivan et al., 2018; Thomas & Blakemore, 2013), the presence of it was rewarded for girls in our present results. Greater masculinity in girls could be implicitly encouraged by pedagogical educators in training or in wider society. There is evidence to suggest that women have become increasingly masculinized in the past decades (Twenge, 1997; Wilde & Diekman, 2005), yet not all investigations yield similar patterns (Eagly et al., 2020; Haines et al., 2016), Changing gender stereotypes (Eagly et al., 2020) may reflect a shift in views of masculinity and femininity.

Sanctioning boys who defy gender stereotypes could be seen as a response to a world in which such individuals experience worse outcomes than their stereotypical counterparts (Folkierska-Żukowska et al., 2022; Hu et al., 2023; Issler et al., 2023). Encouraging girls to take on masculine traits could also be interpreted as awareness of the benefits associated with

this gender role orientation. Studies showed that instrumentality can be a beneficial resource for adolescent girls in academic contexts (Streck et al., 2022), and that the association between agency or masculinity and self-esteem was particularly strong for women (Hirokawa & Dohi, 2007; Streck et al., 2022; Whitley, 1988). Of course, by displaying differential reactions to gender stereotyped behavior, even if to attempt to protect boys from poorer future outcomes or encourage beneficial outcomes in girls, socializing agents are perpetuating the power of stereotypes and hierarchy of gender. Encouraging boys and girls to embrace both masculinity and femininity, free from backlash, could have considerable benefits in academic contexts, where androgyny and gender atypical orientations have been linked to higher achievement, self-esteem, and school-related well-being (Korlat et al., 2022; Yavorsky & Buchmann, 2019; Yu et al., 2020) and for constructing a less gender stereotyped society as a whole (Bem, 1981b).

Limitations and Future Research Directions

Ideally, we would have liked to include a measure of social desirability in order to indicate whether participants may have been providing answers in line with norms of social equity. Future studies may wish to add items measuring social desirability bias, a suggestion laid out previously (Sullivan et al., 2018). Collecting further demographic details of participants, such as their SES, could also be advantageous in the future, as past research showed that high SES individuals displayed different result patterns than low SES individuals for attitudes towards gender atypicality (Meimoun et al., 2023). Similarly, investigating whether the gender of the evaluating party has an impact on results, would be interesting. Due to our overwhelmingly female sample, we would not have sufficient power to detect an effect of participant gender. Past research typically has not found this factor to have a significant impact on most variables (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013).

Whether our findings could generalize to children of other ages is uncertain. Past research has shown that multiple stereotypes of 3- and 4-year-old children were also applied to 7-year-old children, supporting the idea that stereotypes can generalize across age groups (Sullivan et al., 2018, 2022). The theoretical link between atypicality and perceived creativity and self-esteem should apply to other age groups.

We would also state that further expansion of this research could include sampling kindergarten teachers who have passed their examinations. Perhaps there is a difference between pedagogical educators in training and kindergarten teachers working full-time. Examining how experienced teachers react to vignettes of nonconformity, or perhaps even instances of actual nonconformity, such as in past research (Fagot, 1977), could benefit the literature and offer a more naturalistic methodology. We chose to work with vignettes, as this enabled a systematic comparison across conditions, while manipulating only gender and gender stereotyped behavior. We used a between-participants design, rather than a within-participants design, as presenting multiple vignettes to participants could have made our experimental manipulation – and thus our hypotheses – obvious to participants and influenced their responses (Charness et al., 2012).

Future studies should investigate whether opinions expressed by pedagogical educators in training are in line with intended behavior towards gender nonconforming children. While we assessed approval and disapproval by measuring constructs like liking and competence, we have no indication of whether these reactions towards gender nonconforming children will translate into behavioral intentions. Italian pre-service teachers claimed they would adopt a supportive, rather than corrective, stance towards children showing gender nonstereotypical behavior (Bochicchio et al., 2019). Combining the behavioral intent from this investigation with the more affective focus of our study would offer interesting insight.

Comparing our results to past research, while insightful, also comes with a major caveat, as we conducted our research in a different culture to most past research and chose a highly specific sample. While studies conducted in the United States (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013) found results of backlash, investigations originating in Italy and France revealed generally more positive attitudes towards gender atypicality from pre-service teachers (Bochicchio et al., 2019) and adolescents (Meimoun et al., 2023). Along with our findings this calls for more cross-cultural research into perceptions of gender atypicality.

We believe this research, informed by social cognitive theory, applies specifically to enactive experience and direct tuition, as these modes of social influence require feedback from a social agent, in this case pedagogical educators in training. The role of adults in the development of gender typed behaviors and attitudes is explored not just in social cognitive theory, but also in gender schema theory (Martin & Halverson, 1981). According to this theory, socializing agents can serve to label certain activities, behaviors, and materials (clothes, toys) as being gender appropriate or inappropriate, which can then inform a child's gender schema, shaping their future cognitions and behaviors. Experimentally labelling novel toys as "for girls" or "for boys" impacted how much children were interested in or enjoyed a particular toy (Weisgram et al., 2014), while gender inappropriate toys were cast aside like a "hot potato" (Martin et al., 1995, p. 1467). Even the use of novel colors as gender labels can impact children, who reported greater liking of gender appropriate colors (Yeung & Wong, 2018). These findings underscore the relevance of cues from adult socializers for children's attitudes and behaviors.

Conclusion

The results outline that gender stereotypes are a present factor in the lives of adults and children. Pedagogical educators in training reported an interesting pattern, showing a

preference for masculinity over femininity in girls – a pattern worth investigating further. More efforts need to be made in order to achieve the goal of gender equity in the German educational system, as outlined by German political actors (Kultusministerkonferenz, 2016) and the wishes of German parents (Wößmann et al., 2018).

REFERENCES

Abele, A. E., & Wojciszke, B. (2013). The big two in social judgment and behavior. *Social Psychology*, *44*(2), 61–62. https://doi.org/10.1027/1864-9335/a000137

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, *103*(484), 1481–1495. https://doi.org/10.1198/016214508000000841

Antill, J. K., & Cunningham, J. D. (1979). Self-esteem as a function of masculinity in both sexes. *Journal of Consulting and Clinical Psychology*, *47*(4), 783–785. https://doi.org/10.1037/0022-006X.47.4.783

Autorengruppe Fachkräftebarometer (2021*). Fachkräftebarometer Frühe Bildung 2021: Weiterbildungsinitiative Frühpädagogische Fachkräfte* [*Early education professionals barometer 2021: Further training initiative for early education professionals*]. https://www.fachkraeftebarometer.de/fileadmin/Redaktion/Publikation_FKB2017/Publikation_FKB2021/WiFF_FKB_2021_web.pdf

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion.* Rand McNally.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155–162.

Bem, S. L. (1981a). *Bem Sex-Role Inventory: A professional manual*. Consulting Psychologist Press.

Bem, S. L. (1981b). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, *88*(4), 354–364. https://doi.org/10.1037/0033-295X.88.4.354

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, *37*(3), 241–255.

Blakemore, J. E. O. (2003). Children's beliefs about violating gender norms: Boys shouldn't look like girls, and girls shouldn't act like boys. *Sex Roles*, *48*(9–10), 411–419. https://doi.org/10.1023/A:1023574427720

Bochicchio, V., Perillo, P., Valenti, A., Chello, F., Amodeo, A. L., Valerio, P., & Scandurra, C. (2019). Pre-service teachers' approaches to gender-nonconforming children in preschool and primary school: Clinical and educational implications. *Journal of Gay and Lesbian Mental Health*, *23*(2), 117–144. https://doi.org/10.1080/19359705.2019.1565791

Bouchard, C., Coutu, S., Bigras, N., Lemay, L., Cantin, G., Bouchard, M. C., & Duval, S. (2015). Perceived, expressed and observed prosociality among four-year-old girls and boys in childcare centres. *Early Child Development and Care*, *185*(1), 44–65. https://doi.org/10.1080/03004430.2014.903940

Braun, S. S., & Davidson, A. J. (2017). Gender (non)conformity in middle childhood: A mixed methods approach to understanding gender-typed behavior, friendship, and peer preference. *Sex Roles*, *77*(1–2), 16–29. https://doi.org/10.1007/s11199-016-0693-z

Bundesagentur für Arbeit. (2023). *Erzieher/in: Ausbildungsberuf [Pedagogical educator: Vocational training]*. https://web.arbeitsagentur.de/berufenet/beruf/9162

Bussey, K. (2011). Gender identity development. In S. Schwartz, K. Luyckx, & V. Vignoles (Eds.), *Handbook of identity theory and research*. (pp. 603-628). Springer. https://doi.org/10.1007/978-1-4419-7988-9

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*(4), 676–713. https://doi.org/10.1037/0033-295X.106.4.676

Cahill, B., & Adams, E. (1997). An exploratory study of early childhood teachers' attitudes toward gender roles. *Sex Roles*, *36*(7–8), 517–529. https://doi.org/10.1007/BF02766688

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior and Organization*, *81*(1), 1–8. https://doi.org/10.1016/j.jebo.2011.08.009

Cook, R. E., Nielson, M. G., Martin, C. L., & DeLay, D. (2019). Early adolescent gender development: The differential effects of felt pressure from parents, peers, and the self. *Journal of Youth and Adolescence*, *48*(10), 1912–1923. https://doi.org/10.1007/s10964-019-01122-y

Coyle, E. F., & Fulcher, M. (2022). Social influences on gender development: Theory and context. In D. P. VanderLaan & W. I. Wong (Eds.), *Gender and sexuality development: Contemporary theory and research* (pp. 101–124). Springer. https://doi.org/10.1007/978-3-030-84273-4_4

Coyle, E. F., Fulcher, M., & Trübutschek, D. (2016). Sissies, mama's boys, and tomboys: Is children's gender nonconformity more acceptable when nonconforming traits are positive? *Archives of Sexual Behavior*, *45*, 1827–1838. https://doi.org/10.1007/s10508-016-0695-5

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, *40*, 61–149.

Dumas, D., & Dunbar, K. N. (2016). The creative stereotype effect. *PLoS ONE*, *11*(2), 1–13. https://doi.org/10.1371/journal.pone.0142567

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U.S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen [Gender stereotypes: Roles, identities and prejudice]. In R. Becker & B. Kortendiek (Eds.), *Handbuch Frauen- und Geschlechterforschung: Theorie, Methoden, Empirie* (pp. 178–189). Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91972-0_20

Egan, S. K., & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*, *37*(4), 451–463. https://doi.org/https://doi.org/10.1037/0012-1649.37.4.451

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 646–718). John Wiley & Sons, Inc. https://doi.org/10.4324/9781315560984-2

Espinoza, P., Arêas da Luz Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: Teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, *17*(1), 105–126. https://doi.org/10.1007/s11218-013-9226-6

Fagot, B. I. (1977). Consequences of moderate cross-gender behavior in preschool children. *Child Development*, *48*(3), 902–907.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Feinman, S. (1981). Why is cross-sex-role behavior more approved for girls than for boys? A status characteristic approach. *Sex Roles*, *7*(3), 289–300. https://doi.org/10.1007/BF00287543

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69. https://doi.org/10.1007/BF00311015

Folkierska-Żukowska, M., Rahman, Q., & Dragan, W. (2022). Childhood gender nonconformity and recalled perceived parental and peer acceptance thereof, internalized homophobia, and psychological well-being outcomes in heterosexual and gay men from Poland. *Archives of Sexual Behavior*, *51*(4), 2199–2212. https://doi.org/10.1007/s10508-021-02245-9

Good, J. J., & Sanchez, D. T. (2010). Doing gender for different reasons: Why gender conformity positively and negatively predicts self-esteem. *Psychology of Women Quarterly*, *34*(2), 203–214. https://doi.org/10.1111/j.1471-6402.2010.01562.x

Goodman, A., Lamping, D. L., & Ploubidis, G. B. (2010). When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the strengths and difficulties questionnaire (SDQ): Data from British parents, teachers and children. *Journal of Abnormal Child Psychology*, *38*(8), 1179–1191. https://doi.org/10.1007/s10802-010-9434-x

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586.

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337–1345. https://doi.org/10.1177/1073191117698757

Haines, E. L., Deaux, K., & Lofaro, N. (2016). The times they are a-changing … or are they not? A comparison of gender stereotypes, 1983–2014. *Psychology of Women Quarterly*, *40*(3), 353–363. https://doi.org/10.1177/0361684316634081

Halim, M. L., Ruble, D. N., & Amodio, D. M. (2011). From pink frilly dresses to „one of the boys": A social-cognitive analysis of gender identity development and gender bias. *Social and Personality Psychology Compass*, *5*(11), 933–949. https://doi.org/10.1111/j.1751-9004.2011.00399.x

Halim, M. L., Ruble, D. N., Tamis-LeMonda, C. S., Zosuls, K. M., Lurye, L. E., & Greulich, F. K. (2014). Pink frilly dresses and the avoidance of all things „girly": Children's appearance rigidity and cognitive theories of gender development. *Developmental Psychology*, *50*(4), 1091–1101. https://doi.org/10.1037/a0034906

Heilbrun, A. B. (1981). Gender differences in the functional linkage between androgyny, social cognition, and competence. *Journal of Personality and Social Psychology*, *41*(6), 1106–1118. https://doi.org/10.1037/0022-3514.41.6.1106

Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The multiple dimensions of gender stereotypes: A current look at men's and women's characterizations of others and themselves. *Frontiers in Psychology*, *10*, 376558. https://doi.org/10.3389/fpsyg.2019.00011

Hill, M. E., & Augoustinos, M. (1997). Re-examining gender bias in achievement attributions. *Australian Journal of Psychology*, *49*(2), 85–90. https://doi.org/10.1080/00049539708259858

Hirokawa, K., & Dohi, I. (2007). Agency and communion related to mental health in Japanese young adults. *Sex Roles*, *56*(7–8), 517–524. https://doi.org/10.1007/s11199-007-9190-8

Hsu, N., Badura, K. L., Newman, D. A., & Speach, M. E. P. (2021). Gender, "masculinity," and "femininity": A meta-analytic review of gender differences in agency and communion. *Psychological Bulletin*, *147*(10), 987–1011. https://doi.org/10.1037/bul0000343

Hu, T., Jin, F., & Deng, H. (2023). Association between gender nonconformity and victimization: A meta-analysis. *Current Psychology*. https://doi.org/10.1007/s12144-023-04269-x

IBM Corp. (2020). *IBM SPSS Statistics for Windows, Version 27.0.* (27). IBM Corp.

Issler, T. C., Ferreira de Sá, D., Michael, T., & Schäfer, S. K. (2023). The relationship between childhood gender nonconformity, aversive childhood experiences, and mental health in heterosexual and non-heterosexual cisgender men: The buffering effect of sense of coherence. *Stress and Health*, *39*(4), 782-797. https://doi.org/10.1002/smi.3227

Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, *20*(3), 309–323. https://doi.org/10.1007/BF03173559

Kilianski, S. E. (2008). Who do you think I think I am? Accuracy in perceptions of others' self-esteem. *Journal of Research in Personality*, *42*(2), 386–398. https://doi.org/10.1016/j.jrp.2007.07.004

Klasen, H., Woerner, W., Wolke, D., Meyer, R., Overmeyer, S., Kaschnitz, W., Rothenberger, A, & Goodman, R. (2000). Comparing the German versions of the strengths and difficulties questionnaire (SDQ-Deu) and the child behavior checklist. *European Child and Adolescent Psychiatry*, *9*(4), 271–276.

Koenig, A. M. (2018). Comparing prescriptive and descriptive gender stereotypes about children, adults, and the elderly. *Frontiers in Psychology*, *9*, 1086. https://doi.org/10.3389/fpsyg.2018.01086

Koglin, U., Barquero, B., Mayer, H., Scheithauer, H., & Petermann, F. (2007). Deutsche Version des Strengths and Difficulties Questionnaire (SDQ-Deu): Psychometrische Qualität der Lehrer-/Erzieherversion für Kindergartenkinder [German version of the strengths and difficulties questionnaire (T4-16 – SDQ)]. *Diagnostica*, *53*(4), 175–183. https://doi.org/10.1026/0012-1924.53.4.175

Korlat, S., Holzer, J., Schultes, M.-T., Buerger, S., Schober, B., Spiel, C., & Kollmayer, M. (2022). Benefits of psychological androgyny in adolescence: The role of gender role self-concept in school-related well-being. *Frontiers in Psychology, 13*, 856758. https://doi.org/10.3389/fpsyg.2022.856758

Krahé, B., Berger, A., & Möller, I. (2007). Entwicklung und Validierung eines Inventars zur Erfassung des Geschlechtsrollen-Selbstkonzepts im Jugendalter [Development and validation of an inventory for measuring gender role self-concept in adolescence]. *Zeitschrift Für Sozialpsychologie*, *38*(3), 195–208. https://doi.org/10.1024/0044-3514.38.3.195

Koordinationsstelle: Chance Quereinstieg/Männer in Kitas (2019). *Gendersensible Ausbildung von Erzieherinnen und Erziehern [Gender-sensitive training of pedagogical educators]*. https://www.bmfsfj.de/resource/blob/176938/95f26041aeec807c02181801a2b4737d/gendersensible-ausbildung-von-erzieherinnen-und-erziehern-handreichung-data.pdf

Kultusministerkonferenz. (2016). *Leitlinien zur Sicherung der Chancengleichheit durch geschlechtersensible schulische Bildung und Erziehung [Guidelines for ensuring equal opportunities through gender-sensitive school education and upbringing]*. https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2016/2016_10_06-Geschlechtersensible-schulische_Bildung.pdf

Kwan, K. M. W., Shi, S. Y., Nabbijohn, A. N., MacMullin, L. N., VanderLaan, D. P., & Wong, W. I. (2020). Children's appraisals of gender nonconformity: Developmental pattern and intervention. *Child Development*, *91*(4), e780–e798. https://doi.org/10.1111/cdev.13316

Liben, L. S., & Bigler, R. S. (2002). The developmental course of gender differentiation: Conceptualizing, measuring, and evaluating constructs and pathways. *Monographs of the Society for Research in Child Development*, *67*(2), 1–146.

Marsh, H. W., Antill, J. K., & Cunningham, J. D. (1987). Masculinity, femininity, and androgyny: Relations to self-esteem and social desirability. *Journal of Personality*, *55*(4), 661–685. https://doi.org/10.1111/j.1467-6494.1987.tb00457.x

Martin, A. E., & Slepian, M. L. (2021). The primacy of gender: Gendered cognition underlies the big two dimensions of social cognition. *Perspectives on Psychological Science*, *16*(6), 1143–1158. https://doi.org/10.1177/1745691620904961

Martin, C. L. (1990). Attitudes and expectations about children with nontraditional and traditional gender roles. *Sex Roles*, *22*(3–4), 151–166. https://doi.org/10.1007/BF00288188

Martin, C. L. (1995). Stereotypes about children with traditional and nontraditional gender roles. *Sex Roles*, *33*(11–12), 727–751. https://doi.org/10.1007/BF01544776

Martin, C. L., Eisenbud, L., & Rose, H. (1995). Children's gender-based reasoning about toys. *Child Development*, *66*(5), 1453–1471. https://doi.org/10.1111/j.1467-8624.1995.tb00945.x

Martin, C. L., & Halverson, C. F. (1981). A schematic processing model of sex typing and stereotyping in children. *Child Development*, *52*(4), 1119–1134.

Meimoun, E., Bonnot, V., Berenguer, J., & Aelenei, C. (2023). How are gender counter-stereotypical adolescents evaluated by their peers? Investigating the backlash effect in French schools. *Social Psychology of Education*. https://doi.org/10.1007/s11218-023-09807-w

Messner, M. (2011). Gender ideologies, youth sports, and the production of soft essentialism. *Sociology of Sport Journal*, *28*(2), 151–170. https://doi.org/10.1123/ssj.28.2.151

Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules: Status incongruity backlash against modest men. *Psychology of Men & Masculinity*, *11*(2), 140ß151. https://doi.org/10.1037/a0018093

Myers, D., Abell, J., Kolstad, A., & Sani, F. (2010). *Social psychology*. McGraw Hill.

Nauts, S. (2015). *Backlash for gender atypicality*. Radboud University Nijmegen.

Nett, T., Dorrough, A., Jekel, M., & Glöckner, A. (2020). Perceived biological and social characteristics of a representative set of German first names. *Social Psychology*, *51*(1), 17–34. https://doi.org/10.1027/1864-9335/a000383

Petermann, U., Petermann, F., & Schreyer, I. (2010). The German Strengths and Difficulties Questionnaire (SDQ) validity of the teacher version for preschoolers. *European Journal of Psychological Assessment*, *26*(4), 256–262. https://doi.org/10.1027/1015-5759/a000034

Piché, C., & Plante, C. (1991). Perceived masculinity, femininity and androgyny among primary school boys. *European Journal of Psychology of Education*, *6*(4), 423–435. https://doi.org/10.1007/BF03172775

Plaza, M., Boiché, J., Brunel, L., & Ruchaud, F. (2017). Sport = male… But not all sports: Investigating the gender stereotypes of sport activities at the explicit and implicit levels. *Sex Roles*, *76*(3–4), 202–217. https://doi.org/10.1007/s11199-016-0650-x

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, *26*(4), 269–281. https://doi.org/10.1111/1471-6402.t01-1-00066

Prioletta, J., & Davies, A. W. (2022). Femmephobia in kindergarten education: Play environments as key sites for the early devaluation of femininity and care. *Contemporary Issues in Early Childhood*, *0*(0), 1–15. https://doi.org/10.1177/14639491221137902

Proudfoot, D., Kay, A. C., & Koval, C. Z. (2015). A gender bias in the attribution of creativity: Archival and experimental evidence for the perceived association between masculinity and creative thinking. *Psychological Science*, *26*(11), 1751–1761. https://doi.org/10.1177/0956797615598739

Quenneville, S., Talwar, V., & Bosacki, S. (2022). Teacher ratings and adolescent students' perceived social behaviours and gender-role orientations. *Journal of Gender Studies*, *31*(4), 444–456. https://doi.org/10.1080/09589236.2021.1988530

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, *57*(4), 743–762. https://doi.org/10.1111/0022-4537.00239

Rudman, L. A., & Glick, P. (2010). *The social psychology of gender: How power and intimacy shape gender relations*. Guilford Press.

Rudman, L. A., Moss-Racusin, C. A., Glick, P., & Phelan, J. E. (2012). Reactions to vanguards: Advances in backlash theory. In P. Devine & A. Plant (Eds.), *Advances in experimental social psychology*, (1st ed., Vol. 45, pp. 167–227). Elsevier Inc. https://doi.org/10.1016/B978-0-12-394286-9.00004-4

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Sanborn-Overby, U. A., & Powlishta, K. K. (2020). Differential reactions to male and female gender-role violations: Testing the sexual orientation hypothesis. *Archives of Sexual Behavior*, *49*(7), 2531–2546. https://doi.org/10.1007/s10508-020-01803-x

Sandnabba, N. K., & Alberg, C. A. (1999). Parents' attitudes and expectations about children's cross-gender behavior. *Sex Roles*, *40*, 249–263.

Sczesny, S., Nater, C., & Eagly, A. H. (2019). Agency and communion: Their implications for gender stereotypes and gender identities. In A. E. Abele & B. Wojciszke (Eds.), *Agency and communion in social psychology* (pp. 103–116). Routledge. https://doi.org/10.4324/9780203703663

Skinner, O. D., Kurtz-Costes, B., Wood, D., & Rowley, S. J. (2018). Gender typicality, felt pressure for gender conformity, racial centrality, and self-esteem in African American adolescents. *Journal of Black Psychology*, *44*(3), 195–218. https://doi.org/10.1177/0095798418764244

Skoe, E. E. A., Cumberland, A., Eisenberg, N., Hansen, K., & Perry, J. (2002). The influences of sex and gender-role identity on moral cognition and prosocial personality traits. *Sex Roles*, *46*(9–10), 295–309. https://doi.org/10.1023/A:1020224512888

Spence, J. T., Helmreich, R. L., & Stapp, J. (1974). *The personal attributes questionnaire: A measure of sex role stereotypes and masculinity-femininity*. University of Texas.

Streck, H., Nishen, A. K., & Kessels, U. (2022). Instrumentality gives girls the edge: Gender-differential relations between instrumentality, achievement motivation, and self-esteem. *Sex Roles*, *86*(5–6), 379–394. https://doi.org/10.1007/s11199-021-01270-1

Sullivan, J., Ciociolo, A., & Moss-Racusin, C. A. (2022). Establishing the content of gender stereotypes across development. *PLoS ONE*, *17(7),* e0263217. https://doi.org/10.1371/journal.pone.0263217

Sullivan, J., Moss-Racusin, C., Lopez, M., & Williams, K. (2018). Backlash against gender stereotype-violating preschool children. *PLoS ONE*, *13*(4), 1–24. https://doi.org/10.1371/journal.pone.0195503

Thomas, R. N., & Blakemore, J. E. O. (2013). Adults' attitudes about gender nonconformity in childhood. *Archives of Sexual Behavior*, *42*(3), 399–412. https://doi.org/10.1007/s10508-012-0023-7

Tobin, D. D., Menon, M., Menon, M., Spatta, B. C., Hodges, E. V., & Perry, D. G. (2010). The intrapsychics of gender: A model of self-socialization. *Psychological Review*, *117*(2), 601–622.

Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles*, *36*, 305–325.

Van Tilburg, W. A. P., & Igou, E. R. (2014). From Van Gogh to Lady Gaga: Artist eccentricity increases perceived artistic skill and art appreciation. *European Journal of Social Psychology*, *44*(2), 93–103. https://doi.org/10.1002/ejsp.1999

Watson, D., Suls, J., & Haig, J. (2002). Global self-esteem in relation to structural models of personality and affectivity. *Journal of Personality and Social Psychology*, *83*(1), 185–197. https://doi.org/10.1037/0022-3514.83.1.185

Weisberg, R. (2010). The study of creativity: From genius to cognitive science. *International Journal of Cultural Policy*, *16*(3), 235–253. https://doi.org/10.1080/10286630903111639

Weisgram, E. S., Fulcher, M., & Dinella, L. M. (2014). Pink gives girls permission: Exploring the roles of explicit gender labels and gender-typed colors on preschool children's toy preferences. *Journal of Applied Developmental Psychology*, *35*(5), 401–409. Elsevier Inc. https://doi.org/10.1016/j.appdev.2014.06.004

Whitley, B. E. (1988). Masculinity, femininity, and self-esteem: A multitrait-multimethod analysis. *Sex Roles*, *18*(7–8), 419–431. https://doi.org/10.1007/BF00288393

Whitley, B. E., & Gridley, B. E. (1993). Sex role orientation, self-esteem, and depression: A latent variables analysis. *Personality and Social Psychology Bulletin*, *19*(4), 363–369. https://doi.org/10.1177/0146167293194001

Wilde, A., & Diekman, A. B. (2005). Cross-cultural similarities and differences in dynamic stereotypes: A comparison between Germany and the United States. *Psychology of Women Quarterly*, *29*(2), 188–196. https://doi.org/10.1111/j.1471-6402.2005.00181.x

Wößmann, L., Lergetporer, P., Grewenig, E., Kersten, S., & Werner, K. (2018). Was denken die Deutschen zu Geschlechterthemen und Gleichstellung in der Bildung? – Ergebnisse des ifo Bildungsbarometers 2018 [What do Germans think about gender issues and equality in education? – Results of the ifo Education Survey 2018]. *Ifo Schnelldienst*, *71*(17), 15–30.

Xiao, S. X., Cook, R. E., Martin, C. L., & Nielson, M. G. (2019). Characteristics of preschool gender enforcers and peers who associate with them. *Sex Roles*, *81*(11–12), 671–685. https://doi.org/10.1007/s11199-019-01026-y

Yavorsky, J. E., & Buchmann, C. (2019). Gender typicality and academic achievement among American high school students. *Sociological Science*, *6*, 661–683. https://doi.org/10.15195/v6.a25

Yeung, S. P., & Wong, W. I. (2018). Gender labels on gender-neutral colors: Do they affect children's color preferences and play performance? *Sex Roles*, *79*(5–6), 260–272. https://doi.org/10.1007/s11199-017-0875-3

Young, R., & Sweeting, H. (2004). Adolescent bullying, relationships, psychological well-being, and gender-atypical behavior: A gender diagnosticity approach. *Sex Roles*, *50*(7–8), 525–537. https://doi.org/10.1023/B

Yu, J., McLellan, R., & Winter, L. (2020). Which boys and which girls are falling behind? Linking adolescents' gender role profiles to motivation, engagement, and achievement. *Journal of Youth and Adolescence*, 1–17. https://doi.org/10.1007/s10964-020-01293-z

Zeigler-Hill, V., Besser, A., Myers, E. M., Southard, A. C., & Malkin, M. L. (2013). The status-signaling property of self-esteem: The role of self-reported self-esteem and perceived self-esteem in personality judgments. *Journal of Personality*, *81*(2), 209–220. https://doi.org/10.1111/j.1467-6494.2012.00790.x

Zeigler-Hill, V., & Myers, E. M. (2012). A review of gender differences in self-esteem. In S. P. McGeown (Ed.), *Psychology of gender differences* (pp. 131–143). Nova.

Zentner, M., & Von Aufsess, C. (2022). Is being gender nonconforming distressing? It depends where you live: Gender equality across 15 nations predicts how much gender nonconformity is related to self-esteem. *Psychological Medicine*, *52*(10), 1857–1865. https://doi.org/10.1017/S0033291720003645

Zucker, K. J. (1977). Sex-typing of children's toys by college students and fourth-graders. *Catalogue of Selected Documents in Psychology*, *7*, 6–7.

Zucker, K. J., Wilson-Smith, D. N., Kurita, J. A., & Stern, A. (1995). Children's appraisals of sex-typed behavior in their peers. *Sex Roles*, *33*(11), 703–725.

Appendix

English Vignettes (translated from German)

Introduction

In the following you will read a short story set in a Kindergarten.

Try to put yourself into the situation and picture the child from the story. Afterwards we will ask you to please answer some questions about the child.

Today was Natalie's first day as a pre-service educator in the Kindergarten where she previously completed an internship. In the kitchen she runs into a Kindergarten teacher, who has been working at the Kindergarten for a while.

Feminine Girl

Kindergarten teacher: Oh, hello Natalie! How was everything on your first day? Did everything go well with the children?

Natalie: Good, thanks! Overall, it went very well! I already know a lot of the kids from my internship here, and they know me too. The only one I didn't know was Linda, one of the three-year-olds. What is she like? Today on the playground, she played with dolls the whole time. And then she wanted to play in the toy kitchen. When we were drawing later, she drew a dress.

Kindergarten teacher: Yes, that is very typical. Linda also likes princesses and likes to dance ballet.

Feminine Boy

Kindergarten teacher: Oh, hello Natalie! How was everything on your first day? Did everything go well with the children?

Natalie: Good, thanks! Overall, it went very well! I already know a lot of the kids from my internship here, and they know me too. The only one I didn't know was Dominik, one of the three-year-olds. What is he like? Today on the playground, he played with dolls the whole time. And then he wanted to play in the toy kitchen. When we were drawing later, he drew a dress.

Kindergarten teacher: Yes, that is very typical. Dominik also likes princesses and likes to dance ballet.

Masculine Girl

Kindergarten teacher: Oh, hello Natalie! How was everything on your first day? Did everything go well with the children?

Natalie: Good, thanks! Overall, it went very well! I already know a lot of the kids from my internship here, and they know me too. The only one I didn't know was Linda, one of the three-year-olds. What is she like? Today on the playground, she played with a wooden sword the whole time. And then she wanted to play 'cops and robbers.' When we were drawing later, she drew a rocket ship.

Kindergarten teacher: Yes, that is very typical. Linda also plays with cars a lot and likes showing her soccer-cards.

Masculine Boy

Kindergarten teacher: Oh, hello Natalie! How was everything on your first day? Did everything go well with the children?

Natalie: Good, thanks! Overall, it went very well! I already know a lot of the kids from my internship here, and they know me too. The only one I didn't know was Dominik, one of the three-year-olds. What is he like? Today on the playground, he played with a wooden sword the whole time. And then he wanted to play 'cops and robbers.' When we were drawing later, he drew a rocket ship.

Kindergarten teacher: Yes, that is very typical. Dominik also plays with cars a lot and likes showing his soccer-cards.

Manuscript 2

Nice, but not smart?

Attributional backlash from displaying prosocial behavior in the classroom

Streck, H., & Kessels, U. (2023). Nice, but not smart? Attributional backlash from displaying prosocial behavior in the classroom. *Social Psychology of Education*, *26*(6), 1621-1649. https://doi.org/10.1007/s11218-023-09811-0

Note: This is the first-author version of the paper submitted to Social Psychology of Education. Formatting that diverges from APA 7 follows journal-specific guidelines.

Abstract

According to attributional theory, when the application of effort leads to success we praise the achievement. Effort and ability, however, are seen as compensatory and thus, paradoxically, being praised can lead to attributions of low ability. Our study investigates whether praise, not for academic performance, but for social classroom behavior, would also incur attributional backlash. We examined whether prosociality relates to attributions of high effort and low ability, mediated by expected teacher praise and happiness. In adolescence, prosocial behavior is displayed more by females and aligns with femininity. We conducted an experimental vignette study with 324 German ninth graders to examine whether prosocial students experience a denigration of achievement via expected teacher reaction. Multilevel modelling showed that compared to nondescript students, prosocial students were judged to receive good grades as a result of effort and less due to ability, but this was not related to expected teacher reactions. Prosocial students were also judged to be more likeable and popular. Examination of gender-related outcomes showed that prosocial students were believed to be more feminine, but also more masculine than the nondescript student. Female prosocial targets were thought to be more typical, but not as occurring more frequently than their male counterparts. The results are discussed in reference to the paradox of praise. The limitations and implications of the research are discussed, particularly regarding female students' achievements.

***Keywords**:* academic achievement, attribution, gender, paradoxical effects of praise, prosocial behavior

1 Introduction

Attribution theory states that the causes we ascribe to outcomes have a significant impact on how we interpret and judge said outcomes (Heider, 1958). Specifically, teachers have reacted differently depending on whether a student's performance can be attributed to ability or effort (Butler, 1994; Prawat et al., 1983). Research has shown that students are praised when their performance is perceived to be a result of high effort and not high ability (Covington & Omelich, 1979; Weiner & Kukla, 1970). While it is a commonly held belief that praise is a positive experience for students (see Brummelman & Dweck, 2020), many studies have shown that praise in school can incur a negative side effect, known as the "paradox of praise" (see Graham & Chen, 2020; Meyer, 1982; Meyer et al., 1979; Möller, 2005). This paradox occurs as a result of our implicit belief that effort and ability are compensatory (Heider, 1958; Möller, 2005): someone with higher ability needs less effort to solve a task than someone with low ability and vice versa. Due to this "principle of compensation", praising can signify to others that the achieved outcome is due to effort and not ability (Binser & Försterling, 2004; Meyer et al., 1979; Möller, 2005). This essentially reduces the students' achievement to mere effort and diminishes ascriptions of ability, thus praised students may be seen as less able (Meyer et al., 1979; Möller 2005). The paradoxical meanings of teacher praise can be interpreted by students (Miller & Hom, 1996; Möller, 2005), especially by students in adolescence (Barker & Graham, 1987; Möller, 2005). Importantly, it is not just reactions to *academic behavior* that can elicit such attributions: prior research has revealed that *social behavior* in the classroom can evoke ascriptions of ability or effort, via expected reactions by teachers (Kessels & Heyder, 2020). This previous research focused on negative classroom behavior and expected reprimands, which lead students to attribute a fictional disruptive student's low grade to low effort, rather than low ability. Rather than a negative classroom behavior, the present paper investigates positive, prosocial behavior

in the classroom from an attributional perspective, aiming to illuminate whether displaying prosocial behavior relates to the denigration of achievement via expected praise.

In our introduction, we will give an overview on attributional theory and the special relation between effort and ability. We further outline praise and the norm of effort in school, showing the paradoxical effect praise has on attributions. We will ask whether these paradoxical effects of praise will generalize to praise received for prosocial behavior at school. We conclude by examining how prosocial behavior is perceived by others. Overall, the present study investigates whether prosocial behavior leads to ascriptions of high effort and low ability via expected teacher reactions, specifically praise, and how prosocial students are perceived in terms of social and gender-related outcomes.

2 How Does Classroom Behavior Lead to Effort and Ability Attributions?

The causes we ascribe to certain outcomes have an impact on public and private reactions to said outcomes (Heider, 1958; Weiner, 1986). Heider's attribution theory (1958), later expanded upon by Weiner (1986, 2000), forms the basis for investigations into the causes people believe to underlie our actions. Task difficulty, ability, effort (Heider, 1958), and luck (Weiner, 1986) are presented as key determinants of success or failure. Weiner (1986) proposed that the perceived *locus*, *controllability,* and *stability* of causes account for how causal attributions are made. Ability, an internal, stable, but not controllable attribute, is contrasted with effort, which is characterised as internal, flexible, and controllable. Importantly, effort and ability are believed to be compensatory: someone with high ability may solve easy tasks while exerting very little effort. An individual with low ability however, would need to compensate for this by applying more effort to solve the task (Meyer, 1992; Möller, 2005; Weiner & Kukla, 1970).

Applying attributional theory to classroom contexts has a rich history (see Wang & Hall, 2018) and investigations have shown the relevance of attributional theory from the perspective of teachers and students. Butler (1994) asked teachers to react to students described as showing low effort or low ability and found distinct pattern in responses in line with attributional theory. More significantly, these teacher reactions were shown to students, who interpreted the teacher reactions as cues of ability or effort (Butler, 1994). Thus, not only do teachers enact specific behavior in line with their attributions, such as pity, reprimands, or praise, but students receiving these cues interpret how teachers attributed their performance. Since these cues are public, it is also possible for observing peers of these students to make attributional inferences about their fellow classmates. It follows that behavior in classrooms can trigger certain teacher reactions, which are interpreted by students as indicators of teachers' effort or ability attributions (Butler, 1994; Kessels & Heyder, 2020; Meyer et al., 1979; Miller & Hom, 1996; Möller, 2005).

Importantly, it is not just on-task, academic behavior that can incur these teacher reactions and subsequent student interpretations: social behavior in the classroom can also trigger attributions. Prior research has revealed that disrupting the class prompted students to expect corresponding teacher emotions and reactions, which then had an impact on attributions of effort and ability (Kessels & Heyder, 2020). This experimental vignette study has shown that disruptive behaviors lead to expectations of teacher reprimands, which lead to ascriptions of low effort, rather than low ability, in cases of academic failure (Kessels & Heyder, 2020). Students described as equally low performing, but behaving unobtrusively, were not seen as evoking reprimands to the same extent, therefore their low performance was also less strongly attributed to mere lack of effort (Kessels & Heyder, 2020). We believe that this study could be mirrored with desirable, instead of disruptive, classroom behavior. While Kessels and Heyder (2020) revealed how low performing students can reap attributional

benefits from displaying anti-social behavior, we focus on the possible attributional backlash high performing students might risk when behaving prosocially. As their prosocial behavior will likely be praised, this might impact what fellow students consider to be the causes of their good grades, resulting in negative inferences about prosocial students' abilities.

3 Praise in the Classroom

We can understand "praise" in a school context as an expression of "approval or admiration" (Brophy, 1981, p. 5). Praise goes beyond simple feedback and is said to include a display of teacher emotion, such as pride, or delight (Brophy, 1981). Most instances of praise are seen in elementary school and then decrease in later grades (Jenkins et al., 2015), although adolescents also appreciated and responded well to praise both in familial (Padilla-Walker & Carlo, 2004) and academic contexts (Fefer et al., 2016; Hallinan, 2008).

3.1 Displaying Effort Leads to Praise From Teachers

Student accomplishments are accompanied by corresponding emotions and feelings on the part of the teacher, depending on their attributions (Butler, 1994; Hareli & Weiner, 2002; Weiner, 2000). While perceived lack of effort elicits teachers' anger (Butler, 1994; Kessels & Heyder, 2020), displays of effort trigger feelings of happiness (Prawat et al., 1983; Weiner, 2007). At school, a primary reason for receiving praise is the application of effort. Effort is praiseworthy as it is a pathway to success perceived to be under the actors' control. According to Weiner (1995) our positive associations with effort are deeply ingrained in many of our cultures and institutions, which perceive diligence, hard work, and effort as inherently moral features. It has thus become a norm in our institutions, including school, where effort is expected (Matteucci & Gosling, 2004) and appreciated (Matteucci et al., 2008; H. Wang & Hall, 2018) by teachers who expressed greater liking for effortful students (Saidah et al., 2019). Effort has been proposed as "an implicit rule of conduct" (Matteucci & Gosling, 2004, p. 162) in school. Teachers ranked effort and diligence, rather than competence, as more important for attaining success in school and expected effortful students to be better adjusted

towards the demands of school (Matteucci & Gosling, 2004), despite not displaying high ability. Overall, teachers were more likely to report valuing effort over ability (Dompnier & Pansu, 2010; Matteucci et al., 2008; Saidah et al., 2019) and displaying a lack of effort in school was penalized more harshly than displaying a lack of ability (H. Wang & Hall, 2018; Weiner & Kukla, 1970) for all students. Many studies show that praise serves as an indicator for effort, as successful students who achieved their grades as a result of effort are met with teacher praise (Meyer et al., 1979; Miller & Hom, 1996; Möller, 2005), whereas success due to high ability results in fewer instances of praise (Matteucci et al., 2008; Meyer et al., 1979; Möller, 2005; Saidah et al., 2019). Across multiple experimental vignettes studies by Weiner and Kukla (1970) that varied high versus low effort and ability of target students, participants evaluated students described as exerting effort more favourably. Students presented as low in ability, but high in effort received the most rewards from participants. These and other studies (Covington & Omelich, 1979) highlight that praise follows displays of effort, rather than ability.

3.2 Praise from Teachers Leads to Attributional Backlash: High Effort and Low Ability

While on the surface praise may seem like a desirable social exchange (Padilla-Walker & Carlo, 2004), research dating to the late 1970s has examined how praise can be affiliated with negative outcomes because of the associated attributions (Meyer, 1982; Meyer et al., 1979). As instructors are more likely to praise for effort than ability (Covington & Omelich, 1979; Weiner & Kukla, 1970), distributing praise signifies that a good performance is a result of high-effort. Given the compensatory nature between effort and ability (Heider, 1958; Möller, 2005; Weiner, 1986; Weiner & Kukla, 1970), praised individuals are thus judged to be lower in ability (Meyer et al., 1979). Investigating praise from an attributional perspective has yielded much experimental evidence for this ‘paradox of praise’, indicating that participants ascribe higher effort, but lower ability to those students who receive praise instead of neutral feedback (for a review, see Graham & Chen, 2020; Meyer et al., 1979;

Möller, 2005). As outlined above, students are also aware of attributions of ability and effort in the classroom and they interpret teacher communications, such as direct feedback or displays of emotions, accordingly (Butler, 1994; Kessels & Heyder, 2020; Möller, 2005). These findings show that teacher reactions and behaviors in the classroom can have an impact on how students attribute their peers' performances.

Overall it may be said that praise is typically given to displays of effort for academic achievements. We are, however, interested in praise distributed for non-academic behavior, and whether this also has a paradoxical effect on attributions for academic outcomes. Non-academic classroom behavior for which a student may receive praise would be supportive, helpful behavior, benefitting other students (Haydon et al., 2020; Ramaswamy & Bergin, 2009). As such, teachers' reactions to a prosocial student should resemble their reactions to a student whose performance they attribute to high effort. We have thus chosen to investigate whether prosocial behavior will elicit an expectation of teacher approval, such as happiness and praise, which should then induce an attributional pattern favouring effort over ability in peers.

4 Prosociality Leads to Ascriptions of Popularity and Femininity

In our social world, we often see people displaying behavior that benefits others (Pfattheicher et al., 2022), while incurring a cost for themselves. In adults, such prosocial acts may be chivalrous or entail caring for others (Eagly, 2009). In children, such acts can encompass sharing and comforting others (Gerbino et al., 2018). Research shows that prosocial behavior has multiple consequences, ranging from being praised to associations with characteristics and psychological outcomes, such as popularity, likeability, and femininity. In addition to the attributional mechanism outlined above, our study investigates how prosocial peers are perceived by students.

Displaying prosocial behavior is received well by others and is associated with beneficial psychological outcomes for the individual. In a longitudinal study of adolescents, prosocial behavior predicted higher likeability and popularity (Lu et al., 2018). Children and adolescents engaging in prosocial behavior were rated as more popular (Kornbluh & Neal, 2016). Prosocial behavior was associated with better quality friendships in adolescent boys (Son & Padilla-Walker, 2020) and in general was related to better relationships to peers (Lai et al., 2020; M. Wang et al., 2019).

Researchers have claimed that "major psychological causes of prosocial behaviors include warmth-related emotion (e.g., empathy; Batson, 2011)" (Kawamura et al., 2021, p. 453), with experimental evidence showing prosocial actors being ascribed higher warmth (Kawamura et al., 2021; Klein et al., 2014). Warmth is characterized by a kind and friendly disposition directed towards others (Cuddy et al., 2008) and as prosocial behavior is defined as acts benefitting an 'other', the close relation between the two constructs is unsurprising (Eagly, 2009). Warmth has not only been associated with prosociality, but is also closely tied to the construct of femininity (Martin & Slepian, 2021), a pattern consistent across cultures (Asbrock, 2010; Bosson et al., 2022; Ebert et al., 2014).

Both men and women and boys and girls display prosocial behavior (Atkinson et al., 2021; Eagly, 2009), although specific helping behaviors can differ between genders (Becker & Eagly, 2004; Diekman & Clark, 2015; Hine, 2017). Studies found girls scoring higher than boys on different measures of prosociality (Bøe et al., 2016; Gerbino et al., 2018; Koglin et al., 2007; Lohbeck et al., 2015; see Silke et al., 2018; Van der Graaff et al., 2018; Xiao et al., 2019; Zakriski et al., 2005; Zimmer-Gembeck et al., 2005). Given the connection between prosociality and warmth (Kawamura et al., 2021) and warmth and femininity (Martin & Slepian, 2021) we also find evidence of associations between prosocial behavior and femininity (Eisenberg et al., 2001; Quenneville et al., 2022), although such associations may

depend on the type of prosociality. Agentic helping behaviors, such as being chivalrous or taking physical risks (Becker & Eagly, 2004; Hine, 2017), are more often associated with men, while communion-oriented behaviors, such as being nurturing and providing emotional care, are seen as more feminine (Diekman & Clark, 2015; Eagly, 2009).

Both prosocial and antisocial behavior in school are strongly stereotyped as masculine and feminine, respectively. Antisocial behavior in school is stereotypically seen as masculine (Glock & Kleen, 2017; Heyder et al., 2021) and has been shown to have an associated beneficial attributional pattern of low effort, rather than low ability (Kessels & Heyder, 2020). Whether prosocial behavior, which is seen as feminine (Eisenberg et al., 2001; Quenneville et al., 2022), might have a negative pattern of associated attributions is the focus of the present research. The role of stereotypes on attributional judgments has been previously laid out, revealing the impact that stereotyped identities have on evaluations of performance (Reyna, 2000). The "translation" of teacher reactions into attributional causes can be affected by stereotypes (Reyna, 2000): whether a teachers' reaction to a student contains cues of effort or ability can relate to whether the student in question is stereotyped by the teacher. These reactions impart attributional beliefs to the student, signalling whether they are considered able or effortful, in line with social stereotypes (Reyna, 2000). Our present research relates to this attributional model of stereotypes, but focuses less on stereotyped identities, but rather on behaviors that are stereotypically associated with members of a particular group. We believe students showing feminine behaviors, such as prosociality, are more often met with positive reactions by teachers (Younger et al., 1999). Instances of expected teacher happiness and praise will, in turn, lead fellow students to attribute good performances of these prosocial students to high effort, rather than high ability.

This pattern has been observed in judgments of female achievement, as teachers reported girls' successful performance in mathematics were due to high effort, whereas boys'

maths achievement were due to high ability (Espinoza et al., 2014), and female students' failure was more likely to be attributed to a lack of ability (Tiedemann, 2000). The ascription of effort and ability thus has special relevance for the denigration of female achievement. The present study investigates if the display of prosocial behavior even contributes to unfavourable attributional patterns of high performance at school.

5 Study Overview

The present vignette study will investigate whether well-performing target students described as displaying prosocial behavior are expected to be met with happiness and praise by teachers, and if this expected teacher reaction will lead students to attribute these good performances to high effort (and not high ability). In other words, will praise for prosocial behavior lead to ascriptions of lower ability in classrooms? Mirroring the underlying theory of Kessels' and Heyder's (2020) paper on attributional inferences made for a low-achieving student behaving anti-socially who is reprimanded by their teachers, we will test whether displaying prosocial behavior has adverse effects on the attributions made for the academic performance of good students. Just as Kessels and Heyder, we will look at the attributions made by students who were presented with a vignette description of a fictional peer student behaving in a specific way. We argue that prosocial behavior displayed by students will be expected to be judged positively by teachers and peers (Birch & Ladd, 1998; Lai et al., 2020; M. Wang et al., 2019; Wentzel & Asher, 1995). This positive judgment should result in the expectation of happiness experienced by the teacher and instances of praise by the teacher (Jenkins et al., 2015). We thus hypothesize that students showing prosocial behavior in a vignette will be perceived as eliciting teacher happiness and praise (H1). As praise follows displays of high effort, rather than ability (Covington & Omelich, 1979; Weiner & Kukla, 1970), we expect the prosocial student (compared to a nondescript vignette counterpart) will be ascribed greater effort for their good academic performance (H2). Research has shown that teacher reactions (Butler, 1994; Möller, 2005) can impact ascriptions of effort and ability

made by peers. Accordingly, we expect that the hypothesized effort-attribution will be mediated by expected teacher happiness and praise (H2a). As putting effort into a task is seen as compensatory for having low ability in that task (Binser & Försterling, 2004; Meyer et al., 1979; Miller & Hom, 1996; Möller, 2005), the prosocial student should also be ascribed lower academic ability (H3). We believe that this lack-of-ability attribution will similarly be mediated by expected teacher reactions of happiness and praise (H3a). Relating to displaying higher effort, but lower ability, prosocial students are expected to be perceived as possessing lower intelligence (H4a), a sign of lower ability.

Given the connotations of prosociality on a theoretical and empirical level, students described as more prosocial should also be ascribed characteristics relating to popularity and likeability. Prosocial students should thus be evaluated as more popular and as having more friends (H5a) and be more well-liked and more desirable to have as a friend (H5b), which is a consistent finding in the literature (Kornbluh & Neal, 2016; Lu et al., 2018).

The close ties of prosociality to warmth (Kawamura et al., 2021), which is associated with femininity (Eagly, 2009) and the finding that girls tend to display more prosocial behavior (Bøe et al., 2016; Gerbino et al., 2018; Xiao et al., 2019) should, on the whole, relate to greater ascribed femininity for the prosocial student compared to the nondescript student. Prosocial students are hypothesized to be perceived as more feminine (H6a) and less masculine (H6b) than nondescript students. We thus hypothesize that prosocial female targets students will be perceived as more common (H6c) and more typical (H6d) than prosocial male targets.

6 Method

6.1 Sample

For the purpose of this study we recruited 9th grade classes from three schools from the highest academic track (“Gymnasien”), in a large city in Germany. We sampled 324 ninth

graders (52% female, 45% male, 2% diverse or missing) from 12 classes. (We originally sampled 326 ninth graders, but due to two cases in which more than 50% of data were missing, these two participants were excluded from the analysis). Most students were female, which is in line with the gender split in the highest academic track (Federal Statistical Office Germany, 2020), from which the sample was drawn. The mean age of the sample was 14.1 (*SD* = 0.6). Most students (89%) were born in Germany, and 53% of our sample reported speaking only German at home. 41% spoke German and another language, or mostly another language at home. All students participated voluntarily during regular school hours in the classroom. Students completed a paper and pencil questionnaire and were supervised by the researcher and a trained assistant. The study was deemed ethical and approved by the Senate Department for Education, Youth, and Family.

6.2 Materials and Pilot Studies

As stimulus material, four vignettes were developed: a high-achieving female student displaying prosocial behavior, a high-achieving nondescript female student, a high-achieving male student displaying prosocial behavior, and a high-achieving nondescript male student. Behaviors fit for the prosocial vignette were constructed by conducting a pilot study in which 28 university students were asked to rate a list of behaviors in terms of how pro- vs. antisocial they were on 7 pt. Likert scales. These behaviors were then rated in a second pilot by 39 university students in order to assess how warm and competent these were (7 pt. Likert scales). The chosen behaviors were rated prosocial (*M*=5.4, *SD*=0.8) and warm (*M*=5.9, *SD*=0.9), while displaying moderate competence (*M*=4.7, *SD*=1.3).

The names for the students described in the vignette were: Benjamin, Dominik, Natalie, and Linda, chosen from a dataset by Nett and colleagues (2020). They were matched in terms of education, attractiveness, intelligence, competence, and warmth, as well as being popular names among 15-year old adolescents.

The prosocial student was described as: [Name] is 15 years old. His/Her grades are good, he/she often gets good passing grades. He/She behaves in a very friendly manner in school. He/She always asks before using someone else's belongings and shares candy with others. When other students are hurt, he/she shows sympathy."

The nondescript student was characterized as: "[Name] is 15 years old. His/Her grades are good, he/she often gets good passing grades. He/She is a quiet boy/girl who behaves rather unremarkably. He/She is very quiet in class and never shows disruptive behavior." This behavioral description is identical to the description in Kessels and Heyder's (2020) control condition, this vignette attempts to show an absence of prosocial behavior, rather than something in opposition to prosociality (Carifio & Lanza, 1989).

6.3 Dependent Variables

Participants were asked to rate the targets on multiple items on a Likert scale ranging from 1-5, with higher scores representing higher agreement. The exact wording is provided in Table 1. Relevant variables included: expected praise by the teacher, expected happiness by the teacher, effort attribution, ability attribution, intelligence, number of friends, popularity, likeability, wanting to be friends with, frequency of students like this, gender typicality, masculinity, and femininity.

Furthermore, perceived masculinity and femininity of the targets was assessed using a measure by Kessels (2005). 30 items such as "helpful," "considerate," "gentle" and "proud," "powerful," "fearless," measuring femininity and masculinity respectively (15 items each), were rated on a Likert scale ranging from 1 (does not apply at all) to 7 (strongly applies).

6.4 Design

Design and procedure follows Kessels' and Heyder's (2020) study. This vignette study has a mixed 2 x 2 design, in which vignette (prosocial, nondescript) is varied within

participants, while target gender (male, female) is varied between participants. The male and female vignettes were distributed roughly equally among male and female participants.

6.5 Procedure

Participating students received a paper questionnaire, which informed them that they will read about two students (either two boys or two girls) who both perform well academically, but behave very differently in school. Participants read a description of one of the students (either prosocial or nondescript, depending on counterbalancing) and were asked to take a moment to picture the student, before answering questions about this student. They then read a description of the second student and again answered a series of questions. Finally, participants answered questions about themselves and provided basic demographic details. The names of the prosocial and nondescript student, as well as the order of presentation, was varied systematically.

7 Results

7.1 Analytical Plan

Data were analyzed with MPlus Version 8.6 (Muthén & Muthén, 2007). Target student behavior was coded dichotomously (0 = nondescript; 1 = prosocial). Due to the nested structure of our data, multilevel modeling with robust standard errors (MLR) were chosen to conduct the analysis. As vignettes were nested within participants, our levels were Level 1: 648 vignettes and Level 2: 324 participants.

7.2 Descriptive Results

A comprehensive list of means and standard deviations for target ratings is provided in Table 1. This table shows results separately for prosocial and nondescript targets, as well as split by gender for questions relating specifically to female or male targets (frequency of similar boys/girls and how similar to a typical boy/girl the target is). The low intraclass correlations displayed in the table shows that most variance is due to variations on level 1 (prosocial vs. nondescript) and not level 2 (participants). Table 2 displays bivariate correlations for the prosocial and nondescript targets.

Table 1
Dependent Variables, Descriptive Statistics, and Intraclass Correlations Item

Item	Endpoints scale (1-5)	M(SD) Prosocial	M(SD) Nondescript	ICC
The teacher praises him/her. (praise)	Rarely – often	4.21 (0.87)	2.82 (1.09)	0.029
The teacher is happy about him/her. (happiness)	Rarely – often	4.30 (0.79)	3.22 (1.01)	0.014
He/she earns good grades because he/she makes an effort. (effort attribution)	Not at all true – totally true	3.97 (0.81)	3.72 (1.01)	-0.003
He/she earns good grades because he/she is talented. (ability attribution)	Not at all true – totally true	3.10 (0.99)	3.31 (1.05)	0.019
How many friends does he/she have in his/her class? (number of friends)	A few – many	4.17 (0.98)	2.15 (0.90)	-0.001
How popular is he/she in his/her class? (popularity)	Not popular – very popular	4.11 (0.92)	2.21 (0.91)	-0.001
How likeable do you find him/her? (likeability)	Not at all – very	4.02 (0.92)	3.64 (0.94)	-0.003
How much would you like to be friends with him/her? (want to be friends)	Not at all – very much	3.64 (0.99)	3.34 (1.08)	-0.003
How high is his/her IQ? (intelligence)	Below average – above average	3.41 (0.62)	3.77 (0.68)	-0.003
How masculine is he/she? (masculinity)	Not masculine – very masculine	4.24 (0.89)	3.14 (0.88)	-0.003
How feminine is he/she? (femininity)	Not feminine – very feminine	5.50 (0.63)	5.28 (0.76)	-0.003
How often do boys like him exist? (frequency if male target)	Rarely – very often	2.64 (0.93)	2.69 (0.94)	-0.005
How often do girls like her exist? (frequency if female target)	Rarely – very often	3.05 (0.94)	3.07 (0.99)	-0.006
How similar is he to a typical boy? (gender typicality if male target)	Not similar – very similar	2.57 (0.89)	2.30 (0.83)	-0.005
How similar is she to a typical girl? (gender typicality if female target)	Not similar – very similar	3.31 (0.93)	2.75 (0.91)	-0.006

Note. M = mean, *SD* = standard deviation, *ICC* = Intraclass correlation for vignettes nested in participants (i.e., proportion of variance shared across vignettes within participants).

Table 2

Bivariate Correlations Separated for Prosocial and Nondescript Targets

Variables	1	2	3	4	5	6	7	8	9	10	11
1. Praise	-	.540**	-.074	.196**	.325**	.269**	.058	.153**	.099	.294**	-.011
2. Happiness	.392**	-	-.003	.207**	.247**	.215**	.198**	.192**	.120*	.207**	.100
3. Effort attribution	.168**	.168**	-	-.300**	-.075	-.046	.147**	.087	-.028	.038	.166**
4. Ability attribution	-.051	.011	-.386**	-	.130*	.166**	.116*	.182**	.254**	.141*	.009
5. Number of friends	.101	.163**	.039	.012	-	.652**	.218**	.255**	.007	.466**	-.121*
6. Popularity	.021	.129*	.065	.069	.578**	-	.289**	.242**	.003	.388**	-.080
7. Likeability	-.028	.081	.134*	.122*	.231**	.286**	-	.667**	.198**	.178**	.192**
8. Want to be friends	.023	.110*	.104	.149**	.193**	.261**	.670**	-	.174**	.291**	.137*
9. Intelligence	.041	.039	-.217**	.323**	-.102	-.062	-.013	.075	-	.142*	.201**
10. Masculinity	.028	.104	.096	.060	.338**	.343**	.256**	.340**	.027	-	-.255**
11. Femininity	.124*	.091	.213**	.002	-.001	.067	.203**	.099	-.053	-.225**	-

Note. Coefficients above the diagonal represented correlations for the prosocial condition, those below the diagonal correlations for the nondescript condition.

*** $p < .001$. ** $p < .01$. * $p < .05$.

7.3 Ascribed Teacher Reactions and Attributions

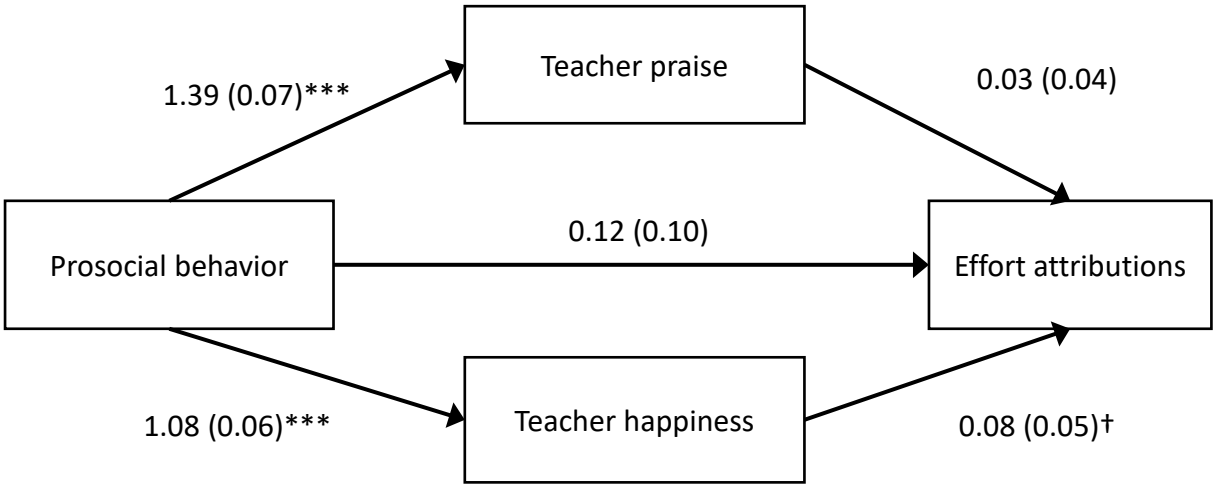
In order to investigate hypotheses 1-3b, two lower level mediation models were constructed, with all effects being fixed. The effect of target student behavior (prosocial vs. nondescript) on effort and ability attributions (Figure 1 and 2, respectively) via expected teacher reactions of praise and happiness was investigated. Note, the coefficients presented here are unstandardized for ease of interpretation.

As seen in Figure 1, target behavior had a significant impact on expectations of teacher happiness (b =1.08, SE = 0.06, p > .001) and praise (b = 1.38, SE = 0.07, p > .001), with more prosocial behavior increasing ratings of happiness and praise, supporting hypothesis 1. In terms of effort attributions, the data reveal a significant positive total effect of target students' behavior (total effect = 0.24, SE = 0.07, p = .001), supporting hypothesis 2a, which claimed that prosocial behavior would increase effort attributions. However, contrary to expectations, this effect of student's behavior on effort attribution is not mediated by teacher reactions happiness (specific indirect effect = 0.08, SE = 0.05, p = .096), or praise (specific indirect effect = 0.04, SE = 0.06, p = .568) and the direct effect of target behavior on effort attribution was not significant (direct effect = 0.12, SE = 0.10, p = .215). Thus, hypothesis 2b, which claimed that effort attribution would be mediated by teacher reactions was not supported. Figure 2 presents findings pertaining to ability attributions and shows that target student behavior had a significant effect on ability attribution (total effect = -0.21, SE = 0.07, p = .002), supporting hypothesis 3a which claimed that prosocial behavior would decrease ability attributions. No significant mediation via teacher reactions was found, showing that neither happiness (specific indirect effect = 0.07, SE = 0.06, p = .212) nor praise (specific indirect effect = 0.01, SE = 0.06, p = .864) related to ability attributions. The direct effect, however, was significant (direct effect = -0.30, SE = 0.09, p = .001). Hypothesis 3b, which predicted a mediation of ability attribution via happiness and praise, was not supported. Finally, prosocial

behavior significantly reduced ascriptions of intelligence ($b = -0.36$ $SE = 0.05$, $p > .001$), supporting hypothesis 4a.

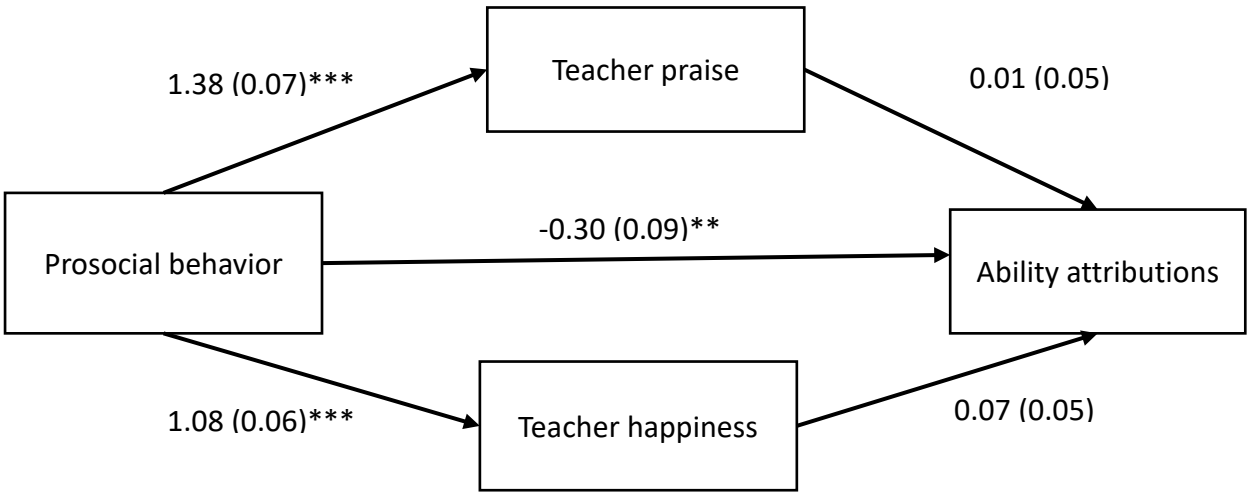
Figure 1

Lower Level Mediation Model for Effort Attributions With Fixed Effects.



Note. Unstandardized coefficients with standard errors in parentheses. $\text{Intercept}_{\text{effort attribution}}$ = 3.41(0.17)***; Level 2 variance = 0.02(0.05). $\text{Intercept}_{\text{praise}}$ = 2.82(0.06)***; $\text{Intercept}_{\text{happiness}}$ = 3.22(0.06)***. Level 1 residual $\text{variance}_{\text{effort attribution}}$ = 0.81(0.06)***; Level 1 residual $\text{variance}_{\text{praise}}$ = 0.97(0.06)***; Level 1 residual $\text{variance}_{\text{happiness}}$ = 0.82(0.05)***.
† p < .10, * $p < .05$., ** $p < .01$., *** $p < .001$.

Figure 2 *Lower Level Mediation Model for Ability Attributions With Fixed Effects.*



Note. Unstandardized coefficients with standard errors in parentheses. $\text{Intercept}_{\text{ability attribution}}$ = 3.07(0.20)***; Level 2 variance = 0.21(0.07)*.$\text{Intercept}_{\text{praise}}$ = 2.82(0.06)***; $\text{Intercept}_{\text{happiness}}$ =3.22(0.06)***. Level 1 residual $\text{variance}_{\text{ability attribution}}$ = 0.81(0.07)***; Level 1 residual $\text{variance}_{\text{praise}}$ = 0.97(0.06)***; Level 1 residual $\text{variance}_{\text{happiness}}$ = 0.82(0.05)***. † p < .10, * $p < .05$., ** $p < .01$., *** $p < .001$.

7.4 Popularity and Likeability

Target student's behavior was used to predict ascriptions of target popularity and likeability. All covariances are included in the model and Table 3 displays the outcomes of the analyses in full. As predicted by hypothesis 5a, prosocial target students were ascribed a higher number of friends (b = 2.02, SE = 0.08, p > .001 and more popularity (b = 1.92, SE = 0.07, p > .001). The targets' likeability (b = 0.38, SE = 0.07, p > .001) and participants' desire to be friends with the target (b = 0.31, SE = 0.08, p > .001), were hypothesized to be higher for prosocial targets, and this was supported by the data, supporting hypothesis 5b.

Table 3

Random Intercept Model Predicting Popularity, Likeability, and Intelligence

	Outcomes				
	Number of friends	Popularity	Likeability	Want to be friends	Intelligence
Intercept	2.15 (0.05)***	2.20 (0.05)***	3.64 (0.06)***	3.34 (0.06)***	3.77 (0.04)***
Prosocial → outcome	2.02 (0.08)***	1.92 (0.07)***	0.38 (0.07)***	0.31 (0.08)***	-0.36 (0.05)***
Level 1 residual variance	0.86 (0.08)***	0.79 (0.05)***	0.86 (0.07)***	0.88 (0.06)***	0.34 (0.03)***
Level 2 variance	0.04 (0.04)	0.05 (0.03)	0.01 (0.03)	0.18 (0.05)***	0.08 (0.03)**

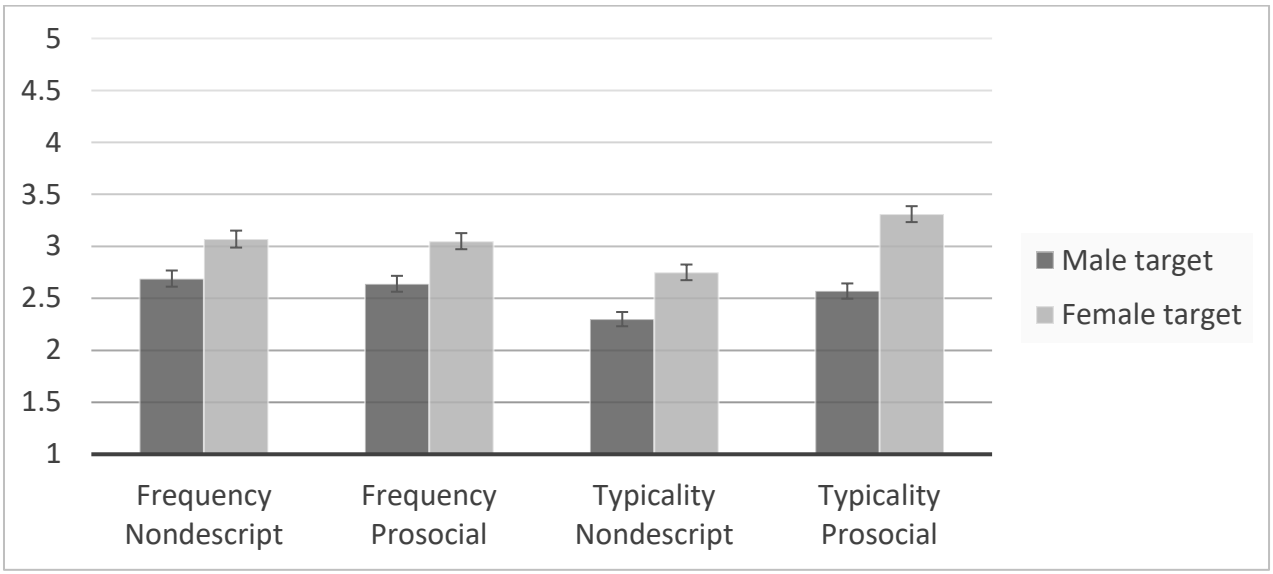
Note. Unstandardized coefficients and standard error in parentheses. * $p < .05$., ** $p < .01$., *** $p < .001$.

7.5 Gender-related Outcomes

In order to test gender-related outcomes, a cross-level interaction model including target behavior (Level 1) and target gender (Level 2) was specified. Level 1, Level 2 and their interactions were allowed to covary and all results are presented in Table 4. Target student gender was coded dichotomously (0 = male, 1 = female). We also tested whether displays of prosocial behavior resulted in higher ascriptions of femininity (6a) and lower ascriptions of masculinity (6b) in a fixed effects model, the results of which are displayed in Table 5. Prosocial behavior increased both perceived femininity (b = 0.22, SE = 0.05, p > .001) and masculinity (b = 1.10, SE = 0.07, p > .001), supporting hypothesis 6a, but going against hypothesis 6b. The analysis shows that prosocial female students were not perceived as statistically significantly more frequent (b = 0.04, SE = 0.15, p = .790) than prosocial male students, which is not in line with hypothesis 6c. However, in line with hypothesis 6d, prosocial female students were perceived as more typical than prosocial male students (b = 0.28, SE = 0.13, p = .034) (Figure 3).

Figure 3

Interaction Effects Between Targets' Gender (Male vs. Female) and Behavior (Prosocial vs. Nondescript) in Predicting Frequency and Typicality.



Note. Scales ranged from 1 to 5, error bars show standard errors.

Table 4

Cross-Level Interaction Model Predicting Target Frequency and Gender Typicality

	Outcomes	
	Frequency	Gender typicality
Intercept	2.69 (0.08)***	2.30 (0.07)***
Prosocial → outcome	-0.05 (0.11)	0.28 (0.09)**
Female target → outcome	0.37 (0.11)***	0.45 (0.10)***
Prosocial*Female target → outcome	0.04 (0.15)	0.28 (0.13)*
Level 1 residual variance	0.83 (0.07)***	0.62 (0.07)***
Level 2 residual variance prosocial → outcome	0.00 (0.08)	0.05 (0.06)
Level 2 residual variance	0.07 (0.06)	0.15 (0.05)**

Note. Unstandardized coefficients and standard errors in parentheses. * $p < .05$., ** $p < .01$., *** $p < .001$.

Table 5

Random Intercept Model Predicting Masculinity and Femininity

	Outcomes	
	Masculinity	Femininity
Intercept	3.15 (0.05)***	5.28 (0.04)***
Prosocial → outcomes	1.10 (0.07)***	0.22 (0.05)***
Level 1 residual variance	0.64 (0.06)***	0.38 (0.04)***
Level 2 variance	0.16 (0.05)**	0.12 (0.03)***

Note. Unstandardized coefficients and standard error in parentheses.
* $p < .05$. ** $p < .01$. *** $p < .001$.

8 Discussion

The present vignette study aimed to investigate prosocial behavior through an attributional lens, in order to establish whether displaying such behavior in classrooms is related to denigration of achievement, by attributing successes to effort rather than ability. The results showed that prosocial students were indeed believed to be more effortful and lower in ability than their non-descript counterparts by participating students, but expected teacher reactions did not mediate these relations. Furthermore, our analysis revealed the prosocial student to be perceived as less intelligent than the non-descript student, supporting the general impression of prosocial students as possessing lower ability. We also found prosocial students to be perceived as more popular, as having more friends, more likeable and being more desirable to have as a friend. Prosocial targets were perceived as more feminine than non-descript students, but contrary to our hypotheses, they were also ascribed more masculinity than non-descript students. Lastly, prosocial female targets were not found to be perceived as more frequent than male prosocial targets, but prosocial female target were rated as more typical than male prosocial targets in line with our prediction.

8.1 Prosocial Behavior Leads to Attributional Backlash, but is Not Mediated by Expected Praise

Our hypotheses regarding attributions were supported, we did not, however, find that expected teacher reactions of happiness and praise were related to these attributional patterns, as we had hypothesized. Kessels and Heyder (2020) found that disruptive, low achieving students received lack-of-effort rather than lack-of-ability attributions, mediated by expected teacher reprimands. In our study, expected teacher reactions were not an integral part of this attributional process. The present study thus differs from past research on the paradox of praise, in which praise led to attributions of high effort and lower ability (Graham & Chen, 2020; Meyer at al, 1979; Möller, 2005).

The paradox of praise was presented as a possible pathway by which prosocial students may experience a denigration of their achievement. The present study could not, however, establish that expected praise is the path through which this denigration occurs. There are several possible reasons why the present study found no link between expected teacher reactions and effort and ability attributions. Firstly, literature examining the paradox of praise outlines that the praise in question is applied to the achieved outcome directly, by praising the performance in a particular task (Meyer at al, 1979; Möller, 2005). In research on this attributional pattern, participants are typically provided vignettes in which fictional students receive neutral feedback or praise following a moderately difficult task (Meyer at al, 1979; Möller, 2005). However, praise in our experiment is not actually given by teachers, rather students were asked whether praise would occur. Furthermore, here, the praise is not stated to relate to the students' grades directly, but instead could be the result of non-academic, prosocial classroom behavior. Our finding that the nondescript student who achieved equally high grades as the prosocial student was perceived as receiving less praise confirms this interpretation. The present study was the first to investigate attributions following praise for behavior not directly tied to students' performance on a specific task. Perhaps, and contrary to our expectations, praise needs to be explicitly applied to the achieved outcome in order to serve as a cue for high effort and thus low-ability. Applying attributional theory to social classroom behavior was substantiated by Kessels and Heyder (2020), who showed that disruptive behavior led to distinct attributions. Perhaps disruptive behavior and prosocial behavior are not diametrically opposed to each other when it comes to the attribution of performance results, as we previously believed, but instead take on different, not complementing, roles in the classroom. Firstly, teachers report to respond more frequently and more visibly to disruptive conduct than to favourable conduct in school (Beaman & Wheldall, 2000). Disruptive behavior is met with anger (de Ruiter et al., 2019; Hagenauer et al., 2015) and reprimands (Kulinna, 2008), while prosocial behavior elicits happiness and praise (Rudy

& Grusec, 2020). While reprimands for disruptions are commonplace in school (Clunies-Ross et al., 2008), praise for prosocial behavior is seen less often (Beaman & Wheldall, 2000; Clunies-Ross et al., 2008) and decreases in adolescence (Fefer et al., 2016). Perhaps teachers' reactions to prosocial behavior are not very common in classrooms, maybe because teachers are aware of possible drawbacks for students who are singled out for their goodness ("teachers' pet"). Secondly, disruptions in the classroom are a hindrance to a conducive learning environment (Blank & Shavit, 2016; Sortkær & Reimer, 2018) and this is especially true for the disruptive students themselves (Becherer et al., 2021; Zimmermann et al., 2013), whereas prosocial acts are desirable, yet not seen as necessary for facilitating students' learning. Thus, even without mentally taking into account the norms of effort and the paradox of praise, it might be easier for participating students to relate teachers' reactions to disruptive behavior to negative academic outcomes, while praising prosocial behavior may be seen as less relevant for performance and grades.

While the paradox of praise may not be a fitting explanation for the present pattern of results, our findings are not entirely without precedent. The two fundamental social categories, warmth and competence (Bakan, 1966), have been stated to be compensatory, where high warmth is believed to be associated with lower competence when given limited information (Holoien & Fiske, 2013; Judd et al., 2005; Kervyn et al., 2010; Yzerbyt et al., 2005). In our vignettes, the prosocial student was presented as warmer than the nondescript student, and the students' competence was held constant, as both the nondescript and prosocial students were presented as earning good grades. Results show, however, that the warmer, prosocial student was perceived as lower in ability and intelligence than the non-descript student. Our findings highlight that people perceived as high on one of the fundamental dimensions may automatically be perceived as lower on the other dimension.

8.2 Prosocial Students Are Seen as More Popular and Likeable

While ascribed unfavourable attributional patterns, prosocial students had advantageous outcomes for popularity and likeability. Prosocial students were perceived as desirable friends, and as being more popular and likeable than nondescript students. A longitudinal investigation has shown higher popularity and likeability of prosocial youth (Lu et al., 2018) and cross-sectional studies (Kornbluh & Neal, 2016) revealed that friendships and peer relationships are stronger for prosocial adolescents (Lai et al., 2020; Son & Padilla-Walker, 2020; M. Wang et al., 2019). Our study further supports these findings.

8.3 Prosocial Students Are Seen as More Feminine and Masculine

In line with our hypotheses and previous research, prosocial students were perceived as more feminine than non-descript students. The link between femininity and prosocial behavior has previously been established by empirical investigations (Eisenberg et al., 2001; Hine, 2017; Quenneville et al., 2022), which have found fewer (Hine, 2017), weak (Eisenberg et al., 2001), or no (Quenneville et al., 2022) associations between prosocial behavior and masculinity. Researchers state that this is not due to masculine incompatibility with such behavior, but rather to the operationalization of prosociality as feminine (Hine, 2017). Indeed, while many prosocial acts are associated with femininity, certain types of prosocial acts can be typed masculine (Hine, 2017). An association between masculinity and prosocial behavior was found in Chinese adolescents (Ma, 2005) showing that both femininity and masculinity related to prosocial behavior. Our present research supports a link between masculinity and prosocial behavior, finding that prosocial targets were ascribed higher masculinity than non-descript students, contrary to our hypotheses. One possible explanation might be that the type of prosocial behavior described in the vignette was very proactive, and that the scale measuring ascribed masculinity comprised instrumental traits related to action and decision-making.

The present study also found that prosocial female targets were believed to be more typical than male prosocial targets, reflecting the literature on gender differences in adolescent prosocial behavior. Previous studies find female adolescents to score higher than their male peers on prosociality (Bøe et al., 2016; Gerbino et al., 2018; Lohbeck et al., 2015; Van der Graaff et al., 2018; Xiao et al., 2019). Our sample may reflect such gender differences in their judgement that prosocial behavior enacted by female adolescents is more typical. Interestingly, female prosocial targets were not seen as occurring more frequently than male prosocial targets. Why were female prosocial students seen as more typical, but not as more frequent? Adolescents may be signaling their awareness of gender stereotypes surrounding prosociality. However, when examining the frequency of this behavior they might be more inclined to draw upon their own experiences and see fewer actual differences in the prosocial behavior of boys and girls. Carlo & Randall (2002) revealed that adolescent boys score higher on measures of "public" prosociality, but more private prosocial behaviors were more often seen in girls. The prosocial behavior described in the vignettes might be considered public prosocial behavior, which could be associated with a particularly "male" type of prosociality.

8.4 Implications

Without calling into question the many positive outcomes of prosociality, the present research reveals some potential negative consequences of displaying prosocial behavior. Multiple intervention studies have aimed at increasing prosocial behavior in educational institutions (Alessandri et al., 2017; Caprara et al., 2014, 2015; Kilian et al., 2006; Kilian & Kilian, 2011; Mesurado et al., 2019; Ramaswamy & Bergin, 2009). Such interventions may not have considered the potential attributional backlash students perceived as prosocial might endure. Encouraging prosocial behavior should aim at attenuating the negative impact on attributions of ability by emphasizing that being caring and prosocial does not imply that one is less clever. Likely, students and teachers are not aware of this potential side-effect of

prosocial behavior and their (automatic) attributional patterns. An awareness about the effect would benefit teachers, whose regular communication about students could impart cues of causal attributions without their explicit knowledge. Past research has revealed gender differences in how academic candidates are described, revealing female applicants to be described with more communal terms (Madera et al., 2009) and "grindstone" words (Akos & Kretchmar, 2016). Our results suggest that communication about students, such as written reports and letters of recommendation (Akos & Kretchmar, 2016; Madera et al., 2009), should take care to avoid language that conveys unintended meaning about the targets' warmth and competence (Ebert et al., 2014; Judd et al., 2005).

As prosocial behavior is typed as feminine (Diekman & Clark, 2015; Eagly, 2009; Eisenberg et al., 2001; Quenneville et al., 2022) and seen more in female adolescents (Bøe et al., 2016; Gerbino et al., 2018; Lohbeck et al., 2015; Van der Graaff et al., 2018; Xiao et al., 2019), the finding that displaying such behavior is associated with higher ascriptions of effort and lower ascriptions of ability, despite equal academic performance, has special relevance in a school context. Past research has outlined that particularly female students' achievements are attributed more to effort, especially in masculine subjects (Espinoza et al., 2014; Fennema et al., 1990; Räty et al., 2002; Tiedemann, 2000). The present study reveals one additional possible reason for the denigration of achievement: gender stereotyped classroom behavior. Girls enact feminine behavior, such as compliancy (Jones & Myhill, 2004) and prosocial behavior (Bøe et al., 2016; Gerbino et al., 2018; Lohbeck et al., 2015; Van der Graaff et al., 2018; Xiao et al., 2019), to express femininity, just as boys enact disruptive behavior to strengthen their masculinity (Heyder et al., 2021). But while displays of disruptive behavior can have beneficial associations (Kessels & Heyder, 2020), the enactment of femininity seems to have negative attributional effects. These findings may seem to contrast past research which has presented the better fit between school and femininity or being female (Heyder & Kessels, 2013) and an incompatibility between engagement at school and masculinity (Heyder

& Kessels, 2015, 2017; Jackson & Dempster, 2009; for an overview, see Kessels et al., 2014). However, even with established trends of girls outperforming boys in terms of academic achievement (O'Dea et al., 2018; Voyer & Voyer, 2014), their accomplishments do not seem to contribute to their equal status. Literature on attributions show that students (Butler, 2014; Stetsenko et al., 2000), parents (Räty et al., 2002), and teachers (Espinoza et al., 2014; Fennema et al., 1990; Tiedemann, 2000) alike endorse that girls' academic achievements are the result of effort, rather than ability. The present results highlight that simply showing positive classroom behavior, which does not even comprise diligence and effort, but is merely prosocial in nature, could exasperate this denigrating pattern of beliefs about girls' achievements.

8.5 Limitations and Future Directions

This study gives some important insight into how attributions may be made based on classroom behavior, we are, however, limited by our use of vignettes. While the experimental design maximizes internal validity and allows for testing causal relations, the limited information provided in a vignette may not translate to actual peer and teacher interactions, in which parties are aware of a wide variety of information. Future examinations could combine experimental designs with more naturalistic settings. Quenneville and colleagues (2022) have previously investigated teacher ratings of their own students' prosocial behavior and gender role self-concept, by asking teachers to rate their students on a variety of scales. Adopting such a design, including questions relating to whether students' actual performances can be attributed to effort or ability and controlling for actual performance may offer further insight into whether teachers relate prosocial behavior of their own students to causes of their academic success. Investigating actual classroom settings would also benefit the literature on the paradox of praise, in which past research has shown less clear effects when using more open-ended measures (Binser & Försterling, 2004; Hofer & Pikowsky, 1988; Weich & Rheinberg, 1988). We would also be interested in investigating whether teachers are aware of

how praise in their classroom, even for non-academic performance, could be (mis-)interpreted by their students.

Another factor that may have had an impact on our results is the construction of our vignettes. Experimental vignettes in educational research can be beneficial (Skilling & Stylianides, 2020), although including suitable control vignettes, in which the experimental manipulation is absent, rather than presented in opposition, is challenging (Carifio & Lanza, 1989). In our study, our vignettes may have unintentionally signalled personality traits, such as agency or extroversion. The prosocial student takes action by helping classmates, makes the decision to share resources and is not shy about speaking up in order to help, while the non-descript student could be characterised as a "wallflower". The control vignette may have been perceived as introverted or shy, compared to our prosocial vignette. The impact this may have had on ascribed ability and effort is uncertain: social withdrawal relates the genius stereotype (Baudson, 2016), although teachers rated shy students as lower in intelligence (Coplan et al., 2011). In peer evaluations, shy students were also rated as less intelligent than their social peers (Ding et al., 2015; Zava et al., 2020). It is therefore notable that in our present study nondescript students, who might have appeared introverted, were viewed as more able when compared to prosocial students. Viewing the experimental vignette as more extroverted could also account for the positive social outcomes ascribed to this student. Extroverted individuals are typically highly social (Ashton et al., 2002) and have a larger friendship circle (Harris & Vazire, 2016; Pollet et al., 2011), mirroring our results for popularity and likeability of prosocial students. The prosocial vignette also captures the independent and self-assured nature of instrumentality, making ascriptions of masculinity more likely. As acts of prosociality can be viewed through a gendered lens (Diekman & Clark, 2015; Eagly, 2009), future investigations could control for the perceived masculinity and femininity of certain helping-behaviors (Hine, 2017). We may have observed a different pattern of results, had our prosocial vignette engaged in different kinds of helping behavior.

We are limited by our recruitment of students from the highest academic track in a large city. While almost half of students attend this track (Berlin Senate Department for Education, Youth and Family, 2023), our findings may not necessarily generalize to students of other academic tracks. The present study was further limited by its design, in which target behavior varied within subjects, but target gender varied between subjects. Within-variations of treatments might lead participants to ascribe more meaning to these varying parameters, leading to a boost in power (Charness et al., 2012), while a parameter that is held constant (in our case, gender of target) will receive less attention from the participants (cf. Kessels & Heyder, 2020). Creating multiple statistically comparable, realistic vignettes (Eckerd et al., 2021) and varying these by gender would have greatly increased the complexity of this investigation.

One additional point of interest we were not able to investigate empirically is whether students who describe themselves as being prosocial would also attribute more effort than ability to the prosocial student in the vignette. In other words, are prosocial students themselves making the attributions that could denigrate their achievements, or are such attributions made by others who would not be characterised as particularly prosocial? A prior study indicates that prosocial students attributed their successes in mathematics to effort to a greater extent than non-prosocial students, but that they attributed their achievements in language to both ability and effort (Redondo et al., 2014). Such findings support the notion that students' actual prosociality shapes their attributional beliefs. Future studies may investigate both participants' own prosocial tendencies, but also perceived identification with the target. High identification or perceived similarity with the target student may lead to more favourable interpretations of the target (Ajzen, 1974; Hampton et al., 2019).

8.6 Conclusion

Our present study examined an attributional backlash to enacting prosocial behavior in school, by examining whether displaying prosocial behavior can lead to a denigration of achievement. We found that prosocial students are more likely to have their good grades attributed to effort, rather than ability, when compared to nondescript students, but that teacher praise was not a significant mediator in this process. Rather, prosocial behavior alone was enough to signify that good academic performance is more likely to be due to effort, rather than innate ability. The findings are especially relevant within the context of gender equity, as previous research has shown that attribution of achievement to effort, rather than ability particularly affects female students (Butler, 2014; Espinoza et al., 2014; Fennema et al., 1990; Stetsenko et al., 2000; Tiedemann, 2000) and has been discussed as an important reason for girls' reluctance to join STEM (Chestnut et al., 2018; Kessels, 2015), as STEM is perceived as requiring not dedication, but genius (Leslie et al., 2015). The present findings highlight that displaying caring, prosocial behavior, which seems to be more in line with the female gender role (Eisenberg et al., 2001; Quenneville et al., 2022), leads into the same attribution trap as being highly engaged in class. Overall, research into classroom attributions and student behavior shows that more than one road leads to the denigration of female students' accomplishments.

References

Ajzen, I. (1974). Effects of information on interpersonal attraction: Similarity versus affective value. *Journal of Personality and Social Psychology, 29*(3), 374–380. https://doi.org/10.1037/h0036002

Akos, P., & Kretchmar, J. (2016). Gender and ethnic bias in letters of recommendation: Considerations for school counselors. *Professional School Counseling*, *20*(1), 102–114. https://doi.org/10.5330/1096-2409-20.1.102

Alessandri, G., Zuffianò, A., & Perinelli, E. (2017). Evaluating intervention programs with a pretest-posttest design: A structural equation modeling approach. *Frontiers in Psychology*, *8*, 223. https://doi.org/10.3389/fpsyg.2017.00223

Asbrock, F. (2010). Stereotypes of social groups in Germany in terms of warmth and competence. *Social Psychology*, *41*(2), 76–81. https://doi.org/10.1027/1864-9335/a000011

Ashton, M. C., Lee, K., & Paunonen, S. V. (2002). What is the central feature of extraversion? Social attention versus reward sensitivity. *Journal of Personality and Social Psychology, 83*(1), 245–252. https://doi.org/10.1037/0022-3514.83.1.245

Atkinson, C., Buie, H., Sandstrom, G., Aknin, L., & Croft, A. (2021). Testing the GRIP: An empirical examination of the gender roles inhibiting prosociality model. *Sex Roles*, *85*(7–8), 440–462. https://doi.org/10.1007/s11199-021-01229-2

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion.* Rand McNally.

Barker, G. P., & Graham, S. (1987). Developmental study of praise and blame as attributional cues. *Journal of Educational Psychology*, *79*(1), 62–66. https://doi.org/10.1037/0022-0663.79.1.62

Baudson, T. G. (2016). The mad genius stereotype: Still alive and well. *Frontiers in Psychology, 7*, 368. https://doi.org/10.3389/fpsyg.2016.00368

Beaman, R., & Wheldall, K. (2000). Teachers' use of approval and disapproval in the classroom. *Developments in Educational Psychology*, *20*(4), 153–180. https://doi.org/10.4324/9780203874677

Becherer, J., Köller, O., & Zimmermann, F. (2021). Externalizing behaviour, task-focused behaviour, and academic achievement: An indirect relation? *British Journal of Educational Psychology*, *91*(1), 27–45. https://doi.org/10.1111/bjep.12347

Becker, S. W., & Eagly, A. H. (2004). The heroism of women and men. *American Psychologist*, *59*(3), 163–178. https://doi.org/10.1037/0003-066X.59.3.163

Berlin Senate Department for Education, Youth and Family. (2023). Blickpunkt Schule: Tabellen - Allgemeinbildende Schulen 2022/2023 [Report on school: Tables - General education 2022/2023].

https://www.berlin.de/sen/bildung/schule/bildungsstatistik/blickpunkt-allgemeinbildende-schulen-2022-23.pdf?ts=1679576643

Binser, M. J., & Försterling, F. (2004). Paradoxe Auswirkungen von Lob und Tadel: Personale und situative Moderatoren [Paradoxical effects of praise and blame: Personal and situative moderators]. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, *36*(4), 182–189. https://doi.org/10.1026/0049-8637.36.4.182

Birch, S. H., & Ladd, G. W. (1998). Children's interpersonal behaviors and the teacher-child relationship. *Developmental Psychology*, *34*(5), 934–946. https://doi.org/10.1037/0012-1649.34.5.934

Blank, C., & Shavit, Y. (2016). The association between student reports of classmates' disruptive behavior and student achievement. *AERA Open*, *2*(3), 1–17. https://doi.org/10.1177/2332858416653921

Bøe, T., Hysing, M., Skogen, J. C., & Breivik, K. (2016). The strengths and difficulties questionnaire (SDQ): Factor structure and gender equivalence in Norwegian adolescents. *PLoS ONE*, *11*(5), e0152202. https://doi.org/10.1371/journal.pone.0152202

Bosson, J. K., Wilkerson, M., Kosakowska-Berezecka, N., Jurek, P., & Olech, M. (2022). Harder won and easier lost? Testing the double standard in gender rules in 62 countries. *Sex Roles*, *87*(1–2), 1–19. https://doi.org/10.1007/s11199-022-01297-y

Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, *51*(1), 5–32. https://doi.org/10.3102/00346543051001005

Brummelman, E., & Dweck, C. S. (2020). Paradoxical effects of praise: A transactional model. In E. Brummelman (Ed.), *Psychological perspectives on praise* (pp. 55–64). Routledge. https://doi.org/10.4324/9780429327667-9

Butler, R. (1994). Teacher communications and student interpretations: Effects of teacher responses to failing students on attributional inferences in two age groups. *British Journal of Educational Psychology*, *64*(2), 277–294. https://doi.org/10.1111/j.2044-8279.1994.tb01102.x

Butler, R. (2014). Motivation in educational contexts. Does gender matter? *Advances in Child Development and Behavior*, *47*, 1–41. https://doi.org/10.1016/bs.acdb.2014.05.001

Caprara, G. V., Kanacri, B. P. L., Gerbino, M., Zuffianò, A., Alessandri, G., Vecchio, G., Caprara, E., Pastorelli, C., & Bridglall, B. (2014). Positive effects of promoting prosocial behavior in early adolescence: Evidence from a school-based intervention. *International Journal of Behavioral Development*, *38*(4), 386–396. https://doi.org/10.1177/0165025414531464

Caprara, G. V., Luengo Kanacri, B. P., Zuffianò, A., Gerbino, M., & Pastorelli, C. (2015). Why and how to promote adolescents' prosocial behaviors: Direct, mediated and moderated effects of the CEPIDEA school-based program. *Journal of Youth and Adolescence*, *44*(12), 2211–2229. https://doi.org/10.1007/s10964-015-0293-1

Carifio, J., & Lanza, M. (1989). On the need to use control vignettes in social science research. *Paper Presented at the Annual Meeting of the New England Research Organization*, 2–32. https://www.unhcr.org/publications/manuals/4d9352319/unhcr-protection-training-manual-european-border-entry-officials-2-legal.html?query=excom 1989

Carlo, G., & Randall, B. A. (2002). The development of a measure of prosocial behaviors for late adolescents. *Journal of Youth and Adolescence*, *31*(1), 31–44. https://doi.org/10.1023/A:1014033032440

Charness, G., Gneezy, U., & Kuhn, M. A. (2012). Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior and Organization*, *81*(1), 1–8. https://doi.org/10.1016/j.jebo.2011.08.009

Chestnut, E. K., Lei, R. F., Leslie, S. J., & Cimpian, A. (2018). The myth that only brilliant people are good at math and its implications for diversity. *Education Sciences*, *8*(2), 65. https://doi.org/10.3390/educsci8020065

Clunies-Ross, P., Little, E., & Kienhuis, M. (2008). Self reported and actual use of proactive and reactive classroom management strategies. *Educational Psychology*, *28*(6), 693–710. https://doi.org/https://doi.org/10.1080/01443410802206700

Coplan, R. J., Hughes, K., Bosacki, S., & Rose-Krasnor, L. (2011). Is silence golden? Elementary school teachers' strategies and beliefs regarding hypothetical shy/quiet and exuberant/talkative children. *Journal of Educational Psychology*, *103*(4), 939–951. https://doi.org/10.1037/a0024551

Covington, M. V., & Omelich, C. L. (1979). It's best to be able and virtuous too: Student and teacher evaluative responses to successful effort. *Journal of Educational Psychology*, *71*(5), 688–700. https://doi.org/10.1037/0022-0663.71.5.688

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology*, *40*, 61–149.

de Ruiter, J. A., Poorthuis, A. M. G., & Koomen, H. M. Y. (2019). Relevant classroom events for teachers: A study of student characteristics, student behaviors, and associated teacher emotions. *Teaching and Teacher Education*, *86*, 102899. https://doi.org/10.1016/j.tate.2019.102899

Diekman, A. B., & Clark, E. K. (2015). Beyond the damsel in distress: Gender differences and similarities in enacting prosocial behavior. In D. A. Schroeder & W. G. Graziano (Eds.), *The Oxford handbook of prosocial behavior* (pp. 376–391). Oxford University Press.

Ding, X., Coplan, R. J., Sang, B., Liu, J., Pan, T., & Cheng, C. (2015). Young Chinese children's beliefs about the implications of subtypes of social withdrawal: A first look at

social avoidance. *British Journal of Developmental Psychology*, *33*(2), 159–173. https://doi.org/10.1111/bjdp.12081

Dompnier, B., & Pansu, P. (2010). La valeur sociale des explications causales en contexte educatif: Autoprésentation des elèves et représentation des enseignants [Social value ofcausal explanations in educational settings: Pupils' self-presentation and teachers' representations]. *Swiss Journal of Psychology*, *69*(1), 39–51. https://doi.org/10.1024/1421-0185.a000005

Eagly, A. H. (2009). The his and hers of prosocial behavior: An examination of the social psychology of gender. *American Psychologist*, *34*(8), 644–658. https://doi.org/10.1037/0003-066X.64.8.644

Ebert, I. D., Steffens, M. C., & Kroth, A. (2014). Warm, but maybe not so competent? Contemporary implicit stereotypes of women and men in Germany. *Sex Roles*, *70*(9–10), 359–375. https://doi.org/10.1007/s11199-014-0369-5

Eckerd, S., DuHadway, S., Bendoly, E., Carter, C. R., & Kaufmann, L. (2021). On making experimental design choices: Discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes. *Journal of Operations Management*, *67*(2), 261–275. https://doi.org/10.1002/joom.1128

Eisenberg, N., Zhou, Q., & Koller, S. (2001). Brazilian adolescents' prosocial moral judgment and behavior: Relations to sympathy, perspective taking, gender-role orientation, and demographic characteristics. *Child Development*, *72*(2), 518–534. https://doi.org/10.1111/1467-8624.00294

Espinoza, P., Arêas da Luz Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: Teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, *17*(1), 105–126. https://doi.org/10.1007/s11218-013-9226-6

Federal Statistical Office Germany. (2020). *Bildung und Kultur. Allgemeinbildende Schulen [Education and culture. General education]*. www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Publikationen/Downloads-Schulen/allgemeinbildende-schulen-2110100197004.pdf?__blob=publicationFile%0A

Fefer, S., DeMagistris, J., & Shuttleton, C. (2016). Assessing adolescent praise and reward preferences for academic behavior. *Translational Issues in Psychological Science*, *2*(2), 153–162. https://doi.org/10.1037/tps0000072

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69. https://doi.org/10.1007/BF00311015

Gerbino, M., Zuffianò, A., Eisenberg, N., Castellani, V., Luengo Kanacri, B. P., Pastorelli, C., & Caprara, G. V. (2018). Adolescents' rosocial behavior predicts good grades beyond intelligence and personality traits. *Journal of Personality*, *86*(2), 247–260. https://doi.org/10.1111/jopy.12309

Glock, S., & Kleen, H. (2017). Gender and student misbehavior: Evidence from implicit and explicit measures. *Teaching and Teacher Education, 67*, 93–103. https://doi.org/10.1016/j.tate.2017.05.015

Graham, S., & Chen, X. (2020). An attributional approach to teacher praise. In E. Brummelman (Ed.), *Psychological perspectives on praise* (pp. 19–26). Rand McNally. https://doi.org/10.4324/9780429327667-4

Hampton, A. J., Fisher Boyd, A. N., & Sprecher, S. (2019). You're like me and I like you: Mediators of the similarity–liking link assessed before and after a getting-acquainted social interaction. *Journal of Social and Personal Relationships, 36*(7), 2221–2244. https://doi.org/10.1177/0265407518790411

Harris, K., & Vazire, S. (2016). On friendship development and the Big Five personality traits. *Social and Personality Psychology Compass, 10*(11), 647–667. https://doi.org/10.1111/spc3.12287

Hagenauer, G., Hascher, T., & Volet, S. E. (2015). Teacher emotions in the classroom: associations with students' engagement, classroom discipline and the interpersonal teacher-student relationship. *European Journal of Psychology of Education*, *30*(4), 385–403. https://doi.org/10.1007/s10212-015-0250-0

Hallinan, M. T. (2008). Teacher influences on students' attachment to school. *Sociology of Education*, *81*(3), 271–283. https://doi.org/10.1093/oso/9780190120801.003.0005

Hareli, S., & Weiner, B. (2002). Social emotions and personality inferences: A scaffold for a new direction in the study of achievement motivation. *Educational Psychologist*, *37*(3), 183–193. https://doi.org/10.1207/S15326985EP3703_4

Haydon, T., Musti-Rao, S., Kennedy, A., Murphy, M., Hunter, W., & Boone, J. (2020). Using teacher praise with middle and high school students. *Beyond Behavior*, *29*(2), 108–115. https://doi.org/10.1177/1074295620915883

Heider, F. (1958). *The psychology of interpersonal relations*. John Wiley & Sons Inc.

Heyder, A., van Hek, M., & Van Houtte, M. (2021). When gender stereotypes get male adolescents into trouble: A longitudinal study on gender conformity pressure as a predictor of school misconduct. *Sex Roles*, *84*, 61–75. https://doi.org/https://doi.org/10.1007/s11199-020-01147-9

Heyder, A., & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles*, *69*(11–12), 605–617. https://doi.org/10.1007/s11199-013-0309-9

Heyder, A., & Kessels, U. (2015). Do teachers equate male and masculine with lower academic engagement? How students' gender enactment triggers gender stereotypes at school. *Social Psychology of Education*, *18*(3), 467–485. https://doi.org/10.1007/s11218-015-9303-0

Heyder, A., & Kessels, U. (2017). Boys don't work? On the psychological benefits of

showing low effort in high school. *Sex Roles*, *77*(1–2), 72–85. https://doi.org/10.1007/s11199-016-0683-1

Hine, B. A. (2017). Identifying the male prosocial niche: the gender-typing of prosocial behaviour across childhood and adolescence. *European Journal of Developmental Psychology*, *14*(2), 206–220. https://doi.org/10.1080/17405629.2016.1186009

Hofer, M., & Pikowsky, B. (1988). Wie Jugendliche bei freier Antwortmöglichkeit Lehrersanktionen deuten [How young people interpret teachers' sanctions when free to respond]. *Zeitschrift Für Pädagogische Psychologie*, *2*, 243–250.

Holoien, D. S., & Fiske, S. T. (2013). Downplaying positive impressions: Compensation between warmth and competence in impression management. *Journal of Experimental Social Psychology*, *49*(1), 33–41. https://doi.org/10.1016/j.jesp.2012.09.001

Jackson, C., & Dempster, S. (2009). 'I sat back on my computer … with a bottle of whisky next to me': Constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *Journal of Gender Studies*, *18*, 341–356. https ://doi.org/10.1080/09589 23090 32600 19.

Jenkins, L. N., Floress, M. T., & Reinke, W. (2015). Rates and types of teacher praise: A review and future directions. *Psychology in the Schools*, *52*(5), 463–476. https://doi.org/10.1002/pits.21835

Jones, S., & Myhill, D. (2004). 'Troublesome boys' and 'compliant girls': Gender identity and perceptions of achievement and underachievement. *British Journal of Sociology of Education*, *25*(5), 547–561.

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*(6), 899–913. https://doi.org/10.1037/0022-3514.89.6.899

Kawamura, Y., Ohtsubo, Y., & Kusumi, T. (2021). Effects of cost and benefit of prosocial behavior on reputation. *Social Psychological and Personality Science*, *12*(4), 452–460. https://doi.org/10.1177/1948550620929163

Kervyn, N., Yzerbyt, V. Y., & Judd, C. M. (2010). Compensation between warmth and competence: Antecedents and consequences of a negative relation between the two fundamental dimensions of social perception. *European Review of Social Psychology*, *21*(1), 155–187. https://doi.org/10.1080/13546805.2010.517997

Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, *20*(3), 309–323. https://doi.org/10.1007/BF03173559

Kessels, U. (2015). Bridging the gap by enhancing the fit: How stereotypes about STEM clash with stereotypes about girls. *International Journal of Gender Science and Technology*, *7*(2), 281–296.

Kessels, U., & Heyder, A. (2020). Not stupid, but lazy: Psychological benefits of disruptive

classroom behavior from an attributional perspective. *Social Psychology of Education*, *23*, 583–613. https://doi.org/10.1007/s11218-020-09550-6

Kessels, U., Heyder, A., Latsch, M., & Hannover, B. (2014). How gender differences in academic engagement relate to students' gender identity. *Educational Research*, *56*(2), 220–229.

Kilian, J. M., Fish, M. C., & Maniago, E. B. (2006). Making schools safe: A system-wide school intervention to increase student prosocial behaviors and enhance school climate. *Journal of Applied School Psychology*, *23*(1), 1–30. https://doi.org/10.1300/J370v23n01_01

Kilian, J. M., & Kilian, D. W. (2011). A school intervention to increase prosocial behavior and improve academic performance of at-risk students. *Improving Schools*, *14*(1), 65–83. https://doi.org/10.1177/1365480211399748

Klein, N., Epley, N., Porter, M., Shaw, E., & Westmorland, W. (2014). The topography of generosity: Asymmetric evaluations of prosocial actions. *Journal of Experimental Psychology: General*, *143*(6), 2366–2379. https://doi.org/10.1037/xge0000025.supp

Koglin, U., Barquero, B., Mayer, H., Scheithauer, H., & Petermann, F. (2007). Deutsche Version des Strengths and Difficulties Questionnaire (SDQ-Deu): Psychometrische Qualität der Lehrer-/Erzieherversion für Kindergartenkinder [German version of the Strengths and Difficulties Questionnaire (T4–16-SDQ): Psychometric quality of the teacher version for preschoolers]. *Diagnostica*, *53*(4), 175–183.

Kornbluh, M., & Neal, J. W. (2016). Examining the many dimensions of children's popularity: Interactions between aggression, prosocial behaviors, and gender. *Journal of Social and Personal Relationships*, *33*(1), 62–80. https://doi.org/10.1177/0265407514562562

Kulinna, P. H. (2008). Teachers' attributions and strategies for student misbehavior. *Journal of Classroom Interaction*, *42*(2), 21–30.

Lai, W., Yang, Z., Mao, Y., Zhang, Q., Chen, H., & Ma, J. (2020). When do good deeds lead to good feelings? Eudaimonic orientation moderates the happiness benefits of prosocial behavior. *International Journal of Environmental Research and Public Health*, *17*(11), 1–12. https://doi.org/10.3390/ijerph17114053

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262–265. https://doi.org/10.4135/9781483392240.n8

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 [The German self-report version of the strengths and difficulties questionnaire (SDQ-DeuS): Psychometric properties, factor structure and critical values]. *Diagnostica*, *61*(4), 222–235. https://doi.org/10.1026/0012-1924/a000153

Lu, T., Li, L., Niu, L., Jin, S., & French, D. C. (2018). Relations between popularity and prosocial behavior in middle school and high school Chinese adolescents. *International Journal of Behavioral Development*, *42*(2), 175–181. https://doi.org/10.1177/0165025416687411

Ma, H. K. (2005). The relation of gender-role classifications to the prosocial and antisocial behavior of chinese adolescents. *Journal of Genetic Psychology*, *166*(2), 189–202. https://doi.org/10.3200/GNTP.166.2.189-202

Madera, J. M., Hebl, M. R., & Martin, R. C. (2009). Gender and letters of recommendation for academia: Agentic and communal differences. *Journal of Applied Psychology*, *94*(6), 1591–1599. https://doi.org/10.1037/a0016539

Martin, A. E., & Slepian, M. L. (2021). The primacy of gender: gendered cognition underlies the big two dimensions of social cognition. *Perspectives on Psychological Science*, *16*(6), 1143–1158. https://doi.org/10.1177/1745691620904961

Matteucci, M. C., & Gosling, P. (2004). Italian and French teachers faced with pupil's academic failure: The "norm of effort." *European Journal of Psychology of Education*, *19*(2), 147–166. https://doi.org/10.1007/BF03173229

Matteucci, M. C., Tomasetto, C., Selleri, P., & Carugati, F. (2008). Teacher judgments and pupils' causal explanations: Social valorization of effort-based explanations in school context. *European Journal of Psychology of Education*, *23*(4), 421–432. https://doi.org/10.1007/BF03172750

Mesurado, B., Distefano, M. J., Robiolo, G., & Richaud, M. C. (2019). The hero program: Development and initial validation of an intervention program to promote prosocial behavior in adolescents. *Journal of Social and Personal Relationships*, *36*(8), 2566–2584. https://doi.org/10.1177/0265407518793224

Meyer, W. U. (1982). Indirect communications about perceived ability estimates. *Journal of Educational Psychology*, *74*(6), 888–897. https://doi.org/10.1037/0022-0663.74.6.888

Meyer, W. U. (1992). Paradoxical effects of praise and criticism on perceived ability. *European Review of Social Psychology, 3*(1), 259–283.

Meyer, W. U., Bachmann, M., Biermann, U., Hempelmann, M., Ploger, F.-O., & Spiller, H. (1979). The informational value of evaluative behavior: Influences of praise and blame on perceptions of ability. *Journal of Educational Psychology*, *71*(2), 259–268. https://doi.org/10.1037/0022-0663.71.2.259

Miller, A. T., & Hom, H. L. (1996). Conceptions of ability and the interpretation of praise, blame, and material rewards. *Journal of Experimental Education*, *65*(2), 163–177. https://doi.org/10.1080/00220973.1997.9943790

Möller, J. (2005). Paradoxical effects of praise and criticism: Social, dimensional and temporal comparisons. *British Journal of Educational Psychology*, *75*(2), 275–295. https://doi.org/10.1348/000709904X24744

Muthén, L. K., & Muthén, B. O. (2007). *Mplus* (8.7). Muthén & Muthén.

Nett, T., Dorrough, A., Jekel, M., & Glöckner, A. (2020). Perceived biological and social characteristics of a representative set of German first names. *Social Psychology*, *51*(1), 17–34. https://doi.org/10.1027/1864-9335/a000383

O'Dea, R. E., Lagisz, M., Jennions, M. D. & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-06292-0

Padilla-Walker, L. M., & Carlo, G. (2004). "It's not fair!" Adolescents' constructions of appropriateness of parental reactions. *Journal of Youth and Adolescence*, *33*(5), 389–401. https://doi.org/10.1023/B:JOYO.0000037632.46633.bd

Pfattheicher, S., Nielsen, Y. A., & Thielmann, I. (2022). Prosocial behavior and altruism: A review of concepts and definitions. *Current Opinion in Psychology*, *44*, 124–129. https://doi.org/10.1016/j.copsyc.2021.08.021

Pollet, T. V., Roberts, S. G. B., & Dunbar, R. I. M. (2011). Extraverts have larger social network layers: but do not feel emotionally closer to individuals at any layer. *Journal of Individual Differences, 32*(3), 161–169. https://doi.org/10.1027/1614-0001/a000048

Prawat, R. S., Byers, J. L., & Anderson, A. H. (1983). An attributional analysis of teachers' affective reactions to student success and failure. *American Educational Research Journal*, *20*(1), 137–152. https://doi.org/10.3102/00028312020001137

Quenneville, S., Talwar, V., & Bosacki, S. (2022). Teacher ratings and adolescent students' perceived social behaviours and gender-role orientations. *Journal of Gender Studies*, *31*(4), 444–456. https://doi.org/10.1080/09589236.2021.1988530

Ramaswamy, V., & Bergin, C. (2009). Do reinforcement and induction increase prosocial behavior? Results of a teacher-based intervention in preschools. *Journal of Research in Childhood Education*, *23*(4), 527–538. https://doi.org/10.1080/02568540909594679

Räty, H., Vänskä, J., Kasanen, K., & Kärkkäinen, R. (2002). Parents' explanations of their child's performance in mathematics and reading: A replication and extension of Yee and Eccles. *Sex Roles*, *46*(3–4), 121–128. https://doi.org/10.1023/A:1016573627828

Redondo, J., Ingles, C. J., & García-Fernández, J. M. (2014). Conducta prosocial y autoatribuciones académicas en educación secundaria obligatoria [Prosocial behavior and academic self-attributions in secondary education]. *Anales de Psicologia*, *30*(2), 482–489. https://doi.org/10.6018/analesps.30.2.148331

Reyna, C. (2000). Lazy, dumb, or industrious: When stereotypes convey attribution information in the classroom. *Educational Psychology Review, 12*(1), 85–110. https://doi.org/10.1023/A:1009037101170

Rudy, D., & Grusec, J. E. (2020). Praise and prosocial behavior. In E. Brummelman (Ed.), *Psychological perspectives on praise* (pp. 103–110). Routledge. https://doi.org/10.4324/9780429327667-16

Saidah, B., Louvet, E., & Pansu, P. (2019). Are students who make an effort perceived as

successful or just liked by their teachers? *Social Psychology of Education*, *22*(2), 405–419. https://doi.org/10.1007/s11218-019-09481-x

Silke, C., Brady, B., Boylan, C., & Dolan, P. (2018). Factors influencing the development of empathy and pro-social behaviour among adolescents: A systematic review. *Children and Youth Services Review*, *94*, 421–436. https://doi.org/10.1016/j.childyouth.2018.07.027

Skilling, K., & Stylianides, G. J. (2020). Using vignettes in educational research: a framework for vignette construction. *International Journal of Research and Method in Education*, *43*(5), 541–556. https://doi.org/10.1080/1743727X.2019.1704243

Son, D., & Padilla-Walker, L. M. (2020). Happy helpers: A multidimensional and mixed-method approach to prosocial behavior and its effects on friendship quality, mental health, and well-being during adolescence. *Journal of Happiness Studies*, *21*(5), 1705–1723. https://doi.org/10.1007/s10902-019-00154-2

Sortkær, B., & Reimer, D. (2018). Classroom disciplinary climate of schools and gender–evidence from the Nordic countries. *School Effectiveness and School Improvement*, *29*(4), 511–528. https://doi.org/10.1080/09243453.2018.1460382

Stetsenko, A., Little, T. D., Gordeeva, T., Grasshof, M., & Oettingen, G. (2000). Gender effects in children's beliefs about school performance: A cross-cultural study. *Child Development*, *71*(2), 517–527. https://doi.org/10.1111/1467-8624.00161

Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics*, *41*, 191–207.

Van der Graaff, J., Carlo, G., Crocetti, E., Koot, H. M., & Branje, S. (2018). Prosocial behavior in adolescence: Gender differences in development and links with empathy. *Journal of Youth and Adolescence*, *47*(5), 1086–1099. https://doi.org/10.1007/s10964-017-0786-1

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174– 1204. https://doi.org/10.1037/a0036620.

Wang, H., & Hall, N. C. (2018). A systematic review of teachers' causal attributions: Prevalence, correlates, and consequences. *Frontiers in Psychology*, *9*, 2305. https://doi.org/10.3389/fpsyg.2018.02305

Wang, M., Wang, J., Deng, X., & Chen, W. (2019). Why are empathic children more liked by peers? The mediating roles of prosocial and aggressive behaviors. *Personality and Individual Differences*, *144*, 19–23.

Weich, K.-W., & Rheinberg, F. (1988). Wie gefährlich ist Lob? Eine Untersuchung zum "paradoxen Effekt" von Lehrersanktionen [How risky is it to praise students? A study concerning "paradoxical effects" of teachers' sanctions]. *Zeitschrift Für Pädagogische Psychologie*, *2*(4), 227–233.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. Springer.

Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct.* Guilford Press.

Weiner, B. (2000). Intrapersonal and interpersonal theories of motivation from an attributional perspective. *Educational Psychology Review*, *12*(1), 1–14. https://doi.org/10.1023/A:1009017532121

Weiner, B. (2007). Examining emotional diversity in the classroom: An attribution theorist considers the moral emotions. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education.* (pp. 75–88). AcademicPress.

Weiner, B., & Kukla, A. (1970). An attributional analysis of achievement motivation. *Journal of Personality and Social Psychology*, *15*(1), 1–20. https://doi.org/10.1037/h0029211

Wentzel, K. R., & Asher, S. R. (1995). The academic lives of neglected, rejected, popular, and controversial children. *Child Development*, *66*(3), 754–763.

Xiao, S. X., Hashi, E. C., Korous, K. M., & Eisenberg, N. (2019). Gender differences across multiple types of prosocial behavior in adolescence: A meta-analysis of the prosocial tendency measure-revised (PTM-R). *Journal of Adolescence*, *77*, 41–58. https://doi.org/10.1016/j.adolescence.2019.09.003

Younger, M., Warrington, M., & Williams, J. (1999). The gender gap and classroom interactions: Reality and rhetoric? *British Journal of Sociology of Education*, *20*(3), 325–341.

Yzerbyt, V. Y., Provost, V., & Corneille, O. (2005). Not competent but warm... Really? Compensatory stereotypes in the French-speaking world. *Group Processes and Intergroup Relations*, *8*(3), 291–308. https://doi.org/10.1177/1368430205053944

Zakriski, A. L., Wright, J. C., & Underwood, M. K. (2005). Gender similarities and differences in children's social behavior: Finding personality in contextualized patterns of adaptation. *Journal of Personality and Social Psychology*, *88*(5), 844–855. https://doi.org/10.1037/0022-3514.88.5.844

Zava, F., Watanabe, L. K., Sette, S., Baumgartner, E., Laghi, F., & Coplan, R. J. (2020). Young children's perceptions and beliefs about hypothetical shy, unsociable, and socially avoidant peers at school. *Social Development*, *29*(1), 89–109. https://doi.org/10.1111/sode.12386

Zimmer-Gembeck, M. J., Geiger, T. C., & Crick, N. R. (2005). Relational and physical aggression, prosocial behavior, and peer relations: Gender moderation and bidirectional associations. *Journal of Early Adolescence*, *25*(4), 421–452. https://doi.org/10.1177/0272431605279841

Zimmermann, F., Schütte, K., Taskinen, P., & Köller, O. (2013). Reciprocal effects between adolescent externalizing problems and measures of achievement. *Journal of Educational Psychology*, *105*(3), 747–761. https://doi.org/10.1037/a0032793

Manuscript 3

Instrumentality Gives Girls the Edge:

Gender-Differential Relations Between Instrumentality,

Achievement Motivation, and Self-Esteem

Streck, H., Nishen, A. K., & Kessels, U. (2022). Instrumentality gives girls the edge: Gender-differential relations between instrumentality, achievement motivation, and self-esteem. *Sex Roles*, *86*(5), 379-394. https://doi.org/10.1007/s11199-021-01270-1

Note: This is the first-author version of the paper submitted to Sex Roles. Formatting that diverges from APA 7 follows journal-specific guidelines.

Abstract

Gender differences in school are often discussed in reference to a particular type of masculinity, negative masculinity, which is often conceptualized as detrimental to success. Another type of masculinity, instrumentality, has rarely been studied in schools even though instrumental characteristics are often exalted outside the academic context. The current study focuses on potential benefits that students may reap from instrumentality. The extent to which an instrumental self-concept is directly and indirectly associated with achievement motivation and self-esteem was examined for adolescent boys and girls in a structural equation model (SEM). A sample of German ninth graders ($N = 355$) completed self-report measures pertaining to their gender role self-concept, hope for success, fear of failure, and global and academic contingent self-esteem. The SEM revealed that instrumentality was associated with lower fear of failure and higher hope for success for both male and female adolescents. High scores in instrumentality were associated with greater self-esteem and lower academic contingent self-esteem. The association between instrumentality and global self-esteem was stronger for adolescent girls, and the indirect association between instrumentality and fear of failure through global self-esteem was significant only for girls. Results indicate that instrumentality can be an asset for students and that female students especially reap the benefits of an instrumental self-concept. The results are discussed in reference to the dangers of emphasizing solely the association between negative masculinity and academic failure, and the importance of studying relations with gender role self-concept separately for male and female adolescents.

Keywords: Masculinity, masculinities, instrumentality, gender role self-concept, achievement motivation, global self-esteem, contingent self-esteem, school

Introduction

School represents a meaningful period in the lives of adolescents, yet there are significant gender disparities in how students fare in school, with adolescent boys displaying worse outcomes than girls (van Hek et al., 2018; Voyer & Voyer, 2014). Some educational researchers have focused on gender and gender role self-concepts as a possible root for these outcomes. Investigations into how masculinities may be interacting with the demands of school have focused mainly on how negative aspects of masculinity are associated with detrimental outcomes (Lyng, 2009). However, the generalization that a masculine self-concept itself does not fit school requirements (cf. Kessels et al., 2014) seems premature. The present study investigates whether one dimension of masculinity—instrumentality—relates to possible positive outcomes in students. These outcomes comprise constructs relating to self-esteem and motivation, and their relation will also be investigated more closely as part of the present study. In addition, we will examine whether the relations between instrumentality and achievement motivation and self-esteem differ for male and female students. This will expand the current understanding of how the endorsement of gendered traits might have different implications for different gender groups. Overall, the present research investigates whether displaying traits traditionally associated with masculinity—namely, instrumental traits—can be an asset in school and whether these effects differ across gender.

Gender Role Self-Concepts

We employ the term *gender* to refer to social groups of men and women or boys and girls, acknowledging the social construct of gender as different from biological sex (American Psychological Association, 2020). In researching gender stereotypes (i.e., the culturally shared characteristics ascribed to men and women based on their gender membership; Myers et al., 2010, p. 467), studies have shown that people perceive agency, which includes traits such as ambition, confidence, and courage, to be more characteristic of men. Additionally, people

perceive communion, which includes traits such as sensitivity and patience, to be more characteristic of women (Eagly et al., 2020). Other constructs described as masculinity and femininity respectively are instrumentality, which refers to action and self-confidence (Spence & Helmreich, 1980), and expressiveness (Spence et al., 1974), which refers to an awareness of the emotions of others and kindness (Korlat et al., 2021; Spence & Helmreich, 1980). Such descriptors can be used to capture both impressions of others and groups as a whole, as in research on gender stereotypes, but can also be related to how one perceives oneself, in the form of attribute self-perceptions (Tobin et al., 2010). When referring to self-attributions of traits traditionally associated with masculinity and femininity, the term gender role self-concept is often used.

In order to assess different dimensions of these gender role self-concepts, scales measuring instrumentality and expressiveness (Spence et al., 1974), agency and communion (Abele & Wojciszke, 2014; Bakan, 1966), or more generally masculinity and femininity (Bem, 1974, 1981), respectively, are widely used in research. Usually, these instruments are constructed by asking research participants how desirable (Bem, 1974) or typical (Spence & Helmreich, 1980) the given traits are for a man or a woman (or for an adolescent boy or girl, respectively [Kessels, 2005; Krahé et al., 2007]).

The content of what is deemed feminine and masculine is fluid across time (Eagly et al., 2020). Especially with regard to masculinity, theorists have criticized the trait approach as inadequately portraying masculinity as a "fixed character type" (Connell & Messerschmidt, 2005, p. 847). Much literature has emphasised that masculine gender roles interact with "social, cultural, and contextual norms" (American Psychological Association Boys and Men Guidelines Group, 2018, p. 6). Several researchers recognize that multiple "masculinities" can be experienced in different ways depending on ethnicity, socioeconomic status, religion, and a host of other social group memberships (Smiler, 2004; Thompson et al., 1992). While we

acknowledge the existing theoretical conceptions on masculinities, our psychological research focuses explicitly on the impact of gendered attribute self-perceptions, as measured by endorsement of instrumental traits, for both male and female students. Throughout the paper, we use the terms employed by the original researchers (e.g., instrumentality, agency, or masculinity) when describing earlier research results.

Masculinities in School

As introduced above, disparities in school performance have been analyzed with regard to gender. The feminization of school hypothesis emerged as one explanation for the academic underachievement of boys. It broadly refers to both the overrepresentation of female staff in school (Verniers et al., 2016) and the overall association of school with femininity (Heyder & Kessels, 2013). Studies showed that students perceive school and learning as something feminine (Heyder & Kessels, 2013; Jackson & Dempster, 2009). Indeed, male adolescents can actually increase their ascribed masculinity by not engaging at school or by being disruptive (Heyder & Kessels, 2017; Kessels & Heyder, 2020). Researchers have even posited that misconduct in school was closely tied to "enactment of masculinity in adolescence" (Heyder et al., 2021, p. 70). As such, the perceived feminization of school may lead boys to perceive a general incongruence between their own gender and academic engagement (Kessels et al., 2014; Verniers et al., 2016).

The lower fit between "male and masculine" (Heyder & Kessels, 2015, p. 467) and school is not limited to student perception, as teachers also seem to associate boys with more problems. In studies simply utilizing the label of "boy" without any further characteristics, teachers perceived male students as troublesome and disruptive (Glock & Kleen, 2017; Jones & Myhill, 2004). They also reported more feelings of rejection towards boys they described as more masculine (Piché & Plante, 1991), even when masculinity as defined in this previous investigation included both negative and positive traits.

Most of the research on the role of maleness and masculinity for academic outcomes at school has looked at "boys behaving 'laddish' or 'macho'" (Lyng, 2009, p. 463), focusing on a particular subset of masculinities in school. This subset includes being loud (Heyder & Kessels, 2015), aggressive (Krahé et al., 2007), or not respecting rules and authorities (Jackson & Dempster, 2009). Studies have investigated this specific negative type of masculinity and its role in boys' underachievement, finding that negative masculinity (Kessels & Steinmayr, 2013), "laddishness," (Jackson, 2002; Jackson & Dempster, 2009), and traditional masculinity (Huyge et al., 2015; Yavorsky et al., 2015) relate to boys' poorer academic behavior and achievement. Additional research shows that an orientation towards traditional masculine gender roles relates to worse academic performance (Hadjar & Lupatsch, 2010; Yavorsky & Buchman, 2019; Yavorsky et al., 2015). Stronger beliefs in traditional gender roles have also been linked with lower levels of school belonging, particularly for boys (Huyge et al., 2015). Another study revealed that negative, but not positive, masculinity was associated with boys' disadvantageous academic help-seeking behavior and worse grades. The absence of effects of negative femininity in this study was interpreted as strong evidence that the determining factor driving the detrimental outcomes was not the negativity of traits but rather the gender typicality (Kessels & Steinmayr, 2013). Research with teachers also points at a lower perceived fit between maleness or masculinity and school. Teachers ascribed the least academic engagement to adolescent boys who enacted socially undesirable masculinity but did not react to adolescent girls showing socially undesirable femininity to the same extent (Heyder & Kessels, 2015).

Given these findings, it may be tempting to view masculinity as a detriment in school. However, since most studies have explicitly focused on negative or "laddish" masculinity (Heyder & Kessels, 2015; Huyge et al., 2015; Jackson, 2002; Jackson & Dempster, 2009; Kessels & Steinmayr, 2013; Yavorsky et al., 2015), such a conclusion discounts the fact that

there are more ways of defining and enacting masculinities. Some characteristics associated with masculinities have a positive relation to academic performance. For instance, one study revealed that assertiveness is seen as a trait required for success in school and was associated more with male than female students (Verniers et al., 2016).

In addition, studies citing masculinity as a negative asset at school stand in stark contrast with the many findings in adults on the benefits of masculine traits, defined as agency and instrumentality. In adult samples, higher instrumentality is related to lower anxiety and higher global self-esteem (Sharpe et al., 1995); and agency (Abele et al., 2016; Gebauer et al., 2013; Wojciszke et al., 2011) or masculinity (Whitley & Gridley, 1993) are positively related to global self-esteem. Professional success is also positively related to agency (Abele et al., 2008) and masculinity (Koenig et al., 2011). So far, research on instrumentality in the school context is scarce, although one study showed it was associated with higher self-esteem in Korean students (Choi et al., 2010).

The present study will first investigate whether a specific facet of masculinity, instrumentality, relates to positive outcomes in students. Secondly, we will study whether these relations differ for male and female students.

Does the Impact of Masculinity and Femininity Vary With Gender?

The results described above highlight the relevance of agency and instrumentality as important aspects of masculinities for well-being and professional success, though it would be simplistic to assume that this relation is equal for men and women or boys and girls. Studies that investigate both gender and gender role self-concept show that traits associated with masculinity and femininity can have differential relations in women and men. Indeed, the associations with masculinity and femininity seem to be even stronger when they do not align with gendered expectations and are thus experienced counter-stereotypically. The relation between masculinity and global self-esteem is often stronger in women than in men (Whitley,

1988), a finding also seen in the relation between agency and global self-esteem (Hirokawa & Dohi, 2007). While some researchers find higher well-being in men than women scoring high on femininity (Matud et al., 2019), others report no relation between femininity and self-esteem (Whitley, 1988) and communion and self-esteem in men (Hirokawa & Dohi, 2007; Wojciszke et al., 2011). Whether these trends of gender-differential effects extend to students remains to be investigated.

In adolescence, gender typicality is rewarded by higher peer status (Egan & Perry, 2001), and this was found to be especially true for boys (Jewell & Brown, 2014). Our research is based on the idea that precisely because instrumentality is perceived as typical for male, but less typical for female individuals (given the construction of the respective measurement tools), the self-reported instrumentality of female adolescents should be more predictive. While male adolescents scoring high on instrumentality may simply be applying gender stereotypes to themselves, girls' endorsement of instrumentality might indicate that they agree with these items because it reflects traits they feel they possess. As "there is little data regarding the implications of masculinity for women" (Smiler, 2006, p. 622), the present research aims to illuminate how instrumentality in the school context impacts both male and female adolescents.

Relations of Gender Role Self-Concept to Students' Motivation and Self-Esteem

Important variables relevant to school students include achievement motivation as well as global and contingent self-esteem. Achievement motivation captures how students engage with challenges and process successes and failures (McClelland et al., 1953). Global self-esteem relates to an overall estimation of the self (Rosenberg, 1965), while contingent self-esteem captures the degree to which self-esteem is dependent on external factors (Crocker, Luhtanen, et al., 2003). The relevance of these constructs is supported by their relation to well-being (Burwell & Shirk, 2006; Otterpohl et al., 2020) and their continued

influence into adulthood (Abele et al., 2016; Crocker & Park, 2004; Gebauer et al., 2013; Otterpohl et al., 2020; Whitley & Gridley, 1993; Wojciszke et al., 2011). The relations of these variables with gender role self-concepts have not been sufficiently investigated in adolescent samples, a gap this paper attempts to address. A visual representation of the proposed directions and relations between constructs can be found in Figure 1.

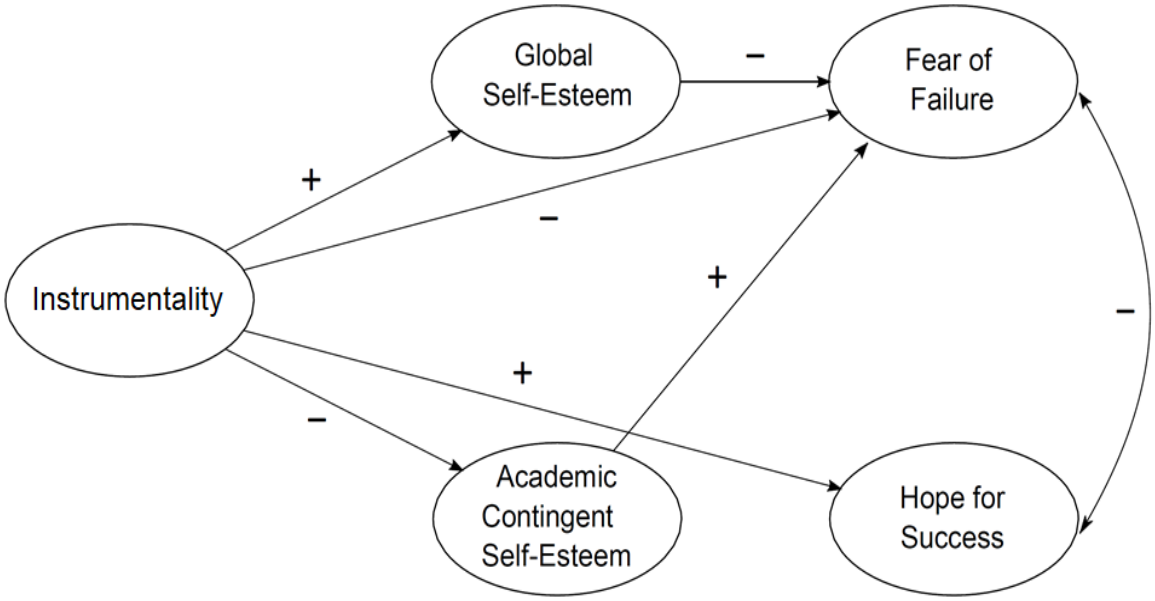


Fig. 1 Theoretical Path Model of the Relations Between Instrumentality, Self-Esteem, Academic Contingent Self-Esteem, and Motivational Constructs (created with Inkscape).

Gender Role Self-Concepts and Achievement Motivation

The most prominent model in research focusing on gender differences in students' motivation is the Eccles-Parsons et al. (1983) expectancy-value model of achievement-related choices, persistence, and performance (e.g., Eccles et al., 1983; Eccles & Wigfield, 2020; Wigfield & Eccles, 2000). This model has been applied to various domains, with many studies focusing on girls' disadvantage in STEM resulting from their lower expectancy of success and task values related to STEM subjects. However, little attention has been given to the impact of students' gender role self-concepts (Eccles & Wigfield, 2020). Recent research

on other motivational constructs such as achievement goal orientations (Butler, 2019) or controlled versus autonomous motivational regulation (Vantieghem & Van Houtte, 2018) state that investigations of gender differences tend to be scarce (Butler, 2019) or "secondary" (Vantieghem & Van Houtte, 2018, p. 381) to the central research question, rather than the focus.

The motivational construct examined in the present study is achievement motivation, comprising fear of failure and hope for success (McClelland et al., 1953). Fear of failure can be characterized as experiencing discomfort at the thought of failing to display one's abilities, while hope for success encompasses the desire to improve one's abilities (Engeser, 2005; Steinmayr et al., 2019; Steinmayr & Spinath, 2009a). Hope for success was found to contribute positively to general school performance (Steinmayr & Spinath, 2009b), while fear of failure was associated with higher anxiety, self-handicapping, and low resilience (Martin & Marsh, 2003). A meta-analysis (Severiens & Dam, 1998) as well as a study from Germany (Engeser, 2005) highlighted that women scored higher on fear of failure than men. Comparable findings have been obtained among German high school students (Steinmayr & Spinath, 2008) and Norwegian elementary school students (Gjesme, 1983), while these studies showed no gender differences in hope for success.

Little is known about how gender role self-concepts relate to hope for success or fear of failure in middle adolescence. A review of studies conducted with college-aged samples highlights that individuals who rated themselves as holding high levels of stereotypically masculine traits had a more advantageous achievement motivation than those who reported higher levels of stereotypically feminine traits (Elmen, 1991; Major, 1979), a finding seen also for individuals with higher agency (Strage, 1997). One early study with college students (Orlofsky & Stake, 1981) revealed that self-attributed masculinity seemed to be a more powerful predictor of achievement motivation than gender. Participants scoring high in

masculinity also scored higher in hope for success (Henschen et al., 1982) as well as scoring lower in fear of failure (Orlofsky & Stake, 1981).

Gender Role Self-Concepts and Self-Esteem

Self-esteem is an important resource in both adults and adolescents (Fairlamb, 2020). *Global self-esteem* in a school context has been shown to relate to higher grades (Crocker, Karpinski, et al., 2003), higher motivation, and lower school burnout (Herrmann et al., 2019). In contrast, having high academic *contingent self-esteem* places additional stress on students to achieve good results, as not only their academic reputation but their feeling of self-worth is perceived to be at risk (Burwell & Shirk, 2006; Crocker & Park, 2004; Fairlamb, 2020; Herrmann et al., 2019). Indeed, researchers have found relations between highly contingent self-esteem and depressive symptoms (Burwell & Shirk, 2006; Otterpohl et al., 2020) and lower global self-esteem (Moore & Smith, 2018; Schöne et al., 2015; Schöne & Stiensmeier-Pelster, 2016). The following sections will illuminate how self-esteem has been investigated in student samples in relation to gender and gender role self-concepts.

Global Self-Esteem. Mirroring research with adult samples, studies typically show that adolescent boys score higher on measures of global self-esteem than girls (Bleidorn et al., 2016; Block & Robins, 1993; Kling et al., 1999; Knox et al., 1998; Schöne & Stiensmeier-Pelster, 2016). Extending focus to gender role self-concept has revealed the positive relation between masculinity and global self-esteem in adolescents (Allgood-Merten & Stockard, 1991; Buckley & Carter, 2005; Cate & Sugawara, 1986), a relation also found for instrumentality (Choi et al., 2010) and agency (Stein et al., 1992).

Some early studies have revealed that the association between masculinity and global self-esteem is stronger for female than male adolescents (Allgood-Merten & Stockard, 1991; Cate & Sugawara, 1986), parallel to findings in adult samples with measures of masculinity (Matud et al., 2019; Whitley, 1988) and agency (Hirokawa & Dohi, 2007). Recently, Fu and

Wang (2020) argued that there is a lack of more current research concerning the relation between gender role self-concept and self-esteem in adolescence.

Academic Contingent Self-Esteem. Adolescent girls score higher on academic contingent self-esteem than boys (Herrmann et al., 2019; Moore & Smith, 2018; Schöne et al., 2015; Schöne & Stiensmeier-Pelster, 2016; van der Kaap-Deeder et al., 2016). However, the relation between gender role self-concept and contingent self-esteem is less established. An investigation on women in Poland did not find an association between masculinity, as measured by an adaptation of the BSRI (Bem, 1974), and academic contingent self-esteem (Mandal & Moroń, 2019), but research with adolescents in schools is needed.

One may assume that a negative association between academic contingent self-esteem and instrumentality would emerge in student samples, given the stereotype of school and learning being feminine (Heyder & Kessels, 2013, 2015). Students scoring high on instrumentality might be less likely to tie their self-esteem to their academic performance, while students scoring high in expressiveness may stake more of their self-esteem on achievements in this supposedly feminine sphere. The inverse relation between global and contingent self-esteem (Moore & Smith, 2018; Schöne et al., 2015; Schöne & Stiensmeier-Pelster, 2016) and the strong association of global self-esteem with masculinity (Allgood-Merten & Stockard, 1991; Buckley & Carter, 2005; Cate & Sugawara, 1986) and instrumentality (Choi et al., 2010) further leads us to believe that instrumentality would show a negative relation with academic contingent self-esteem.

Study Overview and Hypotheses

The present study investigated the importance of instrumental gender role self-concepts (*instrumental self-concepts*) for secondary school students' achievement motivation and self-esteem while comparing whether these relations differ for male and female students (see Figure 1). We also considered achievement motivation and self-esteem simultaneously,

as these constructs have been shown to relate to one another. Given the existing findings in adult samples on the positive outcomes of masculinity, agency, and instrumentality (Abele, 2003; Abele & Candova, 2007; Abele et al., 2008; Hirschy & Morris, 2002; Matud et al., 2019; Whitley, 1984), we constructed a model in which global self-esteem, academic contingent self-esteem, fear of failure, and hope for success were regressed on instrumentality.

We expected that students scoring high on instrumentality will likely also display higher global self-esteem, as traits associated with masculinity are highly valued and have been associated with higher self-esteem (Allgood-Merten & Stockard, 1991; Buckley & Carter, 2005; Cate & Sugawara, 1986, Choi et al., 2010; Stein et al., 1992). We further hypothesized that high scores in instrumentality would be associated with low academic contingent self-esteem, due to findings that school and learning is perceived as a feminine domain (Jackson & Dempster, 2009; Heyder & Kessels, 2013). Furthermore, we expected that instrumentality would be positively associated with hope for success and negatively associated with fear of failure, as instrumentality captures traits like independence and fearlessness and the respective relations had been reported in young adults (Elmen, 1991; Henschen et al., 1982; Orlofsky & Stake, 1981; Major, 1979; Strage, 1997).

Our model will also investigate relations between the two self-esteem constructs and fear of failure. We expect that high global self-esteem should be associated with lower fear of failure since individuals who have a positive view of the self are expected to be secure in their abilities. This relation has been found in the past, with researchers describing the strong negative relation between global self-esteem and fear of failure (e.g., Jöstl et al., 2012; Radziwiłłowicz & Macias, 2014). Academic contingent self-esteem is expected to relate positively to fear of failure, as students who place their sense of self-worth in academic domains are more likely to report distress at the idea of failing in this domain. Furthermore,

researchers have stated that fear of failure would be particularly pronounced in individuals with highly contingent self-esteem (Crocker & Park, 2004). Research offers robust evidence that fear of failure is related to the different aspects of self-esteem (Aktop & Erman, 2006; Jöstl et al., 2012; Radziwiłłowicz & Macias, 2014) but less strong evidence for how hope for success relates to self-esteem (Aktop & Erman, 2006). Consequently, an indirect relation via global self-esteem and academic contingent self-esteem is hypothesized only for fear of failure.

Overall, we investigate how students' instrumental self-concept is related to fear of failure via global self-esteem and academic contingent self-esteem and how their instrumental self-concept relates to hope for success. Additionally, based on research showing larger relations of counter-stereotypical self-descriptions on a variety of outcomes in adult samples (Hirokawa & Dohi, 2007; Whitley, 1988; Wolfram et al., 2009), we expected the relations of instrumentality with the other constructs to be moderated by participants' gender, with stronger relations to be found among female students than male students. We also conduct two exploratory analyses using grades and expressiveness as additional variables. We regressed fear of failure and hope for success on grade average and also constructed a separate model in which both self-esteem constructs and achievement motivation were regressed on expressiveness.

Method

Participants and Procedure

In total, 648 9th grade students participated in one larger study. The scales for the motivational constructs (hope for success and fear of failure) had been administered to a subsample of 288 students. The students completed a paper and pencil questionnaire while supervised by two trained student assistants. The data collection took place in the classroom

in the middle of the year and was conducted instead of a regular lesson. All participants attended the "Gymnasium," the highest academic track in Germany. As girls are overrepresented in this track (53% female students; Federal Statistical Office Germany, 2020), this imbalance was mirrored in our sample too, resulting in a relatively low number of male adolescents for our main analyses. In order to attain a balanced sample, 60 male adolescents were additionally chosen at random from the group of students that had not been given the measures of the motivational constructs and added to the sample. Three students did not indicate their gender and were excluded from the analyses. Thus, the final sample consisted of 355 students from six schools.

On average, the students were 14.2 years old (*SD* = 0.5, *range* = 13–17) and half of the sample identified as female (50.6%). In addition to these constructs, participants also provided information on personality (such as the big-five personality traits) and school-related questionnaires (such as perceived support through teachers, academic self-regulation), none of which were used in the present research. Neither socioeconomic status nor parental education were not collected as part of this study.

Measures

Instrumentality and expressiveness were assessed using a German measure (Kessels, 2005) comprising 30 attributes (15 expressive: e.g., "helpful," "considerate," "gentle"; 15 [instrumental]: e.g., "proud," "powerful," "fearless"; for all items, see Table S1 in the Appendix). This scale had been constructed by asking German adolescents how typical different traits were for a girl and for a boy. Students in the present study were asked to indicate the degree to which the traits apply to themselves on a Likert scale ranging from 1 (*does not apply at all*) to 7 (*strongly applies*). Based on inadequate loadings ($\lambda < .45$) in confirmatory factor analyses, six items were removed from the instrumentality scale and six items from the expressiveness scale (all factor analyses were conducted in the second half of

the sample which did not fill in the motivational construct scales). Both final scales showed good reliability (Cronbach's $\alpha_{instrumentality}$ = .86 and Cronbach's $\alpha_{expressiveness}$ = .85).

Global self-esteem was assessed with the Self-Esteem Inventory for Children and Adolescents (SEKJ; Schöne & Stiensmeier-Pelster, 2016), a German scale used to assess self-esteem on a Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Examples from the 10-item scale include *"I like myself,"*and *"I am completely satisfied with myself."* Based on a confirmatory factor analysis, two items that had low loadings and referred to liking specific aspects of oneself (looks and name) were excluded. The final scale had good reliability, Cronbach's α = .88.

Academic contingent self-esteem was measured with the respective subscale from the above-mentioned Self-Esteem Inventory for Children and Adolescents (Schöne & Stiensmeier-Pelster, 2016). This 11-item scale, also presented in German, was scored on a Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*), e.g., *"When I get better grades than my classmates, I feel more worthy."* A confirmatory factor analysis indicated that all items should be retained. Reliability of the scale was very good, Cronbach's α = .91.

Achievement motivation was measured using the German achievement motivation scale (Engeser, 2005), which separates achievement motivation into hope for success and fear of failure, each subscale consisting of 5 items (10 items total). Hope for success was measured with statements such as *"I would like to be given more difficult tasks,"* while fear of failure was characterized by items such as *"If I do not understand a problem immediately I start feeling anxious."* Items were rated on a Likert scale ranging from 1 (*does not apply at all*) to 4 (*strongly applies*). Based on inadequate loadings in a confirmatory factor analysis, three items were removed from the hope for success scale. No changes were indicated for the fear of failure scale. Reliability was adequate for the hope for success scale (Spearman-Brown coefficient = .77) and good for fear of failure (Cronbach's α = .83).

In an exploratory analysis outlined in more detail in the following section, we also included students' achievement. This academic achievement score consisted of self-reported grades in the three main subjects (German, Math, English), which were then averaged. In the German school system, higher grade averages indicate worse performance.

Analysis Plan

We tested our hypotheses using structural equation modelling, which allows us to test our set of hypotheses in one rather than multiple models. Moreover, we used latent variables for all constructs, making our path estimates more reliable. All analyses were conducted using the lavaan package (Rosseel, 2012) in R (R Core Team, 2017). We conducted a multi-group SEM to compare the relations between instrumentality, achievement motivation, and self-esteem constructs separately for male and female students. To reduce the number of paths tested, parcels were used instead of items for constructs with more than five items. The rationale for using parcels and the analysis plan were preregistered and can be accessed via the Open Science Framework (https://osf.io/m8n49/). Please note that we provided a respecified model guided by theory as recommended by Kline (2016), which is the model presented here.

Exploratory Analyses

In addition to this analysis, we conducted exploratory analyses that included two control variables. In one set of analyses, we included grade average as a predictor of fear of failure and hope for success and tested these additional paths for equivalence between the genders. In another set of analyses, we added expressiveness to the model as a predictor for self-esteem, self-esteem contingency, and motivational constructs. These additional paths were tested for equivalence as well.

Missingness

Most commonly, participants did not have any missing values (72.1%), and another 16.9% had missingness by design on the fear of failure and hope for success scales (the 60 male adolescents randomly chosen from the other subsample). The remaining 9.6% had either failed to fill in individual items or did not answer the 11-item scale on contingency of self-esteem. In total, 3.9% of cells were missing. Since we knew that the majority of missingness was missing by design (MCAR), we used FIML (Full-Information Maximum Likelihood) to estimate missing values for all SEM analyses.

Results

Descriptive Statistics

Descriptive statistics as well as correlations between the means of instrumentality and the self-esteem and motivational constructs are presented in Table 1, separated by gender. Correlations indicated that all constructs were (in part marginally) associated with one another for female adolescents, whereas this was not the case for male adolescents. On all scales, male and female students significantly differed in their mean values such that male students described themselves as more instrumental, less expressive, more hopeful, and less fearful. Moreover, they had higher global self-esteem, and this self-esteem was less contingent on academic competence.

Table 1

Correlations and Descriptive Statistics Separated for Male and Female Adolescents

Variables	1	2	3	4	5	6
1. Instrumentality	—	–.161*	.495***	–.278***	–.483***	.275***
2. Expressiveness	.113	—	.058	.040	.134+	.215**
3. Self-Esteem	.345**	.204*	—	–.373***	–.460***	.194*
4. Academic contingency of self-esteem	.129	.051	–.257*	—	.478***	.042
5. Fear of Failure	–.132	.099	–.295**	.454***	—	–.352***
6. Hope for Success	.037	.174+	.076	–.278**	–.361***	—
M (*SD*) female	4.25 (1.09)	5.40 (0.96)	3.21 (0.92)	3.04 (0.88)	2.40 (0.67)	2.77 (0.46)
M (*SD*) male	4.96 (0.93)	5.07 (0.99)	3.75 (0.79)	2.53 (0.93)	2.01 (0.66)	2.97 (0.50)
t-value	–6.39***	3.05**	–5.69***	5.26***	4.80***	–3.45***

Note. Coefficients above the diagonal represented correlations for female adolescents, those below the diagonal correlations for male adolescents. Correlations are based on list-wise deletion.

***$p < .001$. **$p < .01$. *$p < .05$. +$p < .10$.

Main Model: Instrumentality

We regressed global self-esteem, academically contingent self-esteem, fear of failure, and hope for success on instrumentality. In addition, fear of failure was regressed on global self-esteem and academically contingent self-esteem (Figure 2). The χ^2-test for the fully-latent two-group model was significant, $\chi^2(205) = 297.89$, $p < .001$, which could be indicative of poor model fit; however, model fit indices presented a different picture with adequate to good fit, CFI = .96, TLI = .95, RMSEA = .05, SRMR = .09. Thus, the results should be interpreted

with caution. To test whether gender moderated any of the paths, we followed a model trimming approach as described by Kline (2016). From the starting model in which all paths were free to vary across the two groups, nested models were tested in which each regression path was individually constrained to be equal for both genders. In one instance this reduced model fit at $p < .05$, indicating that the coefficients of this path were different for male and female adolescents. For adolescent girls the relation between instrumentality and global self-esteem was stronger, since constraining the coefficient to be equal for both genders worsened model fit, $\chi^2\Delta(1) = 4.89$, $p = .03$. All other regression paths were constrained to be equal. Thus, in the final model we constrained all paths apart from global self-esteem on instrumentality to be equal for male and female adolescents. This model had good model fit when considering the model fit indices, CFI = .96, TLI = .96, RMSEA = .05, SRMR = .09, but the χ^2-test remained significant, $\chi^2(210) = 304.12$, $p < .001$. In the text, we report the values of the coefficients constrained to be equal (i.e., collapsed across both genders), with the exception of the coefficient that differed significantly between boys and girls (see Figure 2 for the unconstrained model, which includes separate values of all coefficients).

The regression coefficients indicated that ascribing more instrumental traits to oneself was associated with greater self-esteem (female adolescents: $b = 0.44$, $SE = 0.07$, $p < .001$; male adolescents: $b = 0.25$, $SE = 0.09$, $p = .005$), and these findings are thus consistent with our hypotheses. For example, an adolescent girl who rated herself one point higher on the scale of instrumentality than the mean for girls also rated her self-esteem about half a point above the mean on the self-esteem scale in this group. An adolescent boy rating himself one point above the mean for instrumentality also had a higher self-esteem score than a boy with an average score on instrumentality —but only about one fourth of a point higher. Consistent with the hypotheses laid out above, both female and male students who scored higher on instrumentality also scored lower on academic contingent self-esteem ($b = -0.20$, $SE = 0.05$,

$p < .001$). As such, a student with instrumentality one point above the mean also had about a fifth of a point lower score on contingency of self-esteem compared to students with average instrumentality. Additionally, greater instrumentality was associated with less fear of failure for both male and female adolescents ($b = -0.14$, $SE = 0.04$, $p = .001$) and was related to greater hope for success ($b = 0.15$, $SE = 0.06$, $p = .012$), which is in line with the hypothesized associations. In turn, higher global self-esteem and lower contingency of self-esteem were associated with less fear of failure ($b = -0.12$, $SE = 0.05$, $p = .010$, and $b = 0.32$, $SE = 0.06$, $p < .001$, respectively), findings which are consistent with our hypotheses.

Next, we examined the indirect effects of instrumentality on fear of failure through global self-esteem. For female adolescents, instrumentality was negatively associated with fear of failure via global-self-esteem, $b = -0.05$, $SE = 0.02$, $p = .014$, bootstrapped 95% CI [–0.10, –0.02]. However, the indirect effect for male adolescents did not reach the traditional significance threshold, $b = -0.03$, $SE = 0.02$, $p = .055$, bootstrapped 95% CI [–0.07, –0.01]. Lastly, the indirect effect via contingency of self-esteem also differed from zero, $b = -0.06$, $SE = 0.02$, $p = .002$, bootstrapped 95% CI [–0.12, –0.03]. Overall, this model explained fear of failure well ($R^2_{females} = .470$; $R^2_{males} = .362$) as well as self-esteem when it came to female students ($R^2_{females} = .302$; $R^2_{males} = .085$). However, self-esteem contingency and hope for success remained largely unexplained for both genders ($R^2_{females} = .094$; $R^2_{males} = .057$, and $R^2_{females} = .085$; $R^2_{males} = .043$, respectively).

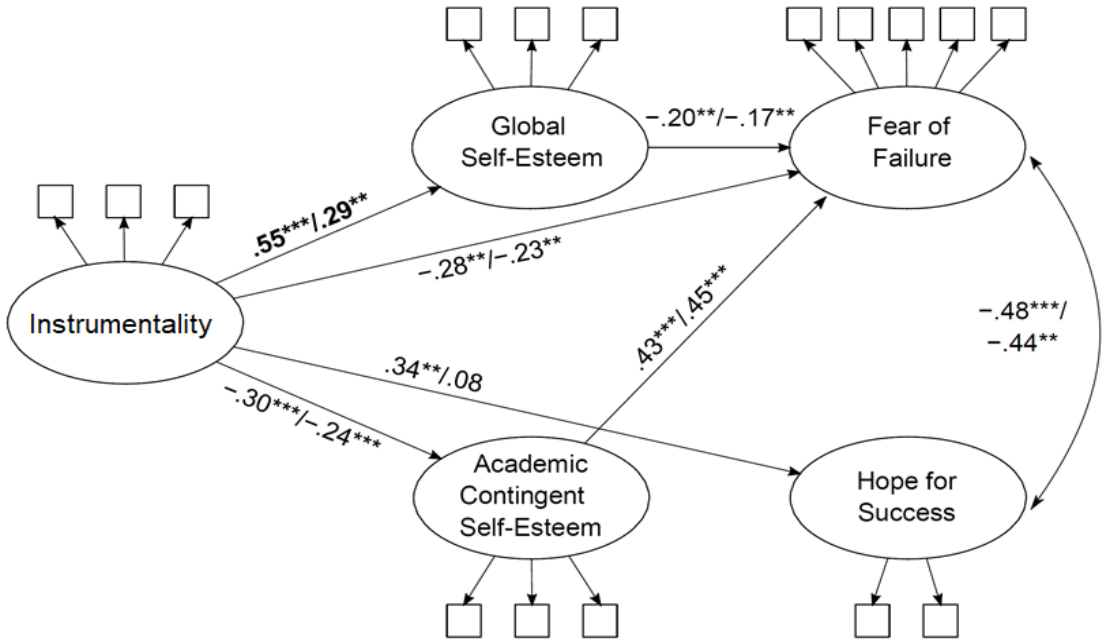


Fig. 2 Main Model Regressing Instrumentality on the Self-Esteem and Motivational Constructs.

Note. Results of the fully-latent two-group model without equality constraints for the path coefficients. Standardized coefficients are first presented for female and then for male adolescents. The coefficients are significantly different for boys and girls only for the path between masculinity and global self-esteem (in bold). $\chi^2(205) = 297.89$, $p < .001$, CFI = .96, TLI = .95, RMSEA = .05, SRMR = .09.

Exploratory Analyses

In addition to our main analysis, we included two additional variables that could relate to the constructs examined in our main model. First, we included expressiveness to test whether this aspect of gender role self-concept might similarly relate to both self-esteem and motivational constructs. In a second model, we added grade average to the model and regressed fear of failure and hope for success on it. The results of the exploratory analyses are presented in Table 2.

Table 2

Standardized Regression Coefficients for the Two Exploratory Models

Variables	Model 1: expressiveness	Model 2: grades
Self-esteem		
Instrumentality	**0.56***/0.28****	**0.55***/0.29****
Expressiveness	0.08/0.09	-
Academic contingency of self-esteem		
Instrumentality	–0.29***/–0.23***	–0.30***/–0.24***
Expressiveness	0.12^{+}/0.12^{+}	-
Fear of failure		
Self-esteem	–0.23**/–0.20**	–0.18*/–0.16*
Academic contingency of self-esteem	0.41***/0.42***	0.43***/0.45***
Instrumentality	–0.25**/–0.21**	–0.30***/–0.24***
Expressiveness	0.15*/0.16*	-
Grade average [a]	-	0.26***/0.26***
Hope for success		
Instrumentality	0.33**/0.23**	0.29**/0.21**
Expressiveness	0.18^{+}/0.15^{+}	-
Grade average	-	–0.47***/–0.40***
Indirect effects		
Instrumentality -> SE -> FoF	–0.06**/–0.03*	–0.05*/–0.03*
Instrumentality -> ACSE -> FoF	–0.06**	–0.06**
χ^2test	$\chi^2(302) = 426.05^{***}$	$\chi^2(238) = 335.03^{***}$
Model Fit indices	CFI = .96, TLI = .95, RMSEA = .05, SRMR = .09	CFI = .96, TLI = .95, RMSEA = .05, SRMR = .09

Note. Coefficients for female adolescents are presented first, then coefficients for male adolescents. Standardized coefficients that are constrained to be equal may differ slightly between the genders due to differences in variance. None of the relations between expressiveness and grades with the other variables differed significantly by gender. SE = self-esteem, ACSE = Academic contingency of self-esteem, FoF = fear of failure.
[a] Note that in the German school system, higher grades indicate worse performance.
$^{***}p < .001$. $^{**}p < .01$. $^{*}p < .05$. $^{+}p < .10$.

When including expressiveness in the model, it was positively associated with fear of failure such that greater expressiveness was associated with greater fear of failure, though the coefficient was small in size. Expressiveness was not associated with other constructs at

$p < .05$. Including expressiveness in the model did not change the relations between the other constructs, though the size of the coefficients of the paths predicting fear of failure from instrumentality and self-esteem contingency was somewhat smaller.

When grades were included as predictors of fear of failure and hope for success they were associated with fear of failure but more so with hope for success. Specifically, worse grades were related to lower hopes for success and greater fear of failure. Including grades in the model did not affect the size of other path coefficients.

Discussion

We investigated whether having a strong instrumental self-concept would be associated with achievement motivation and two aspects of self-esteem and whether these relations were equally strong for both adolescent girls and boys. We found that a stronger instrumental self-concept was associated with greater global self-esteem and lower academic contingent self-esteem, which, in turn, were each associated with lower fear of failure. One of the examined relations was different in female students than in male students. In particular, the relation between instrumentality and global self-esteem was stronger for adolescent girls. Moreover, instrumentality indirectly related to fear of failure through global self-esteem in female but not male adolescents.

Our descriptive results showed that adolescent boys scored higher on global self-esteem than adolescent girls. The structural equation model indicated that instrumentality was positively associated with global self-esteem in both genders, with this relation being stronger in female students. Our present results support past findings in which male gender (Bleidorn et al., 2016; Block & Robins, 1993; Kling et al., 1999; Knox et al., 1998; Schöne & Stiensmeier-Pelster, 2016) and higher masculinity (Allgood-Merten & Stockard, 1991; Buckley & Carter, 2005; Cate & Sugawara, 1986), agency (Stein et al., 1992), and

instrumentality (Choi et al., 2010) have been found to relate to higher global self-esteem in adolescents.

Past research on academic contingent self-esteem has established that female students tend to score higher in this domain, indicating that they attach their self-esteem to their academic performance to a greater extent than male students (Moore & Smith, 2018; Schöne et al., 2015; Schöne & Stiensmeier-Pelster, 2016; van der Kaap-Deeder et al., 2016). Our present research corroborates these findings, with adolescent girls scoring higher on this measure than boys. Research on the relation between gender role self-concept and academic contingent self-esteem is scarce, though one study using an adult sample found no association between these constructs (Mandal & Moroń, 2019). We hypothesized that students who see themselves as having more instrumental characteristics may also place less importance on academic pursuits. This hypothesis was supported in our analysis. Instrumentality seems to serve as a buffer against grounding one's self-worth in success at school. This could be due to the stereotyping of school and learning as feminine (Jackson & Dempster, 2009; Heyder & Kessels, 2013), which may discourage students displaying traits traditionally associated with masculinity from attaching self-esteem to this domain. If successful academic performance is not as central to one's self-esteem, failure is not as threatening to a positive self-image and fear of failure should therefore be lower, as has been demonstrated in our data. Additionally, the high degree of independence, which is part and parcel of the construct of instrumentality itself, implies that also the self-esteem of persons scoring high on instrumentality should be less contingent on external sources such as academic success.

In line with previous findings (Engeser, 2005; Steinmayr & Spinath, 2008), male adolescents scored higher than female adolescents on hope for success, while female students displayed higher scores on fear of failure. Findings also illustrated that instrumentality was associated with lower fear of failure and higher hope for success, which suggests that

adolescent girls and boys may benefit from a more instrumental self-concept. These results are in line with past research with adults (Orlofsky & Stake, 1981), in which a masculine gender role self-concept was related to an advantageous achievement motivation (Elmen, 1991; Henschen et al., 1982; Major, 1979; Strage, 1997). Our study fills an important gap in the literature, as we are able to provide insight into the relation between gender role self-concepts and achievement motivation in adolescents, an area in which modern studies are lacking.

Our findings also relate to research on motivational constructs which are more domain-specific and are often studied within the expectancy-value model (Eccles & Wigfield, 2020). Academic self-concepts in different domains, commonly used as a proxy for expectation of success, have been studied in relation to students' gender-role self-concept. Some studies found that the endorsement of traits typically associated with femininity relate to higher verbal self-concepts (McGeown & Warhurst, 2020; Pajares & Valiante, 2001), while scores on masculinity display a weaker positive relation with math self-concepts (Wolter & Hannover, 2016; McGeown & Warhurst, 2020). Our study extends these findings by focusing domain-independent motivational constructs. As instrumentality was related to higher hope for success and lower fear of failure, it could follow that instrumentality may relate to greater expectations of success in general (Vollmer, 1986).

Differential Relations by Gender

As predicted, we found a stronger relation between female adolescents' instrumentality and their global self-esteem compared to male adolescents. Moreover, the indirect relation between instrumentality and fear of failure through self-esteem was only significant for female adolescents. Based on the idea of the prescriptive function of gender stereotypes (Eagly, 1983; 1987; Eagly & Karau, 2002), we hypothesized that a student's ascription of counter-stereotypical traits will, in fact, be more predictive for other

psychological outcomes than the ascription of traits that are perceived as typical of one's own gender. It follows that when women or girls endorse traits traditionally associated with masculinity, these should be traits they truly feel reflect their self-image and thus be able to predict their other characteristics, too. In contrast, men or boys endorsing traits associated with masculinity might, in some cases, simply be describing themselves in line with the male gender stereotype, without actually perceiving themselves the way they indicated on measurements. Our findings are in line with research utilizing complex gender role profiles (Lyng, 2009) showing that the most adaptive behavior patterns were exhibited by adolescents who resisted traditional gender norm ideals (Yu et al., 2020) and that gender typicality relates to academic achievement differently for adolescent girls and boys (Yavorsky & Buchmann, 2019).

Limitations and Future Research Directions

Like many psychological investigations, our data are cross-sectional and thus do not allow us to test causal links between the variables laid out above (Kline, 2016; Maxwell & Cole, 2007). Path models are formulated based on assumptions about the directionality of relations, and these assumptions themselves are not directly tested but rely on characteristics of the dataset (e.g., cross-sectional vs. longitudinal data). While some paths are likely only in one direction (e.g., relating to demographic variables such as gender), this is not the case for the constructs in our study. For example, there might exist a self-reinforcing mechanism over time such that self-esteem leads to lower fear of failure, while having lower fear of failure might lead one to ascribe more instrumental traits to oneself when asked (e.g., proud, powerful, fearless). The directionality of the relations can therefore not be conclusively determined based on the results in our dataset (Kline, 2016; Maxwell & Cole, 2007). Future research should examine the question of gender-differential effects of instrumentality on self-esteem and motivational constructs in longitudinal data.

Secondly, the sample was drawn from a large city in Germany, which, while diverse, may not be reflective of all cities within the nation. Perhaps, more importantly, the sample recruited for this study stems from the "Gymnasium," an academic track which only represents one of several types of academic environment. While over 40% of students attend this track (Federal Statistical Office Germany, 2020), these students are typically in higher socioeconomic classes than students from outside this track (Reiss et al., 2016). Generalizing to these other academic tracks may be difficult, as these students may display different patterns of achievement motivation, global self-esteem, and academic contingent self-esteem. Interestingly, gender role self-concept is theorized to vary according to socioeconomic status (SES), with qualitative studies showing higher SES parents encouraging more instrumental traits in their daughters (Friedman, 2013) and lower SES intertwined with more traditional masculinity in male adolescents (Morris, 2008). A recent quantitative study, however, found no relation between SES and gender typicality (Yavorsky & Buchmann, 2019). These results highlight the complicated relation between socioeconomic status and gender role self-concept, which we cannot speak to on an empirical level, given that we were not able to investigate socioeconomic status in this study.

Moreover, it should be kept in mind that the χ^2-test was significant for the SEM, indicating that the observed variance-covariance matrix was not reproduced by the model. Though the reliance on (only) χ^2-tests is somewhat debated (e.g., Barrett, 2007; Markland, 2007) and both absolute and relative fit indices imply good fit, coefficients should be interpreted with care. A replication of this model in a different data set could provide further support for differential relations of gender role self-concept for male and female adolescents. In addition, replications could address potential sources for a lack of fit in the present model. For example, our model may have omitted a relevant variable such as social class that could predict endogenous variables in our model or moderate the relations between those variables

already included in the model. Thus, future research should examine whether additional, theoretically relevant variables could influence self-esteem and/or motivational constructs in school children.

We also wish to stress the importance of developing measures and designs that allow researchers to better detect whether response patterns regarding gender-typed attributes might reveal participants' actual traits and/or whether they reflect participants' knowledge of and willingness to conform to gender stereotypes (Tobin et al., 2010). The present findings indicate that especially counter-stereotypical descriptions might reveal the traits adolescents feel they possess, as responding counter-stereotypically represents a rejection of gender stereotypes. Contrastingly, responses in line with stereotypical descriptions might be a reaction to pressure from gender stereotypes and thus may not show the traits with which adolescents truly identify.

Further research in this area is needed in order to better understand what it actually means when individuals ascribe gender-typed attributes to themselves. Tobin and colleagues (2010) have pointed to the multiple dimensions of gender identity. Their gender self-socialization model (GSSM) explicates how *attribute self-perception* (this is what the instruments used in our study captured) has to be conceptualized as distinct from self-perceived gender typing (*gender identity*). The authors note that people might ascribe instrumental or expressive traits to themselves without considering these traits as masculine or feminine. It may be the case that self-ascriptions of characteristics considered gendered by others may not be categorized as masculine or feminine by the individual making the ascription (Tobin et al., 2010). The complex and personal beliefs surrounding gender and its associated traits may differ from adolescent to adolescent, and capturing these ideas in a few items may fail to depict certain aspects of an individual's concept of gender. While the overarching categories of masculinity and femininity are relevant terms for this paper, given

as they resonate with gender stereotypes in public consciousness (Rudman & Glick, 2008), it is not guaranteed that the traits resonating with adolescents are gendered in their minds.

However, since people on average seem to possess more, deeper, and more fine-grained knowledge about own-gender traits and behavior, the knowledge about own-gender stereotypes might also be larger than the knowledge about other-sex stereotypes (Martin et al., 2002). Thus, regardless of whether a person considers these traits explicitly masculine or feminine, own-gender traits should be more familiar and, since being more familiar leads to more liking and approval (Bornstein, 1989; Zajonc, 1968), being more of an expert in own-gender traits will make the endorsement of these more likely as compared to the endorsement of other-gender traits. Our study adds to our understanding that the endorsement of gendered traits might have different significance for male and female individuals; as for self-esteem, the endorsement of gender-role incongruent traits was more predictive than the endorsement of gender-role congruent traits.

As gender stereotypes are not static and shift across time, women are being ascribed more competence in recent decades, and social attitudes shift towards egalitarianism, we may see stereotype contents changing to reflect these changes, too (Eagly et al., 2020). Of course, as masculinity (as well as femininity) is a complex, fluid construct (Eagly et al., 2020) that can comprise a variety of traits (Berger & Krahé, 2013), not all types or aspects of masculinity may be helpful in academic contexts. One popular strain of research has investigated gender differences in school through the lens of a specific type of masculinity, for example, “laddish,” or “negative,” or “traditional” masculinities, finding such masculinities to be problematic in academia (Glock & Kleen, 2017; Heyder & Kessels, 2013, 2015; Huyge et al., 2015; Jackson, 2002; Jackson & Dempster, 2009; Jones & Myhill, 2004; Kessels & Steinmayr, 2013; Yavorsky et al., 2015). In contrast, the present study examined

instrumentality as one dimension of masculinity and found it to be associated with positive outcomes.

However, some positive aspects associated with instrumentality, such as higher global self-esteem recorded in multiple studies including the present paper, could also have negative effects in an academic context. Seeking help was observed to be less likely in individuals scoring high on self-esteem (Tessler & Schwartz, 1972), highlighting the delicate balance between beneficial and detrimental outcomes of certain traits. Gendered traits have a complex relation with academic performance and while this paper suggests that some aspects of masculinity, such as pride and fearlessness, may be helpful in the school context, this finding should not be extrapolated to other types of masculinities.

Practice Implications

The benefits of resisting strict gender roles has been exalted since the 70s (Bem, 1974) yet has not taken hold in the discourse on feminization of school until recently (Yu et al., 2020). A more careful consideration of the multifaceted impact of masculinities in school would benefit male and female students, teachers, and parents. Adopting a new perspective that emphasizes the benefits of both expressive and instrumental self-concepts in school contexts (cf. Verniers et al., 2016; Yu et al., 2020) could present an advantage to adolescents. Challenging the view that schools are a feminine space may not only disrupt the alienation felt by boys in school (Hadjar & Lupatsch, 2010) but also the seemingly false promise school appears to make to female students: At school, behaviors that are rather typically feminine are reinforced through good grades, while in working life, success is related to quite different traits (Steinmayr & Kessels, 2017). Still, simply suggesting that female students should be encouraged to develop their instrumentality would be simplistic, as the possible impact of instrumentality has to be put into context. Studies taking the intersection of gender and

socioeconomic status (Friedman, 2013) or ethnicity (George, 2015) into account have shown that traits associated with masculinity have the potential to interact with other identities.

Given the growing disillusionment with binary models of gender, it would be most preferable if the perception that certain traits reflect something masculine or feminine that applies exclusively to men or women, respectively, were abandoned (Connell & Messerschmidt, 2005). This would enable people of all genders to describe themselves with any trait they deem appropriate.

Conclusion

The present study investigated the association between instrumentality and achievement motivation in schools both directly and indirectly via multiple self-esteem constructs. While a certain type of masculinity has been portrayed as problematic in the school context, our findings point to positive associations between a different type of masculinity—instrumentality—and adaptive patterns of motivation and self-esteem in adolescents. The fact that one of these associations and an indirect effect were stronger in female than in male adolescents underlines that counter-stereotypical self-perceptions might be more powerful predictors than self-perceptions in line with gender stereotypes.

References

Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology*, *85*(4), 768–776. https://doi.org/10.1037/0022-3514.85.4.768

Abele, A. E., & Candova, A. (2007). Prädiktoren des Belastungserlebens im Lehrerberuf: Befunde einer 4-jährigen Längsschnittstudie [Predicting teachers' stress experience: Findings from a 4-year longitudinal study]. *Zeitschrift für Pädagogische Psychologie*, *21*(2), 107–118. https://doi.org/10.1024/1010-0652.21.2.107

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness-communion with warmth and morality. *Frontiers in Psychology*, *7*, 1810. https://doi.org/10.3389/fpsyg.2016.01810

Abele, A. E., Rupprecht, T., & Wojciszke, B. (2008). The influence of success and failure experiences on agency. *European Journal of Social Psychology*, *38*(3), 436–448. https://doi.org/10.1002/ejsp.454

Abele, A. E., & Wojciszke, B. (2014). Communal and agentic content in social cognition: A dual perspective model. *Advances in Experimental Social Psychology*, *50*, 195–255. https://doi.org/10.1016/B978-0-12-800284-1.00004-7

Aktop, A. & Erman, K. A. (2006). Relationship between achievement motivation, trait anxiety and self-esteem. *Biology of Sport*, *23*(2), 127–141.

Allgood-Merten, B., & Stockard, J. (1991). Sex role identity and self-esteem: A comparison of children and adolescents. *Sex Roles*, *25*(3–4), 129–139. https://doi.org/10.1007/BF00289850

American Psychological Association. (2020). *Publication Manual of the American Psychological Association* (7th ed.). American Psychological Association.

American Psychological Association Boys and Men Guidelines Group. (2018). *APA guidelines for psychological practice with boys and men.* http://www.apa.org/about/policy/psychological-practice-boys-men-guidelines.pdf

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion.* Rand McNally.

Barrett, P. (2007). Structural equation modelling: Adjudging model fit. *Personality and Individual Differences*, *42*(5), 815–824. https://doi.org/10.1016/j.paid.2006.09.018

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155–162. https://doi.org/10.1037/h0036215

Bem, S. L. (1981). *Bem Sex-Role Inventory: A professional manual*. Palo Alto: Consulting Psychologist Press.

Berger, A. & Krahé, B. (2013). Negative attributes are gendered too: Conceptualizing and measuring positive and negative facets of sex-role identity. *European Journal of Social*

Psychology, *43*(6), 516–531. https://doi.org/10.1002/ejsp.1970

Bleidorn, W., Arslan, R. C., Denissen, J. J. A., Rentfrow, P. J., Gebauer, J. E., Potter, J., & Gosling, S. D. (2016). Age and gender differences in self-esteem: A cross-cultural window. *Journal of Personality and Social Psychology*, *111*(3), 396–410. https://doi.org/10.1037/pspp0000078.supp

Block, J., & Robins, R. W. (1993). A longitudinal study of consistency and change in self-esteem from early adolescence to early adulthood. *Child Development*, *64*(3), 909–923. https://doi.org/10.2307/1131226

Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, *106*(2), 265–289. https://doi.org/10.1037/0033-2909.106.2.265

Buckley, T. R., & Carter, R. T. (2005). Black adolescent girls: Do gender role and racial identity impact their self-esteem? *Sex Roles*, *53*(9–10), 647–661. https://doi.org/10.1007/s11199-005-7731-6

Burwell, R. A., & Shirk, S. R. (2006). Self processes in adolescent depression: The role of self-worth contingencies. *Journal of Research on Adolescence*, *16*(3), 479–490. https://doi.org/10.1111/j.1532-7795.2006.00503.x

Butler, R. (2019). Gender, motivation, and society: New and continuing challenges. In E. N. Gonida & M. S. Lemos (Eds.), *Motivation in education at a time of global change: Theory, research, and implications for practice* (Vol. 20, pp. 129–149). Bingley, UK: CPI Group. https://doi.org/10.1108/S0749-742320190000020007

Cate, R., & Sugawara, A. I. (1986). Sex role orientation and dimensions of self-esteem among middle adolescents. *Sex Roles*, *15*(3–4), 145–158. https://doi.org/10.1007/BF00287480

Choi, H., Kim, J. H., Hwang, M. H., & Heppner, M. J. (2010). Self-esteem as a mediator between instrumentality, gender role conflict and depression in male Korean high school students. *Sex Roles*, *63*(5–6), 361–372. https://doi.org/10.1007/s11199-010-9801-7

Connell, R. W. & Messerschmidt, J. W. (2005). Hegemonic masculinity: Rethinking the concept. *Gender and Society*, *19*(6), 829–859. https://doi.org/10.1177/0891243205278639

Crocker, J., Karpinski, A., Quinn, D. M., & Chase, S. K. (2003). When grades determine self-worth: Consequences of contingent self-worth for male and female engineering and psychology majors. *Journal of Personality and Social Psychology*, *85*(3), 507–516. https://doi.org/10.1037/0022-3514.85.3.507

Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A. (2003). Contingencies of self-worth in college students: Theory and measurement. *Journal of Personality and Social Psychology*, *85*(3), 894–908. https://doi.org/10.1037/0022-3514.85.5.894

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*,

130(3), 392–414. https://doi.org/10.1037/0033-2909.130.3.392

Eagly, A. H. (1983). Gender and social influence: A social psychological analysis. *American Psychologist*, *38*(9), 971–981. https://doi.org/10.1037/0003-066X.38.9.971

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, N.J.: Erlbaum.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598. https://doi.org/10.1037//0033-295X.109.3.573

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of U. S. public opinion polls from 1946 to 2018. *American Psychologist*, *75*(3), 301–315. https://doi.org/10.1037/amp0000494

Eccles-Parsons, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75-146). San Francisco, CA: W. H. Freeman

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*(4), 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Egan, S. K., & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*, *37*(4), 451–463. https://doi.org/10.1037/0012-1649.37.4.451

Elmen, J. (1991). Achievement orientation in early adolescence: Developmental patterns and social correlates. *The Journal of Early Adolescence*, *11*(1), 125–151. https://doi.org/10.1177/0272431691111006

Engeser, S. (2005). *Messung des expliziten Leistungsmotivs: Kurzform der Achievement Motives Scale [Measuring the explicit achievement motive: A short version of the Achievement Motives Scale]*. Universität Trier.

Fairlamb, S. (2020). We need to talk about self-esteem: The effect of contingent self-worth on student achievement and well-being. *Scholarship of Teaching and Learning in Psychology*. https://doi.org/10.1037/stl0000205

Federal Statistical Office Germany. (2020). *Bildung und Kultur. Allgemeinbildende Schulen [Education and culture. General education]*. Retrieved from: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Publikationen/Downloads-Schulen/allgemeinbildende-schulen-2110100197004.pdf?__blob=publicationFile

Friedman, H. L. (2013). Tiger girls on the soccer field. *Contexts, 12*(4), 30-35. https://doi.org/10.1177/1536504213511213

Fu, C., & Wang, M. (2020). Parental corporal punishment and girls' self-esteem: The

moderating effects of girls' agency and communion in China. *Sex Roles*, 1–12. https://doi.org/10.1007/s11199-020-01174-6

Gebauer, J. E., Wagner, J., Sedikides, C., & Neberich, W. (2013). Agency-communion and self-esteem relations are moderated by culture, religiosity, age, and sex: Evidence for the "Self-Centrality Breeds Self-Enhancement" principle. *Journal of Personality*, *81*(3), 261–275. https://doi.org/10.1111/j.1467-6494.2012.00807.x

George, J. A. (2015). Stereotype and school pushout: Race, gender, and discipline disparities. *Arkansas Law Review*, *68*(101), 101–129.

Gjesme, T. (1983). Motivation to approach success (TS) and motivation to avoid failure (TF) at school. *Scandinavian Journal of Educational Research*, *27*(3), 145–164. https://doi.org/10.1080/0031383830270302

Glock, S., & Kleen, H. (2017). Gender and student misbehavior: Evidence from implicit and explicit measures. *Teaching and Teacher Education*, *67*, 93–103. https://doi.org/10.1016/j.tate.2017.05.015

Hadjar, A. & Lupatsch, J. (2010). Der Schul(miss)erfolg der Jungen: Die Bedeutung von sozialen Ressourcen, Schulentfremdung und Geschlechterrollen [The lower educational success of boys: The impact of social resources, school alienation and gender role patterns]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *62*(4), 599–622. https://doi.org/10.1007/s11577-010-0116-z

Henschen, K. P., Edwards, S. W., & Mathinos, L. (1982). Achievement motivation and sex-role orientation of high school female track and field athletes versus nonathletes. *Perceptual and Motor Skills, 55*(1), 183–187. https://doi.org/10.2466/pms.1982.55.1.183

Herrmann, J., Koeppen, K., & Kessels, U. (2019). Do girls take school too seriously? Investigating gender differences in school burnout from a self-worth perspective. *Learning and Individual Differences*, *69*, 150–161. https://doi.org/10.1016/j.lindif.2018.11.011

Heyder, A., & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles*, *69(11–12)*, 605–617. https://doi.org/10.1007/s11199-013-0309-9

Heyder, A., & Kessels, U. (2015). Do teachers equate male and masculine with lower academic engagement? How students' gender enactment triggers gender stereotypes at school. *Social Psychology of Education*, *18*(3), 467–485. https://doi.org/10.1007/s11218-015-9303-0

Heyder, A., & Kessels, U. (2017). Boys don't work? On the psychological benefits of showing low effort in high school. *Sex Roles*, *77*(1–2), 72–85. https://doi.org/10.1007/s11199-016-0683-1

Heyder, A., Van Hek, M., & Van Houtte, M. (2021). When gender stereotypes get male adolescents into trouble: A longitudinal study on gender conformity pressure as a

predictor of school misconduct. *Sex Roles*, *84*, 61–75. https://doi.org/10.1007/s11199-020-01147-9

Hirokawa, K., & Dohi, I. (2007). Agency and communion related to mental health in Japanese young adults. *Sex Roles*, *56*(7–8), 517–524. http://doi.org/10.1007/s11199-007-9190-8

Hirschy, A. J., & Morris, J. R. (2002). Individual differences in attributional style: The relational influence of self-efficacy, self-esteem, and sex role identity. *Personality and Individual Differences*, *32*(2), 183–196. https://doi.org/10.1016/S0191-8869(01)00003-4

Huyge, E., Van Maele, D. & Van Houtte, M. (2015). Does students' machismo fit in school? Clarifying the implications of traditional gender role ideology for school belonging. *Gender and Education, 27*(1), 1–18. https://doi.org/10.1080/09540253.2014.972921

Jackson, C. (2002). 'Laddishness' as a self-worth protection strategy. *Gender and Education*, *14*(1), 37–50. https://doi.org/10.1080/09540250120098870

Jackson, C. & Dempster, S. (2009). 'I sat back on my computer … with a bottle of whisky next to me': Constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *Journal of Gender Studies*, *18*(4), 341–356. https://doi.org/10.1080/09589230903260019

Jewell, J. A., & Brown, C. S. (2014). Relations among gender typicality, peer relations, and mental health during early adolescence. *Social Development*, *23*(1), 137–156. https://doi.org/10.1111/sode.12042

Jones, S., & Myhill, D. (2004). 'Troublesome boys' and 'compliant girls': Gender identity and perceptions of achievement and underachievement. *British Journal of Sociology of Education*, *25*(5), 547–561. https://doi.org/10.1080/0142569042000252044

Jöstl, G., Bergsmann, E., Lüftenegger, M., Schober, B. & Spiel, C. (2012). When will they blow my cover? The impostor phenomenon among Austrian doctoral students. *Journal of Psychology*, *220*(2), 109–120. https://doi.org/10.1027/2151-2604/a000102

Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, *20*(3), 309–323. https://doi.org/10.1007/BF03173559

Kessels, U. & Heyder, A. (2020). Not stupid, but lazy: Psychological benefits of disruptive classroom behavior from an attributional perspective. *Social Psychology of Education*, *23*, 583–613. https://doi.org/10.1007/s11218-020-09550-6

Kessels, U., Heyder, A., Latsch, M., & Hannover, B. (2014). How gender differences in academic engagement relate to students' gender identity. *Educational Research*, *56*(2), 220–229. https://doi.org/10.1080/00131881.2014.898916

Kessels, U., & Steinmayr, R. (2013). Macho-man in school: Toward the role of gender role self-concepts and help seeking in school performance. *Learning and Individual Differences*, *23*(1), 234–240. https://doi.org/10.1016/j.lindif.2012.09.013

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin*, *125*(4), 470–500. https://doi.org/10.1037/0033-2909.125.4.470

Knox, M., Funk, J., Elliott, R., & Bush, E. G. (1998). Adolescents' possible selves and their relationship to global self-esteem. *Sex Roles*, *39(1–2)*, 61–77.

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin*, *137*(4), 616–642. https://doi.org/10.1037/a0023557

Korlat, S., Forst, N., Schultes, M., Schober, M. T., Spiel, B. & Kollmayer, M. (2021). Gender role identity and gender intensification: Agency and communion in adolescents' spontaneous self-descriptions. *European Journal of Developmental Psychology*, 1–25. https://doi.org/10.1080/17405629.2020.1865143

Krahé, B., Berger, A., & Möller, I. (2007). Development and validation of an inventory for measuring gender role self-concept in adolescence [Entwicklung und Validierung eines Inventars zur Erfassung der Geschlechtsrollenorientierung im Jugendalter]. *Zeitschrift für Sozialpsychologie*, *38*(3), 195–208. https://doi.org/10.1024/0044-3514.38.3.195

Lyng, S. T. (2009). Is there more to "Antischoolishness" than masculinity? On multiple student styles, gender, and educational self-exclusion in secondary school. *Men and Masculinities*, *11*(4), 462–487. https://doi.org/10.1177/1097184X06298780

Major, B. (1979). Sex-role orientation and fear of success: Clarifying an unclear relationship. *Sex Roles*, *5*(1), 63–70. https://doi.org/10.1007/BF00289346

Mandal, E., & Moroń, M. (2019). Contingencies of self-worth and global self-esteem among college women: The role of masculine and feminine traits endorsement. *Social Psychological Bulletin*, *14*(3), Article e33507. https://doi.org/10.32872/spb.v14i1.33507

Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences*, *42*(5), 851–858. https://doi.org/10.1016/j.paid.2006.09.023

Martin, A. J., & Marsh, H. W. (2003). Fear of failure: Friend or foe? *Australian Psychologist*, *38*(1), 31–38. https://doi.org/10.1080/00050060310001706997

Martin, C. L., Ruble, D. N., & Szkrybalo, J. (2002). Cognitive theories of early gender development. *Psychological Bulletin*, *128*(6), 903–933. https://doi.org/10.1037/0033-2909.128.6.903

Matud, M. P., López-Curbelo, M., & Fortes, D. (2019). Gender and psychological well-being. *International Journal of Environmental Research and Public Health*, *16*(9), 1–11.

https://doi.org/10.3390/ijerph16193531

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12*(1), 23-44. https://doi.org/10.1037/1082-989X.12.1.23

McClelland, D. C., Atkinson, J., Clark, R., & Lowell, E. (1953). *The Achievement Motive.* New York: NY: Appleton-Century-Croft.

McGeown, S. P. & Warhurst, A. (2020). Sex differences in education: exploring children's gender identity. *Educational Psychology*, *40*(1), 103–119. https://doi.org/10.1080/01443410.2019.1640349

Moore, J. S. B., & Smith, M. (2018). Children's levels of contingent self-esteem and social and emotional outcomes. *Educational Psychology in Practice*, *34*(2), 113–130. https://doi.org/10.1080/02667363.2017.1411786

Morris, E. W. (2008). "Rednecks," "rutters," and 'rithmetic: Social class, masculinity, and schooling in a rural context. *Gender and Society*, *22*(6), 728–751. https://doi.org/10.1177/0891243208325163

Myers, D., Abell, J., Kolstad, A., & Sani, F. (2010). *Social psychology*. London: McGraw Hill.

Orlofsky, J. L., & Stake, J. E. (1981). Psychological masculinity and femininity: Relationship to striving and self-concept in the achievement and interpersonal domains. *Psychology of Women Quarterly*, *6*(2), 218–233. https://doi.org/10.1111/j.1471-6402.1981.tb00409.x

Otterpohl, N., Steffgen, S. T., Stiensmeier-Pelster, J., Brenning, K., & Soenens, B. (2020). The intergenerational continuity of parental conditional regard and its role in mothers' and adolescents' contingent self-esteem and depressive symptoms. *Social Development*, *29*(1), 143–158. https://doi.org/10.1111/sode.12391

Pajares, F. & Valiante, G. (2001). Gender differences in writing motivation and achievement of middle school students: A function of gender orientation? *Contemporary Educational Psychology*, *26*(3), 366–381. https://doi.org/10.1006/ceps.2000.1069

Piché, C., & Plante, C. (1991). Perceived masculinity, femininity and androgyny among primary school boys. *European Journal of Psychology of Education*, *6*(4), 423–435. https://doi.org/10.1007/BF03172775

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria. https://www.r-project.org/

Radziwiłłowicz, W. & Macias, M. (2014). Self-esteem and achievement motivation level in overweight and obese adolescents. *Health Psychology Report*, *2*(2), 132–143. https://doi.org/10.5114/hpr.2014.43920

Reiss, K., Sälzer, C., Schiepe-Tiska, A., Klieme, E., & Köller, O. (2016). *PISA 2015. Eine Studie zwischen Kontinuität und Innovation [PISA 2015. A study between continuity*

and innovation]. Münster; New York: Waxmann.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5--12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rudman, L. A. & Glick, P. (2008). *The social psychology of gender: How power and intimacy shape gender relations.* New York, US: Guilford Press.

Schöne, C., & Stiensmeier-Pelster, J. (2016). *Selbstwertinventar für Kinder und Jugendliche – SEKJ [Self-esteem inventory for children and adolescents]*. Göttingen: Hogrefe. https://doi.org/10.13140/RG.2.1.3583.5607

Schöne, C., Tandler, S. S., & Stiensmeier-Pelster, J. (2015). Contingent self-esteem and vulnerability to depression: Academic contingent self-esteem predicts depressive symptoms in students. *Frontiers in Psychology*, *6*, 1573. https://doi.org/10.3389/fpsyg.2015.01573

Severiens, S. & Dam, G. T. (1998). A multilevel meta-analysis of gender differences in learning orientations. *British Journal of Educational Psychology*, *68*(4), 595–608. https://doi.org/10.1111/j.2044-8279.1998.tb01315.x

Sharpe, M. J., Heppner, P. P., & Dixon, W. A. (1995). Gender role conflict, instrumentality, expressiveness, and well-being in adult men. *Sex Roles(1–2)*, *33*, 1–18. https://doi.org/10.1007/BF01547932

Smiler, A. P. (2004). Thirty years after the discovery of gender: Psychological concepts and measures of masculinity. *Sex Roles*, *50*(1–2), 15–26. https://doi.org/10.1023/B:SERS.0000011069.02279.4c

Smiler, A. P. (2006). Living the image: A quantitative approach to delineating masculinities. *Sex Roles*, *55*(9–10), 621–632. https://doi.org/10.1007/s11199-006-9118-8

Spence, J. T. & Helmreich, R. L. (1980). Masculine instrumentality and feminine expressiveness: Their relationships with sex role attitudes and behaviors. *Psychology of Women Quarterly*, *5*(2), 147–163. https://doi.org/10.1111/j.1471-6402.1980.tb00951.x

Spence, J. T., Helmreich, R. L., & Stapp, J. (1974). *The personal attributes questionnaire: A measure of sex role stereotypes and masculinity-femininity*. University of Texas.

Stein, J. A., Newcomb, M. D., & Bentler, P. M. (1992). The effect of agency and communality on self-esteem: Gender differences in longitudinal data. *Sex Roles*, *26*(11–12), 465–483. https://doi.org/10.1007/BF00289869

Steinmayr, R., & Kessels, U. (2017). Good at school = successful on the job? Explaining gender differences in scholastic and vocational success. *Personality and Individual Differences*, *105*, 107–115. https://doi.org/10.1016/j.paid.2016.09.032

Steinmayr, R. & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, *22*, 185–209. https://doi.org/10.1002/per

Steinmayr, R., & Spinath, B. (2009a). What explains boys' stronger confidence in their intelligence? *Sex Roles*, *61*(9–10), 736–749. https://doi.org/10.1007/s11199-009-9675-8

Steinmayr, R., & Spinath, B. (2009b). The importance of motivation as a predictor of school achievement. *Learning and Individual Differences*, *19*(1), 80–90. https://doi.org/10.1016/j.lindif.2008.05.004

Steinmayr, R., Weidinger, A. F., Schwinger, M., & Spinath, B. (2019). The importance of students' motivation for their academic achievement-replicating and extending previous findings. *Frontiers in Psychology*, *10*, 1730. https://doi.org/10.3389/fpsyg.2019.01730

Strage, A. (1997). Agency, communion, and achievement motivation. *Adolescence*, *32*(126), 299–312.

Tessler, R. C. & Schwartz, S. H. (1972). Help seeking, self-esteem, and achievement motivation: An attributional analysis. *Journal of Personality and Social Psychology, 21*(3)*, 318–326*. https://doi.org/10.1037/h0032321

Thompson, E. H., Pleck, J. H. & Ferrera, D. L. (1992). Men and masculinities: Scales for masculinity ideology and masculinity-related constructs. *Sex Roles*, *27*(11–12), 573–607. https://doi.org/10.1007/BF02651094

Tobin, D. D., Menon, M., Menon, M., Spatta, B. C., Hodges, E. V., & Perry, D. G. (2010). The intrapsychics of gender: a model of self-socialization. *Psychological Review*, *117*(2), 601–622. https://doi.org/10.1037/a0018936

Van der Kaap-Deeder, J., Wouters, S., Verschueren, K., Briers, V., Deeren, B., & Vansteenkiste, M. (2016). The pursuit of self-esteem and its motivational implications. *Psychologica Belgica*, *56*(3), 143–168. https://doi.org/10.5334/pb.277

Van Hek, M., Kraaykamp, G., & Pelzer, B. (2018). Do schools affect girls' and boys' reading performance differently? A multilevel study on the gendered effects of school resources and school practices. *School Effectiveness and School Improvement*, *29*(1), 1–21. https://doi.org/10.1080/09243453.2017.1382540

Vantieghem, W., & Van Houtte, M. (2018). Differences in study motivation within and between genders: An examination by gender typicality among early adolescents. *Youth and Society*, *50*(3), 377–404. https://doi.org/10.1177/0044118X15602268

Verniers, C., Martinot, D., & Dompnier, B. (2016). The feminization of school hypothesis called into question among junior and high school students. *The British Journal of Educational Psychology*, *86*(3), 369–381. https://doi.org/10.1111/bjep.12111

Vollmer, F. (1986). Why do men have higher expectancy than women? *Sex Roles*, *14*(7–8),

351–362. https://doi.org/10.1007/BF00288420

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174– 1204. https://doi.org/10.1037/a0036620

Whitley, B. E. (1984). Sex-role orientation and psychological well-being: Two meta-analyses. *Sex Roles*, *12*(1–2), 207–225. https://doi.org/10.1007/BF00288048

Whitley, B. E. (1988). Masculinity, femininity, and self-esteem: A multitrait-multimethod analysis. *Sex Roles*, *18*(7–8), 419–431. https://doi.org/10.1007/BF00288393

Whitley, B. E., & Gridley, B. E. (1993). Sex role prientation, self-esteem, and depression: A latent variables analysis. *Personality and Social Psychology Bulletin*, *19*(4), 363–369.

Wigfield, A., & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Wojciszke, B., Baryla, W., Parzuchowski, M., Szymkow, A., & Abele, A. E. (2011). Self-esteem is dominated by agentic over communal information. *European Journal of Social Psychology*, *41*(5), 617–627. https://doi.org/10.1002/ejsp.791

Wolfram, H. J., Mohr, G., & Borchert, J. (2009). Gender role self-concept, gender-role conflict, and well-being in male primary school teachers. *Sex Roles*, *60*(1–2), 114–127. https://doi.org/10.1007/s11199-008-9493-4

Wolter, I. B. & Hannover, B. (2016). Gender role self-concept at school start and its impact on academic self-concept and performance in mathematics and reading. *European Journal of Developmental Psychology*, *13*(6), 681–703. https://doi.org/10.1080/17405629.2016.1175343

Yavorsky, J. E., & Buchmann, C. (2019). Gender typicality and academic achievement among American high school students. *Sociological Science*, *6*(25), 661–683. https://doi.org/10.15195/v6.a25

Yavorsky, J. E., Buchmann, C. & Miles, A. (2015). High school boys, gender, and academic achievement: Does masculinity negatively impact boys' grade point averages? *Population Association of America*. http://paa2015.princeton.edu/uploads/152814

Yu, J., McLellan, R., & Winter, L. (2020). Which boys and which girls are falling behind? Linking adolescents' gender role profiles to motivation, engagement, and achievement. *Journal of Youth and Adolescence*, 1–17. https://doi.org/10.1007/s10964-020-01293-z

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monographs*, *9(2)*, 1–27. https://doi.org/10.1037/h0025848

Zuckerman, M., Li, C., & Hall, J. A. (2016). When men and women differ in self-esteem and when they don't: A meta-analysis. *Journal of Research in Personality*, *64*, 34–51.

Appendix

Table S1

Instrumentality, Expressiveness Scale Items (Kessels, 2005)

Instrumental	Expressive
furchtlos (fearless)	artig (good)
bereit, etwas zu riskieren (willing to take risks)	hilfsbereit (helpful)
kann Druck gut standhalten (stand up well under pressure)	kinderlieb (love children)
kraftvoll (forceful)	rücksichtsvoll (considerate)
mächtig (powerful)	sanft (gentle)
mutig (courageous)	vernünftig (sensible)
stark (strong)	verständnisvoll (understanding)
stolz (proud)	vorsichtig (careful)
unerschrocken (not excitable)	weichherzig (tender-hearted)
aktiv (active) *	bescheiden (modest)*
fälle leicht Entscheidungen (make decisions easily)*	gefühlsbetont (emotional)*
fühle mich überlegen (feel superior)*	nachgiebig (yielding)*
hartnäckig (persistent)*	ordentlich (neat)*
habe Führungseigenschaften (act as a leader)*	romantisch (romantic)*
zeige geschäftsmäßiges Verhalten (show business-like behavior)*	scheu (shy)*

Note. Items were presented in German (English translation in parentheses).

* Items were not included in the study based on the confirmatory factor analyses described in the methods section.

III. Discussion

Summary of Results

In three manuscripts, the present dissertation investigated gender, gender stereotypes, and gender role self-concepts from the perspective of educators, peers, and students in academic settings. I will briefly outline these works before discussing the overarching theme of gender and gendered characteristics, highlighting gender schema theory and the emerging topic of gender as a status characteristic, which is defined as an individual property that holds information on competence and social evaluations based on group membership (Berger et al., 1977). This discussion then outlines the limitations of this work and delineates opportunities for future research, such as embracing the role of parents and an intersectional approach. The implications of the present work, like the role of femininity and its status in school, and a question of whether society has become more gender schematic are outlined before the conclusion.

Study 1 (*Linda's Cars vs. Dominik's Dolls: How Do Pedagogical Educators in Training React to Children's Violations of Gender Stereotypes?*) investigated the gender stereotypes about 3-year-old children held by adults in Germany and whether pedagogical educators in training would show backlash towards fictional children described as behaving in gender stereotypical or nonstereotypical ways. The extensive first study revealed that gender stereotypes about children show some tendencies towards agentic and communal stereotypes but strongly focus on activities and play preferences. Regarding the attitudes of pedagogical educators in training, the findings of the experimental vignette study showed no backlash towards gender nonstereotypical children regarding "liking," a departure from the existing literature (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013). Interestingly, ratings of perceived competence, self-esteem, and creativity all show a similar pattern, with masculine girls receiving the highest ratings and feminine girls receiving the lowest ratings, while there were no significant differences between masculine and feminine

boys. Feminine boys received positive ratings for prosocial behavior, while gender stereotyped behavior, but not gender, related to internalizing and externalizing tendencies. Overall, this work shows a benefit of masculine and a detriment of feminine stereotyped behavior for girls, rather than a backlash of behaving in feminine stereotyped ways for boys. Previous work found boys receiving worse ratings from adults for displaying femininity (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013), and present work found rewards for girls showing masculinity, such empirical findings stressing the high social value of masculinity and the lower standing of femininity (Feinman, 1981).

The second study *(Nice, But Not Smart? Attributional Backlash from Displaying Prosocial Behavior in the Classroom)* investigated whether displaying prosocial behavior would lead to attributional backlash by peers, as mediated by expected teacher reactions. Participating 9th grade students read vignettes describing academically high-achieving students, with one stated to be prosocial and the other being nondescript, with target gender varied between participants. Participants indicated whether a teacher would react with praise and happiness towards the targets and the extent to which the targets received their good grades as a result of high ability and effort. Further ratings of intelligence, popularity, likability, gender typicality, and frequency of similar same-gender students were also ascertained. We hypothesized that the expected teacher reactions of praise and happiness would mediate the association between target behavior and attributions for academic achievement (following Kessels & Heyder, 2020), as predicted by literature on the paradox of praise (Graham & Chen, 2020; Meyer et al., 1979; Möller, 2005). However, prosocial behavior related to achievement attributions, without the expected teacher reaction mediating the association. Regarding the gender-related variables, participating students found prosocial targets to be more feminine but also more masculine than the nondescript student, the latter finding being unexpected. This work discussed the concerning finding that prosocial behavior,

seen as feminine (Eisenberg et al., 2006; Quenneville et al., 2022), relates to a denigrating attributional pattern. Prior research established that girls' achievements are attributed less to ability but rather to effort (Espinoza et al., 2014; Fennema et al., 1990; Räty et al., 2002; Tiedemann, 2000). Thus, the finding that feminine stereotyped, prosocial behavior relates to lower ability attributions is concerning and relevant particularly for girls, as they show greater prosociality, especially in adolescence (Bøe et al., 2016; Gerbino et al., 2018; Van der Graaff et al., 2018; S. X. Xiao et al., 2019). Beliefs about lower ability can also disincentivize girls from joining STEM, which is stereotyped as requiring "brilliance" (Chestnut et al., 2018; Kessels, 2015), more related to innate ability.

In the third study of this dissertation (*Instrumentality Gives Girls the Edge: Gender-Differential Relations Between Instrumentality, Achievement Motivation, and Self-Esteem),* we focus on students themselves. In this cross-sectional design, 9th grade students reported their gender role self-concept, global and academic contingent self-esteem, as well as their achievement motivation. Structural equation modeling revealed that instrumentality had a positive association with global self-esteem and a negative association with academic contingent self-esteem, thus overall showing advantageous relations between a masculine orientation and the self-esteem constructs. Similarly, an instrumental gender role self-concept related to advantageous achievement motivation, with greater hope for success and lower fear of failure in both boys and girls. A key aspect of this study was the focus on gender differential relations, which revealed that instrumentality related more strongly to global self-esteem in girls than boys. The indirect effect of instrumentality via global self-esteem to fear of failure was also significant in girls but not boys. In other words, having a stronger instrumental self-concept related to greater global self-esteem and through this increased sense of self-esteem, girls experienced an association of lower fear of failure. This work discusses these findings critically in light of the feminization of school hypothesis, which has presented a negative view of masculinity in the academic context. The results of this study,

however, highlight that instrumental gender role self-concepts can have beneficial relations in school, specifically for girls. A connection is also drawn to the gender self-socialization model (Tobin et al., 2010) and the question of whether boys may be indicating greater instrumental orientations due to the pressure to conform to gender expectations. Counter-stereotypical orientations could represent more authentic reflections of internal states and have been shown to be related to advantageous academic outcomes in past works (Yavorsky & Buchmann, 2019; Yu et al., 2020).

Discussing Gender and Gendered Characteristics

In the following section, I will discuss the overarching findings of the three works and examine the theoretical lens provided by gender schema theory. In the subsequent section, gender is examined as a status characteristic, and the findings are applied to the educational context.

Relation to Gender Schema Theory

According to gender schema theory, individuals incorporate and organize information into an existing cognitive structure relating to gender (Bem, 1981, 1983). In the present work, I used gender schema theory as an overarching framework to understand how one's gender schema informs both the response to others' behavior and intrapersonal processes. In each work, the gender schema held by participants is believed to have played a role in how they responded to measures. Overall, arguing that gender stereotypes and gender role self-concepts could have activated the gender schema and associated constructs, like competence, attributions, self-esteem, and achievement motivation, in the academic context. Our work shows that it is not just gender that would activate the schematic associations with other constructs but also stereotyped behavior and traits, the impact of which also interacted with gender in some of our work.

In manuscript 1, girls described as showing masculine stereotyped behaviors were rated as more competent, as having greater self-esteem, and as being more creative, all traits associated with masculinity over femininity (Martin & Slepian, 2021; Proudfoot et al., 2015; Whitley, 1983, 1988). The feminine boy received the greatest scores in terms of prosocial behavior, associated with femininity (Eisenberg et al., 2006; Piché & Plante, 1995; Quenneville et al., 2022). Masculine girls benefited from masculinity more than boys, for whom masculinity and femininity made no significant difference in terms of ratings of competence, creativity, and self-esteem; and feminine boys benefited from femininity more than girls, as feminine girls were not ascribed greater prosocial behavior, instead actually receiving worse ratings in terms of competence, creativity, and self-esteem.

In study 2, participants rated a prosocial target as having attained their academic achievements as resulting from more effort and less from ability than a nondescript target, and expected teacher reactions of praise and happiness did not mediate the relation. Prosocial behavior, which is stereotyped as feminine (Eisenberg et al., 2006; Quenneville et al., 2022), thus related to an attributional pattern more characteristic of girls (Assouline et al., 2021; Bornholt & Möller, 2003; Lohbeck et al., 2017; Mok et al., 2011). A stereotyped behavior activated associations in line with the gender stereotype, in this case prosocial, feminine behavior relating more to effort and less to ability. Interestingly, while the prosocial target was rated as more feminine than the nondescript target, in line with our hypotheses, participants also rated the prosocial target as more masculine, an unexpected finding. Could prosocial behavior be less related to femininity than previously thought (Hine, 2017; Ma, 2005)? As our results were dependent on warm, social behavior (Kervyn et al., 2012, 2016), other classroom behaviors could trigger these attributions, which may be a contributing factor to the female denigrating attributional patterns (Assouline et al., 2021; Bornholt & Möller, 2003; Lohbeck et al., 2017; Mok et al., 2011). Of course, it should be noted that not all feminine behavior is seen equally in the school context, and certain types of femininity do not

suit the academic context. In vignettes, exaggerated femininity, encompassing caring deeply about appearance or behaving in a suggestive manner, was associated with learning impeding behaviors by teachers (Heyder & Kessels, 2015), and in different work, students endorsing traditional femininity, capturing domesticity and a focus on personal appearance, showed poorer academic outcomes (Yu et al., 2020). The gender schema of femininity could be seen as containing complex associations, some of which suit the academic context, while other associations do not suit the educational setting.

Study 3, which investigated gender role self-concepts rather than gender stereotypes, showed that masculine orientations had beneficial associations with self-esteem and motivational constructs, especially for girls. This last study relates most closely to Bem's initial empirical investigation of gender schema theory, as in this early work participants provided an indication of their own gender role self-concept (sex typing) (Bem, 1981). In Bem's research, the following measures were cognitive, while ours related to personal traits, departing from Bem's work. Our work did reveal that an instrumental orientation related to responses more in line with masculinity on measures of self-esteem and achievement motivation (H. Choi et al., 2010; Orlofsky & Stake, 1981; Whitley, 1983, 1988; Yang & Gao, 2021). Past work has shown the consistent relation between masculinity and greater self-esteem (H. Choi et al., 2010; Whitley, 1983, 1988), possibly due to the fact that masculinity is socially valued (Orr & Ben-Eliahu, 1993). Masculinity is said to reflect more socially desirable traits (Pedhazur & Tetenbaum, 1979; Puglisi, 1980), although findings on this are mixed (Marsh et al., 1987) and may depend on whether traits are being evaluated as desirable for the self or for others (Abele & Wojciszke, 2007). Our work clearly supports this pattern, with masculine gender stereotypes and gender role self-concept relating to greater self-esteem.

As delineated above, gender schema theory can illuminate the connection between inter- and intrapersonal mechanisms related to gender and gendered characteristics. Rather than focusing on individual cognitive processes with gender schema theory, Bem later stated to use it to expose "an invisible cultural lens" (Bem, 1993, p. 131). While clear connections can be made from gender schema theory to the results in the present work, we did not assess the activation of gender schemata themselves. Thus, the connection should be seen as tentative, to be examined using more direct evidence. A possible issue that arises from the measurement of gender stereotypes and gender role self-concept is that people may not be relating the presented attributes to gender (Wood & Eagly, 2015). However, people may also be acutely aware of the gender stereotyped nature of the items and thus bias their responses to be more in line with their gender (cf. manuscript 3). After all, violations of gender stereotypes are punished (Eckes, 2008; Rudman & Glick, 2010), and congruity with one's gender and gender schema is beneficial for one's self-esteem (Bem, 1983; Canevello, 2020). The exact nature of the gender schema activation is thus not accessible with the data available to us.

The broad understanding of gender schema theory can be tremendously enlightening, as previous work stated: "we still organize and stereotype our world based on gender, and this gendered organization still has serious consequences, particularly for women and girls" (Starr & Zurbriggen, 2017, p. 575). Using associations of gender and gendered characteristics allowed us to formulate hypotheses about likability, competence, causal attributions, self-esteem, and motivation. Feminine stereotyped behavior related to feminine causal attributions and less favorable perceptions of competence, creativity, and self-esteem. Masculine stereotyped behavior related to beneficial, masculine patterns of ascriptions of competence, creativity, and self-esteem. Had these phenomena been investigated without gender as a lens, we could have missed a key finding: characteristics associated with femininity remain socially undervalued, even in school.

Gender as a Status Characteristic

In investigating how social inequalities are perpetuated in interactions, Berger and colleagues conceptualized the construct of status characteristics (Berger et al., 1972, 1977). Characteristics such as ethnicity, age, and occupation may be status characteristics, as they are believed to relay information (accurately or inaccurately) on competence in certain domains. Status characteristics must have at least two possible states, which are evaluated differently and, importantly, hierarchically, in terms of competence in society.

According to theorists (Berger et al., 1972; Rashotte & Webster, 2005), gender meets the criteria of a status characteristic. The evidence for men holding more status than women in society is most easily illustrated by regarding high status roles and determining which gender is more represented in these positions. Political (Hinojosa, 2021; IPU Parline, 2024), economic (Eurostat, 2023; Orbach, 2017), or academic leadership positions (Federal Statistical Office, 2022; Treviño et al., 2018) all show an overrepresentation of men – there are many more unnamed examples. But not only are men afforded higher status than women, masculinity is also valued over femininity.

The work by Berger et al. on status characteristics (1977) and by Alice Eagly on social role theory (1987) has more recently been integrated into one theoretical view (Best, 2009; Gerber, 2009). According to Gerber (2009), we associate masculine traits with higher status, as the ubiquity of men in high status positions leads to a confluence of viewing masculine traits as more fitting with leadership. Consequently, as women are viewed as lower status, the more expressive and communal traits associated with women become schematically associated with low status. Experimentally, individuals described as holding high status in society were rated as more agentic, while low-status individuals received higher ratings on communality across multiple studies (Conway et al., 1996), irrespective of gender (Conway & Vartanian, 2000, p. 183).

While men experience advantages in the vocational domain, girls experience advantageous outcomes in school. From a human capital theory viewpoint, however, this female educational achievement is most relevant when it can be converted into "income and status at the labour market" (Hadjar et al., 2014, p. 118). Seeing gender from a status characteristic approach could illuminate why we see female individuals outperforming male individuals academically but not vocationally.

We may first need to question how to view academic success. While it is true that girls outperform boys academically (O'Dea et al., 2018; Voyer & Voyer, 2014), this achievement is said to come at a cost (Herrmann et al., 2019). Not only do girls experience greater rates of school burnout and lower well-being (Herrmann et al., 2019), but girls experience worse attributional patterns (Assouline et al., 2021; Bornholt & Möller, 2003; Lohbeck et al., 2017; Mok et al., 2011), lower self-esteem (Kling et al., 1999; Twenge & Campbell, 2001; see Zeigler-Hill & Myers, 2012; Zuckerman et al., 2016), greater academic contingent self-esteem (Schöne et al., 2015; Schöne & Stiensmeier-Pelster, 2016; van der Kaap-Deeder et al., 2016), and more disadvantageous achievement motive (Steinmayr & Spinath, 2008; Wach et al., 2015) than boys. Furthermore, despite outperforming boys academically, the abilities of girls continue to be underestimated. On self-report measures, boys displayed greater confidence in their intelligence than girls in multiple domains, even when controlling for IQ. Parents of the participating adolescents showed a similar pattern, estimating boys' intelligence to be greater compared to that of girls (Steinmayr & Spinath, 2009). Parents and teachers also attributed successful performances to high ability in boys, but to greater effort in girls (Espinoza et al., 2014; Fennema et al., 1990; Räty et al., 2002; Tiedemann, 2000). Are girls truly profiting in this academic system if measures of well-being and self-esteem are not reflecting their high performance and if their achievements and abilities are not recognized and valued by others?

The disconnect between academic and working success may be due to differential reward categories in these domains. Past work, in which personality traits and their relation to academic and vocational success were assessed in a sample of students and adult professionals, showed how the domains of school and work are associated with different personality patterns (Steinmayr & Kessels, 2017). Success in school was related to agreeableness and low need for aggression, which did not predict professional success. Traits like intelligence, conscientiousness, and need for achievement were highlighted as relevant for both academic and professional success (Steinmayr & Kessels, 2017). This suggests that there is some overlap but also some differences between the personality traits required for success in vocational and academic environments. Other work found compliance, assertiveness, and effort/intelligence (the latter a single factor) to be rated as key for academic success, with only assertiveness judged to be more typical of boys, while compliance and effort/intelligence were associated more strongly with girls (Verniers et al., 2016).

These findings could still, however, be seen as neutral, even beneficial for girls – after all, displaying effort and compliance are seen very positively in school (Jones & Myhill, 2004; Matteucci & Gosling, 2004; Saidah et al., 2019). However, it is precisely the traits that serve girls in an academic setting that are detrimental to women's achievement in occupational contexts. "Few company executives, politicians, lawyers and so on would be described as compliant and conformist, though their [Personal Assistants] may well be!" (Myhill, 2002, p. 350). Even effort, held up as a core value and ingrained in academic institutions (Matteucci & Gosling, 2004), may contribute to this issue. Effort, as previously outlined, is seen as compensatory for low ability (Möller, 2005; Weiner & Kukla, 1970). But not only are effort and ability contrasted with one another, they are seen as hierarchical: "effort is virtuous, but it's better to have ability" (Nicholls, 1976, p. 306). Students and teachers alike report valuing effortless achievement (Heyder & Kessels, 2017; Jackson &

Nyström, 2014). In school, effortless achievement emerges as a signifier of "real ability" (Renold & Allan, 2006, p. 467), a more "authentic" achievement than one attained through more feminine, effortful ways (Jackson & Dempster, 2009, p. 348). This is a key path by which success of girls is undermined: girls' achievement is not authentic or natural but rather merely the result of "hard work" (Miller, J. 1996, p. 142). This belief can illuminate why some students, particularly boys, display counterproductive tendencies in school, such as procrastinating, making excuses, not working hard, or disrupting the classroom (Covington, 1984; Kessels & Heyder, 2020). This behavior not only affirms their masculinity (Heyder et al., 2021) but also their status, as they disengage from effortful learning, a low-status behavior.

Status, a factor we did not directly assess or investigate, can unite some of these patterns. Our present findings regarding masculinity and femininity, both as gender role self-concepts and gender stereotypes, show (1) more negative evaluations in terms of competence, creativity, and self-esteem for feminine girls, with more positive judgments of masculine girls on these dimensions by educators, (2) the disadvantageous attributional outcomes for feminine stereotyped, prosocial, classroom behavior by peers, and (3) advantages of a masculine gender role self-concept for self-esteem constructs and achievement motive for boys and girls. Currently, it seems girls are not able to benefit from their brief time at the top of the achievement hierarchy, as their successes in this domain are devalued as being merely due to effort (Espinoza et al., 2014; Fennema et al., 1990; Räty et al., 2002; Tiedemann, 2000), a lower status method of attaining achievement, while they concurrently experience worse mental health and well-being than their male peers (Herrmann et al., 2019; Kessels & Van Houtte, 2021; Schöne & Stiensmeier-Pelster, 2016; Stoeber & Rambow, 2007). Regarding school and scholastic achievement from a status characteristic approach could be illuminating for future investigations.

Limitations and Future Research Directions

Specific limitations of each work were discussed in the respective manuscripts. In this section, I explore more broad limitations of the present work below and present ideas for future research directions.

Limitations

Cross-sectional and experimental work of this nature is inherently susceptible to flaws and shortcomings. I focus specifically on difficulties in empirical educational research and overarching terminological difficulties.

Manuscript 1 and 2 utilized experimental vignettes. This gave us the opportunity to manipulate specific gender stereotyped behaviors and examine their impact on educators and students. Such experimental vignette designs have been critiqued in the past for lacking external validity and not necessarily reflecting real-world processes (Hainmueller et al., 2015, p. 2395). While naturalistic investigations of gender stereotyped behavior in kindergartens or prosocial behavior in classrooms would have been worthwhile, they would have required an immense amount of data collection, more heavily impacting the participants. This level of focus would have been difficult to justify, given that some of this work, specifically study 2, investigated an unconventional attributional focus, with little past work (Kessels & Heyder, 2020). Our studies provide some first indication of the occurrence of the hypothesized effects in these specific samples, laying a foundation for more comprehensive future work.

Two of the manuscripts outlined in this work (manuscript 2 and 3) were conducted with students attending the Gymnasium. A large number of students in Germany attend this highest academic track (Federal Statistical Office, 2024), yet it is not necessarily representative of all students in Germany. Students at the Gymnasium are more likely to be German (vs. having a migrant background) (Federal Statistical Office, 2024), female (vs. male) (Federal Statistical Office, 2023), and of higher SES (vs. lower) (Federal Statistical

Office, 2020). This may present an issue for generalization of our results to other tracks. Indeed, I focus on the question of an intersectional approach, specifically focusing on SES, in the upcoming section on “Future Research Directions”. However, since many students attend the Gymnasium, our findings are generalizable to a large number of students. I would encourage future work to explore gender role self-concepts and associations with self-esteem and motivational constructs in other academic tracks.

I now focus on some difficulty with terminology that affects not only the present work but the wider literature focusing on gender in psychology. Due to a multitude of definitions in the literature, authors have used the same words in referring to vastly different constructs (Egan & Perry, 2001; Tobin et al., 2010). To address this, I used terminology adapted from the GSSM (Tobin et al., 2010), considering gender identity (gender self-categorization) as distinct from gender stereotypes and as distinct from attribute self-perceptions (gender role self-concept). This, however, could be examined critically. The conceptualization of gender identity purely as gender self-categorization would be criticized by researchers as reductive (Tobin et al., 2010). By reducing gender identity to one dimension, namely self-reported gender, it is not capturing the variety of understandings associated with people’s gender identities. While I agree with this critique to an extent, I do note that assessing multiple dimensions invites a host of interpretations and introduces difficulties in quantifying group differences, which lie at the core of our empirical investigations. Others may question this measure for possibly actually capturing biological sex. This becomes slightly more complex in German, where the term “Geschlecht” may refer to both biological sex and gender (Dudenredaktion, n.d.). I would point out that we did not ask participants for any biological gender markers, but that it is highly likely, though not guaranteed, that biological sex and gender overlap in this measure.

We ascertained gender stereotypes via a method of collecting traits and behaviors, and asking individuals to rate the desirability and typicality of these items for male and female children in German society (Koenig, 2018; Rudman et al., 2012; Sullivan et al., 2022) (manuscript 1). Tobin and colleagues (2010) have stated that the above procedure would constitute the traditional approach, which is guided by the researcher rather than the individual. The authors critique the lack of individual perspective and oversimplification in the stereotypes that are endorsed (Tobin et al., 2010). While these critiques of stereotypes are valid, I would stress that a more individualized measurement of gender stereotypes is complex and perhaps overly so for investigations of group differences. The GSSM also recommends not capturing stereotypes across domains (those who believe boys should "excel in sports" may not necessarily believe boys should be adept at science [Tobin et al., 2010, p. 611]). Our findings suggest that feminine stereotyped play-activities, for example, related to feminine stereotyped outcomes, such as being seen as lower in competence (A. E. Martin & Slepian, 2021) and creativity (Proudfoot, 2015), suggesting that stereotypes may be linked across domains, but of course I recognize that this link should not be assumed.

We ascertained gender role self-concept with a measure by Kessels (2005), in which people indicated the fit of gender stereotyped items for themselves. This method, which is similar to the BSRI (Bem, 1974) and PAQ (Spence, 1991), has also been criticized. The core of the argument by Tobin et al. (2010) states that many researchers infer gender identity from these measures of self-attribute associations. In other words, simply identifying traits in oneself, such as is the procedure in the measure by Kessels (2005), the BSRI (Bem, 1974) or PAQ (Spence, 1991), does not necessarily relate to gender identity (Tobin et al., 2010). This is a valid point and thus our work consciously separated gender role self-concept from gender identity. Measuring gender role self-concept as we have done is essentially measuring the extent to which people identify with gender stereotypes. As gender stereotypes have a

descriptive and a pre-/proscriptive function (Eckes, 2008; Rudman & Glick, 2010), it is difficult to judge the fit of stereotypes to oneself in a neutral manner. This relates to a core issue in research on gender: how can one account for the pressure that gender stereotypes place on individuals when measuring stereotypes (Beyer, 1999)? The social desirability aspect inherent in gender stereotypes (Eckes, 2008; Rudman & Glick, 2010) makes an assessment of personal traits understood to be gendered by most people almost inherently biased, as explored in manuscript 3.

Relying strongly on gender stereotypes in one's responses could be a potential issue, but weak activation of gender related knowledge could be equally limiting. In manuscript 2, we hypothesized that feminine, prosocial behavior would lead to greater effort rather than ability ascriptions for high achieving students. While we did find support for this pattern, results also revealed that the prosocial target student was perceived as both more feminine *and* masculine than the nondescript student. These results suggest that prosociality was not strongly only stereotyped as feminine in our sample. Would target prosocial behavior seen as even more feminine, perhaps encompassing greater nurturing, relational characteristics (Eagly, 2009), have yielded stronger direct effects? Ensuring that experimental materials clearly evoke the desired gender stereotype would benefit future literature.

Overall, this work has adopted a more traditional approach to the constructs of gender, gender stereotypes, and gender role self-concepts than recommended by the GSSM (Tobin et al., 2010). However, the careful consideration and delineation of the terms used in this present dissertation is a strength of the work, while still acknowledging the difficulties of assessing gender stereotyped behaviors and traits.

Future Research Directions

In this section I will suggest further research to expand the perspectives presented here, those of educators, peers, and students, to parents in the educational context.

Acknowledging the role of status in gender dynamics directly could also illuminate differences between school and working environments, which seem to operate under different reward categories (Steinmayr & Kessels, 2017).

Multiple perspectives

One aim of this work was to explore gendered characteristics from multiple perspectives, these being educators, peers, and the self. Future work could benefit from including the three perspectives in one study. Measuring students' prosociality, gender role self-concepts, and attributions for their own successes and failures and comparing these values with how their peers (perhaps friendship dyads) and educators view their current and subsequent achievements and attributions for these achievements could be beneficial. Of course, such longitudinal investigations would require a comprehensive dataset. Authors note that in-group favoritism, classroom relationship networks, and GPA could confound such results on peer attributions, for example (Grow et al., 2016; Kisfalusi et al., 2019). For this reason, experimental research can offer meaningful insight but is not necessarily a naturalistic reflection of the academic setting. A method of including multiple perspectives in a single, experimental investigation was used by Schoneveld and Brummelman (2023). This work, in a similar vein to study 2 of the present dissertation, focused on the paradox of praise and attributional outcomes, specifically for students from low socioeconomic backgrounds. While study 2 captured how students expected teachers to react, Schoneveld and Brummelman (2023) directly asked teachers to respond to hypothetical students with different social backgrounds. Such responses were then included in vignettes and shown to students, who were asked how these teacher statements could be understood in terms of attributions (see Butler 1994), showing that indeed, low SES students were given inflated praise by teachers, which was then interpreted as a cue of low ability by students (Schoneveld & Brummelman, 2023). In this way, the researchers managed to circumvent the issue of complex classroom

relationships and include multiple perspectives in a single investigation on attributional backlash.

I, additionally, would recommend adding a fourth perspective to the educational system, that being the parents. Parents shape the home experience of students and provide opportunities for academic but also personal support and guidance (Eccles & Harold, 1993). In line with theoretical models, such as (situated) expectancy value theory (Eccles & Wigfield, 2020) and the parent socialization model (Eccles & Davis-Kean, 2005), parents' values and beliefs can impact the motivational beliefs of their children (Jacobs et al., 2005; Simpkins et al., 2012). Gender and gendered characteristics, specifically gender stereotypes (Eccles et al., 1990; Eccles & Davis-Kean, 2005; Eccles & Wigfield, 2020), play a role here. When parents more strongly endorsed gender stereotypes about boys' greater aptitude in math, it related to how they view their children's abilities in math, which in turn related to how children perceived their own abilities, shaping their performance (Jacobs, 1991). More recent empirical work found that parents' gender stereotyped beliefs related to their sons' reading skills, specifically finding lower reading skills in boys whose parents believed girls to have superior reading competencies (Muntoni & Retelsdorf, 2019). Such findings highlight that gender stereotypes held by relevant socializing agents, such as parents, can shape academic and personal development and could be considered in future ventures.

Intersectionality

Future investigations should be encouraged to include a more intersectional perspective. Intersectionality can be defined as the understanding that membership in social groups can interact and give rise to "different meanings and experiences" (Warner, 2008, p. 454). Gender can interact with sexual identity (Sanborn-Overby & Powlishta, 2020; Vantieghem & Van Houtte, 2020), ethnic background (Buckley & Carter, 2005; Skinner et al., 2018), and socioeconomic status (Friedman, 2013; Kane, 2006; Morris, 2008) and thus impact patterns of outcomes in a complex manner. Research shows that the content of

stereotypes can also differ as a function of gender and other identities (Klysing, 2023). Stereotypes about Black women, for example, could not be ascertained by adding up stereotypes about women and Black people but instead reflect characteristics exclusive to this identity (Ghavami & Peplau, 2013, p. 117). An intersectional psychological approach encompasses careful consideration of multiple identities and how these can be accurately compared with one another (Warner, 2008; Grabe 2020). For the studies presented in this dissertation, it was not possible to collect or analyze multiple social identities beyond gender with enough power to find meaningful effects.

Particularly SES could present interesting avenues for future research, as gender expression and class identity can interact. Qualitative and quantitative work showed parents from higher socioeconomic backgrounds have encouraged some counterstereotypical behavior, particularly in girls (Friedman, 2013), while lower class boys displayed hegemonic masculinity (Kane, 2006; Morris, 2008). Adolescents high in SES also had more positive attitudes towards the competence of a gender atypical student presented in a vignette (Meimoun et al., 2023). Viewing gender with reference to status can offer insight into such findings. For example, a man experiencing lower status due to unfavorable economic and social situations might be more averse to risking more status loss by behaving in a feminine manner. Similarly, he may disapprove of atypicality in others: masculine women could threaten his position, and feminine men undermine one of his remaining status characteristics, his own masculinity. This relates to work on ‘precarious manhood’ (Vandello et al., 2008), stating that masculinity needs to be reaffirmed often, leading to feelings of anxiety (Vandello & Bosson, 2013). Experimentally, men who believed themselves to have performed in a gender atypical manner on a knowledge test, thus feeling more feminine and perhaps lower in status, subsequently indicated greater anxiety (Vandello et al., 2008).

A future consideration could relate to work on priming (Guinote et al., 2015; Zhang et al., 2022), showing that leading participants into believing themselves to hold low or high social status had an impact on their subsequent (prosocial) behavior. It could be worthwhile to prime participants and manipulate their perceived status in society (as opposed to their perceived gender typicality [Vandello et al., 2008]) and subsequently measure their gender role self-concepts and willingness to uphold gender norms. Greater status could relate to more lenient attitudes towards gender atypicality and violations to gender stereotypes (Meimoun et al., 2023; Friedman, 2013). However, it is also possible that individuals believing themselves to hold greater status may show less equitable views; after all, changes to the status quo could result in a greater loss of their resources and power compared to lower status individuals. Indeed, past research has found greater social dominance orientation and anti-egalitarian views in individuals with higher status (Sidanius et al., 2000). Additionally, participants believing themselves to be low status have shown greater communion, more egalitarian views, and prosociality (Guinote et al., 2015; Zhang et al., 2022). The proposed research could offer new experimental insight into questions of status and gendered characteristics, especially if analyzed gender differentially. Manipulations of perceived status and the security of this status could impact men and women differently: men may react more negatively to status loss than women, for whom such an experience may be a more regular occurrence. Significant findings could open the door for interventions aimed at gender inequality – could artificial indications of status lead to real effects capable of advancing change? While the above could expand our understanding of status and gender, it would still leave other identities, such as ethnicity or sexual orientation, unexplored, leaving room for more comprehensive, intersectional investigations.

Implications

There are many implications of the present work, due to the breadth of the research, which investigated the views of educators, peers, and students themselves and expanded on gaps in the literature on gender stereotypes, attributional backlash, and gender differential impacts in the academic context. I will discuss gender stereotypes and gender role self-concepts, before regarding femininity and masculinity in school and the question of whether current developments in gender discourses are increasing the salience of gender.

Gender Stereotypes and Gender Role Self-Concepts

In the introduction of this dissertation, I focused on gender (in)congruity and its associated outcomes. I outlined how adherence to gender stereotypes was associated with beneficial outcomes (Braun & Davidson, 2017; Jewell & Brown, 2014). Violations of gender stereotypes, however, have resulted in backlash (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013). Although adherence to stereotypes had been framed positively in the literature, in the present work, adherence to feminine stereotyped behavior was detrimental for girls. In fact, girls showing masculinity received the most positive evaluations, despite violating gender stereotypes. Violating stereotypes may be negative when the violating behavior is feminine (Coyle et al., 2016; Sullivan et al., 2018; Thomas & Blakemore, 2013), whereas masculine behavior can be beneficial. This is also in line with gender as a status characteristic, which highlights adherence to masculinity as more valuable than adherence to femininity (Feinman, 1981). Such results resemble findings on an androgynous gender role self-concept, which has been presented as advantageous, especially for adolescents and children (N. Choi, 2004; Korlat et al., 2022; Pauletti et al., 2017). It has been theorized that the origin of the benefits of the androgynous orientation may be the high masculinity (W. H. Jones et al., 1978; Korlat et al., 2022; Whitley, 1984). Enacting masculine

gender stereotypes and gender role self-concepts showed beneficial associations in the works of this dissertation, particularly for girls.

The framing of this dissertation drew a distinction between gender stereotypes and gender role self-concepts. This distinction, as seen in the GSSM (Tobin et al., 2010), can be helpful but can also obscure potential similarities. These constructs are designed in similar ways, with participants evaluating characteristics seen as typical and desirable for the two genders and selecting distinctly masculine or feminine items (Bem, 1981; Kessels, 2005; Spence, 1991; Rudman et al., 2012). Going forth, acknowledging the similarities in gender stereotypes and gender role self-concepts, while not treating them as identical, could be beneficial for identifying patterns in results.

Gender Salience in Schools

The literature on gender and gender stereotypes in school has presented femininity as a good fit for school (Heyder & Kessels, 2013; Steinmayr & Kessels, 2017) and masculinity (Jackson & Dempster, 2009; Jones & Myhill, 2004), primarily focusing on negative masculinity (Lyng, 2009), as detrimental. The implications of this could be examined more closely. I will first highlight the implications of the framing of school as feminine, then very briefly question the recommendation of encouraging masculine behaviors, and finally outline the idea of moving beyond gender as a categorizing system.

In manuscript 2, we stated that the educational system should be mindful of encouraging feminine stereotyped social behaviors, such as prosocial behavior, in students, as this type of behavior could lead to unexpected and undesirable attributions for students. This may seem counterintuitive, as much feminine stereotyped behavior seen in classrooms, such as being diligent and effortful (Heyder & Kessels, 2017; Myhill, 2002), is conducive to the learning environment (Anaya & Zamarro, 2023; Park & Kim, 2023; Verniers et al., 2016). However, the framing of school as feminine can be a disservice to students of all genders.

This view alienates male students from this domain (Hadjar & Lupatsch, 2010), and indeed, boys show lower school engagement and lower levels of school belonging than girls (Huyge et al., 2015; Kessels & Van Houtte, 2021; Van Houtte, 2023). While girls do show better academic achievement and earn higher learning certificates (Federal Statistical Office, 2023; O'Dea et al., 2018; Voyer & Voyer, 2014), female students feel pressure to perform academically (Herrmann et al., 2019). At the same time, as previously outlined, their high achievements are portrayed as effortful (Espinoza et al., 2014; Fennema et al., 1990; Räty et al., 2002; Tiedemann, 2000) and thus less desirable (Nicholls, 1976). In a truly feminized domain, one could expect feminine achievement to be valued highly by educators and peers and be associated with girls experiencing greater confidence in their abilities. This is, however, not the case. Instead, boys are more likely to display stronger confidence in their intelligence than girls, even when controlling for measured intelligence (Steinmayr & Spinath, 2009), and boys, more than girls, are more likely to attribute their successes to their abilities (Assouline et al., 2021; Bornholt & Möller, 2003; Lohbeck et al., 2017; Mok et al., 2011).

As the works in this present dissertation show, feminine stereotyped behavior can be linked with denigrating attributions, and masculine stereotyped behavior and gender role self-concepts can be associated with favorable competence ratings by educators and peers and advantageous associations with self-esteem and achievement motivation for students themselves. Given that instrumentality and agency seem advantageous for students in school, as well as for future career success (Abele, 2003), some may ask: Should educational contexts emphasize masculinity more strongly (cf. manuscript 3)? This could increase the well-being of students and also prepare them for the working world. Researchers have already constructed interventions aimed at increasing (positive) masculinity in the school context, with results indicating positive outcomes, such as greater willingness to question harmful gender norms (Namy et al., 2015, p. 216). I would, however, caution against more sweeping

"masculinization" efforts, so as not to cater the school environment from one gender to the other. Rather, students could profit from a de-gendering of school (Kessels, 2002; Kessels & Hannover, 2008). Prosocial behavior (manuscript 2) and other gender stereotyped behavior, such as playing with certain toys (manuscript 1), could be seen as neutral in regard to gender. The decoupling of attributes from gender was discussed by Bem (1981), claiming that gender aschematicity, or a social world in which maleness and femaleness would not provide an expansive network of associations outside of the biological, would be desirable.

The implications of this work apply to both students and educators. Students themselves, particularly adolescents, have the agency to monitor their learning environment (Galanes & Carmack, 2013) and can draw attention to teachers or peers who reinforce gender stereotypes (Ayres & Leaper, 2013). Evidence from an intervention in which students were trained to identify sexism showed these students to respond more to problematic remarks from peers than students not exposed to the intervention (Pahlke et al., 2014). In this case, attention to gender was highlighted rather than downplayed. Reducing the salience of gender has also been explored in past work. Having single-sex classes (Kessels & Hannover, 2008) has been linked with benefits for female students in a male-stereotyped subject, showing a negative association between the girls' accessibility of gender-related self-knowledge and their ability self-concepts in physics.

Educators, whose behavior and views can impact students (Dai et al., 2022; Retelsdorf et al., 2015), can attempt to reduce gender stereotypes in their classrooms. Experimental work in early childhood pedagogy suggests that a gender-neutral environment, in which teachers minimized emphasis on gender stereotypes and gendered language, did lead children to be more open to playing with other-gender children and less likely to apply gender stereotypes (Shutts et al., 2017). Recent findings show encouraging patterns, with teachers reporting lower gender salience, such as lower gender segregation, and egalitarian gender-role attitudes

(Farago et al., 2022). As Bem (1983), who attempted to raise gender aschematic children stated, children will eventually develop a gender schema from interacting with the broader environment. Unless wide swaths of our society, including its institutions, adopt gender neutral practices, it will be difficult to shrink the expansive network of associations we hold with gender.

Gender Salience in Beyond Schools

In recent years, gender has become an increasingly complex category, with researchers calling for consideration of gender beyond the binary (Hyde et al., 2019). Data show a global generational divide (Ipsos, 2023), with more younger people compared to older people identifying as trans* (experiencing a mismatch between biological sex and gender identity [Antidiskriminierungsstelle des Bundes, 2023]) or nonbinary (gender identity in which the person neither identifies as male or female [Hegarty et al., 2018]). Alongside such social changes, people have been discussing the use of gender-neutral pronouns (Bigler & Leaper, 2015; Keener & Kotvas, 2023; Wallner & Eriksson Barajas, 2023) and gender-fair language (Spinelli et al., 2023; Vervecken et al., 2013; H. Xiao et al., 2023). Such developments could be seen in multiple ways: as a society increasing or decreasing the salience of gender.

The use of gender neutral third-person pronouns is becoming increasingly common (Arnold et al., 2021) in certain languages (they/them – English, hen – Swedish, hen/die – Danish [Decock et al., 2023]). In one qualitative work, Swedish students readily used the gender-neutral pronoun in classroom discussions (Wallner & Eriksson Barajas, 2023). A quantitative study with job descriptions using masculine pronouns (he/him/his), both sets of pronouns (he/him/his, she/her/hers), or gender-neutral pronouns (they/them/their) showed women responding with greater felt belongingness in response to the description with the gender-neutral pronouns and both sets of pronouns, while presentation made no significant difference for men (Keener & Kotvas, 2023). Bigler and Leaper (2015) have been proponents

of gender-neutral language, stating that using gendered pronouns could overly emphasize gender when it is not relevant or even damaging to do so.

Discussions about the merits of using gender-fair language (German: “gendern”) also touch on the topic of gender-neutrality (Spinelli et al., 2023; Vervecken et al., 2013; H. Xiao et al., 2023). There are multiple ways proposed to use gendered language in a gender-fair manner: While some advocate for the generic masculine form (“Politiker”) as being seen as gender-neutral, others propose naming both the masculine and feminine forms (“Politiker und Politikerinnen”) as a way to include both genders, while others advocate for the inclusion of symbols to denote a reference to all genders, a commonly used symbol being the asterisk (“Politiker*innen”) (Schach, 2023). Research has studied how the different methods of gender-fair language impact perceivers (Spinelli et al., 2023), one study showing fewer gender stereotyped associations when male-stereotyped occupations were presented in gendered pairs compared to the generic masculine (Vervecken et al., 2013). Researchers have theorized that not viewing the generic masculine as gender-neutral may be a recent phenomenon (H. Xiao et al., 2023). The generic masculine may have previously been interpreted in a gender-neutral manner, but as other gender-fair methods have become more common, exclusively naming the masculine is now more often interpreted as exclusively meaning the masculine. Recent work showed that descriptions of groups written in gender-fair language compared to generic masculine, increased the perceived ratio of women to men (H. Xiao et al., 2023). This work also acknowledges actual gender ratios, raising the question of which type of mental representation is preferable. Should our mental representation of, say, politicians match the observed gender distribution of this profession, or should we aim for gender balance, despite this not necessarily reflecting the status quo?

The topics of additional gender identities and the use of gender-fair language are brought forth here to illustrate that we are currently in a time of increased focus on gender

dynamics. Despite this focus, specifically in the case of gender-fair language, being on more gender equality, one could argue that repeated stress on gender as a dividing category in society could increase the salience of gender and bring with it possible side-effects (Bigler & Leaper, 2015).

Conclusion

This work has discussed gender and gender stereotyped traits and behaviors in the educational context from the perspectives of educators, peers, and the self. Findings showed that not just gender, but also gender stereotypes and gender role self-concept can greatly impact how students are seen by others and how they see themselves, beyond just academic achievement. Overall, the works found detrimental impacts of feminine stereotyped behaviors, which related to denigrating attributional patterns and worse judgments of competence, creativity, and self-esteem compared to masculine behaviors. Such findings could be seen as being at odds with the feminization of school hypothesis, which has been used to outline negative outcomes of masculinity and the positive outcomes of femininity. However, the results are unsurprising when regarding the status differential between masculinity and femininity, with masculinity routinely associated with advantageous outcomes. Such beneficial effects were found in the present work, showing masculine stereotyped behaviors to relate to better ratings of perceived competence, creativity, and self-esteem and a masculine gender role self-concept to be associated with beneficial self-esteem and achievement motivation, particularly for girls.

Going forward, in a society that has an increasingly wider range of gender identities and a focus on gender-inclusive language, it is prudent to remain mindful of the changing nature of gender dynamics and stereotypes. Continuing to collect data in a variety of settings with a multitude of samples can help us understand these changing gender dynamics and apply this knowledge to the educational sector in order to better foster learning and individual development.

References II.

Abele, A. E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology*, *85*(4), 768–776. https://doi.org/10.1037/0022-3514.85.4.768

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751–763. https://doi.org/10.1037/0022-3514.93.5.751

Anaya, L. M., & Zamarro, G. (2023). The role of student effort on performance in PISA: revisiting the gender gap in achievement. *Oxford Economic Papers*, *76*(2), 533-560. https://doi.org/10.1093/oep/gpad018

Antidiskriminierungsstelle des Bundes. (2023). *Geschlecht und Geschlechtsidentität [Gender and gender identity]*. https://www.antidiskriminierungsstelle.de/DE/ueber-diskriminierung/diskriminierungsmerkmale/geschlecht-und-geschlechtsidentitaet/geschlecht-und-geschlechtsidentitaet-node.html

Arnold, J. E., Mayo, H. C., & Dong, L. (2021). My pronouns are they/them: Talking about pronouns changes how pronouns are understood. *Psychonomic Bulletin and Review*, *28*(5), 1688–1697. https://doi.org/10.3758/s13423-021-01905-0

Assouline, S. G., Mahatmya, D., Ihrig, L., & Lane, E. (2021). High-achieving rural middle-school students' academic self-efficacy and attributions in relationship to gender. *High Ability Studies*, *32*(2), 143–169. https://doi.org/10.1080/13598139.2020.1740582

Ayres, M. M., & Leaper, C. (2013). Adolescent girls' experiences of discrimination: An examination of coping strategies, social support, and self-esteem. *Journal of Adolescent Research*, *28*(4), 479–508. https://doi.org/10.1177/0743558412457817

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, *42*(2), 155–162.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, *88*(4), 354–364.

Bem, S. L. (1983). Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. *Signs: Journal of Women in Culture and Society*, *8*(4), 598–616. https://doi.org/10.1086/494131

Bem, S. L. (1993). *The lenses of gender: transforming the debate on sexual inequality*. Yale University Press.

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, *37*(3), 241–255. https://www-jstor-org.ezp2.lib.umn.edu/stable/2093465

Berger, J., M. Hamit, F., Norman, R. Z., & Jr., M. Z. (1977). *Status characteristics and social interaction: An expectation states approach.* Elsevier.

Best, D. L. (2009). Another view of the gender-status relation. *Sex Roles*, *61*(5–6), 341–351. https://doi.org/10.1007/s11199-009-9629-1

Beyer, S. (1999). The accuracy of academic gender stereotypes. *Sex Roles*, *40*(9-10), 787–813. https://doi.org/10.1023/A:1018864803330

Bigler, R. S., & Leaper, C. (2015). Gendered language: psychological principles, evolving practices, and inclusive policies. *Policy Insights from the Behavioral and Brain Sciences*, *2*(1), 187–194. https://doi.org/10.1177/2372732215600452

Bøe, T., Hysing, M., Skogen, J. C., & Breivik, K. (2016). The strengths and difficulties questionnaire (SDQ): Factor structure and gender equivalence in Norwegian adolescents. *PLoS ONE*, *11*(5), e0152202. https://doi.org/10.1371/journal.pone.0152202

Bornholt, L., & Möller, J. (2003). Attributions about achievement and intentions about further study in social context. *Social Psychology of Education*, *6*(3), 217–231. https://doi.org/10.1023/A:1024715609124

Braun, S. S., & Davidson, A. J. (2017). Gender (non)conformity in middle childhood: A mixed methods approach to understanding gender-typed behavior, friendship, and peer preference. *Sex Roles*, *77*(1–2), 16–29. https://doi.org/10.1007/s11199-016-0693-z

Buckley, T. R., & Carter, R. T. (2005). Black adolescent girls: Do gender role and racial identity: Impact their self-esteem? *Sex Roles*, *53*(9–10), 647–661. https://doi.org/10.1007/s11199-005-7731-6

Butler, R. (1994). Teacher communications and student interpretations: Effects of teacher responses to failing students on attributional inferences in two age groups. *British Journal of Educational Psychology*, *64*(2), 277–294. https://doi.org/10.1111/j.2044-8279.1994.tb01102.x

Canevello, A. (2020). Gender schema theory. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1741–1743). Springer Nature. https://doi.org/10.4324/9781315664095

Chestnut, E. K., Lei, R. F., Leslie, S. J., & Cimpian, A. (2018). The myth that only brilliant people are good at math and its implications for diversity. *Education Sciences*, *8*(2), 65. https://doi.org/10.3390/educsci8020065

Choi, H., Kim, J. H., Hwang, M. H., & Heppner, M. J. (2010). Self-esteem as a mediator between instrumentality, gender role conflict and depression in male Korean high school students. *Sex Roles*, *63*(5), 361–372. https://doi.org/10.1007/s11199-010-9801-7

Choi, N. (2004). Sex role group differences in specific, academic, and general self-efficacy. *Journal of Psychology: Interdisciplinary and Applied*, *138*(2), 149–159. https://doi.org/10.3200/JRLP.138.2.149-159

Conway, M., Pizzamiglio, M. T., & Mount, L. (1996). Status, communality, and agency: Implications for stereotypes of gender and other groups. *Journal of Personality and Social Psychology*, *71*(1), 25–38. https://doi.org/10.1037/0022-3514.71.1.25

Conway, M., & Vartanian, L. R. (2000). A status account of gender stereotypes: Beyond communality and agency. *Sex Roles*, *43*(3–4), 181–199. https://doi.org/10.1023/A:1007076813819

Covington, M. V. (1984). The self-worth theory of achievement motivation: Findings and implications. *The Elementary School Journal*, *85*(1), 5–20. https://doi.org/10.1086/461388

Coyle, E. F., Fulcher, M., & Trübutschek, D. (2016). Sissies, mama's boys, and tomboys: Is children's gender nonconformity more acceptable when nonconforming traits are positive? *Archives of Sexual Behavior*, *45*, 1827–1838. https://doi.org/10.1007/s10508-016-0695-5

Dai, A., Li, D., Zhu, H., & Zhang, Z. (2022). How do teachers' gender stereotypes impact students? *Advances in Social Science, Education and Humanities Research*, *637*, 728–733.

Decock, S., Van Hoof, S., Soens, E., & Verhaegen, H. (2023). The comprehensibility and appreciation of non-binary pronouns in newspaper reporting. The case of hen and die in Dutch. *Applied Linguistics*, amad028. https://doi.org/10.1093/applin/amad028

Dudenredaktion. (n.d.). *"Geschlecht" auf Duden online [„Gender" in Duden online]*. https://www.duden.de/rechtschreibung/Geschlecht

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Erlbaum.

Eagly, A. H. (2009). The his and hers of prosocial behavior: An examination of the social psychology of gender. *American Psychologist*, *34*(8), 644–658. https://doi.org/10.1037/0003-066X.64.8.644

Eccles, J. S., & Davis-Kean, P. E. (2005). Influences of parents' education on their children's educational attainments: The role of parent and child perceptions. *London Review of Education*, *3*, 191–204.

Eccles, J. S., & Harold, R. D. (1993). Parent-school involvement during the early adolescent years. *Teachers College Record: The Voice of Scholarship in Education*, *94*(3), 568–587. https://doi.org/10.1177/016146819309400311

Eccles, J. S., Jacobs, J. E., & Harold, R. D. (1990). Gender role stereotypes, expectancy effects, and parents' socialization of gender differences. *Journal of Social Issues*, *46*(2), 183–201. https://doi.org/10.1111/j.1540-4560.1990.tb01929.x

Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A developmental, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, *61*(4), 101859. https://doi.org/10.1016/j.cedpsych.2020.101859

Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen [Gender stereotypes: Roles, identities and prejudice]. In R. Becker & B. Kortendiek (Eds.), *Handbuch Frauen- und Geschlechterforschung: Theorie, Methoden, Empirie* (pp. 178–189). Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-91972-0_20

Egan, S. K., & Perry, D. G. (2001). Gender identity: A multidimensional analysis with implications for psychosocial adjustment. *Developmental Psychology*, *37*(4), 451–463. https://doi.org/https://doi.org/10.1037/0012-1649.37.4.451

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 646–718). John Wiley & Sons, Inc. https://doi.org/10.4324/9781315560984-2

Espinoza, P., Arêas da Luz Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: Teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, *17*(1), 105–126. https://doi.org/10.1007/s11218-013-9226-6

Eurostat. (2023). *Employment by sex, age, professional status and occupation*. https://doi.org/https://doi.org/10.2908/LFSA_EGAIS

Farago, F., Martin, C. L., Granger, K. L., Santos, C. E., & Miller, C. F. (2022). Teachers' gender-role attitudes and gendered classroom practices. *Sex Roles*, *87*(9–10), 471–486. https://doi.org/10.1007/s11199-022-01331-z

Federal Statistical Office. (2020). *Bildungsstand der Bevölkerung: - Ergebnisse des Mikrozensus 2019 [Educational level of the population: Results of the micro-census 2019]*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Bildungsstand/Publikationen/Downloads-Bildungsstand/bildungsstand-bevoelkerung-5210002197004.pdf?__blob=publicationFile

Federal Statistical Office. (2022). *Statistik des Hochschulpersonals: Berichtsjahr 2022 [Higher education staff statistics: Year of report 2022*]. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/statistischer-bericht-hochschulpersonal-2110440227005.html

Federal Statistical Office. (2023). *Allgemeinbildende und Berufliche Schulen - Schnellmeldung Studienberechtigte. Abgangsjahr 2023. [General and vocational schools - Quick report Students eligible for higher education. Graduation year 2023.]*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Publikationen/Downloads-Schulen/statistischer-bericht-studienberechtigte-schnellmeldung-5211001238005.html

Federal Statistical Office. (2024*). Statistischer Bericht - Allgemeinbildende und Berufliche Schulen - Schnellmeldung Schülerinnen und Schüler - Vorläufige Ergebnisse Schuljahr 2023/2024 [Statistical report - General and vocational schools - Quick report on pupils - Preliminary results school year 2023/2024.].* https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Schulen/Publikationen/Downloads-Schulen/statistischer-bericht-sm-schueler-5211003248005.html

Feinman, S. (1981). Why is cross-sex-role behavior more approved for girls than for boys? A status characteristic approach. *Sex Roles*, *7*(3), 289–300. https://doi.org/10.1007/BF00287543

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55–69. https://doi.org/10.1007/BF00311015

Friedman, H. L. (2013). Tiger girls on the soccer field. *Contexts*, 12(4), 30–35. https://doi.org/10.1177/1536504213511213

Galanes, G. J., & Carmack, H. J. (2013). "He's really setting an example": Student contributions to the learning environment. *Communication Studies*, *64*(1), 49–65. https://doi.org/10.1080/10510974.2012.731464

Gerber, G. L. (2009). Status and the gender stereotyped personality traits: Toward an integration. *Sex Roles*, *61*(5–6), 297–316. https://doi.org/10.1007/s11199-008-9529-9

Gerbino, M., Zuffianò, A., Eisenberg, N., Castellani, V., Luengo Kanacri, B. P., Pastorelli, C., & Caprara, G. V. (2018). Adolescents' prosocial behavior predicts good grades beyond

intelligence and personality traits. *Journal of Personality*, *86*(2), 247–260. https://doi.org/10.1111/jopy.12309

Ghavami, N., & Peplau, L. A. (2013). An intersectional analysis of gender and ethnic stereotypes: Testing three hypotheses. *Psychology of Women Quarterly*, *37*(1), 113–127. https://doi.org/10.1177/0361684312464203

Grabe, S. (2020). Research methods in the study of intersectionality in psychology: examples informed by a decade of collaborative work with majority world women's grassroots activism. *Frontiers in Psychology*, *11*. 494309. https://doi.org/10.3389/fpsyg.2020.494309

Graham, S., & Chen, X. (2020). An attributional approach to teacher praise. In E. Brummelman (Ed.), *Psychological perspectives on praise* (pp. 19–26). Rand McNally. https://doi.org/10.4324/9780429327667-4

Grow, A., Takács, K., & Pál, J. (2016). Status characteristics and ability attributions in Hungarian school classes: An exponential random graph approach. *Social Psychology Quarterly*, *79*(2), 156–167. https://doi.org/10.1177/0190272516643052

Guinote, A., Cotzia, I., Sandhu, S., & Siwa, P. (2015). Social status modulates prosocial behavior and egalitarianism in preschool children and adults. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(3), 731–736. https://doi.org/10.1073/pnas.1414550112

Hadjar, A., Krolak-Schwerdt, S., Priem, K., & Glock, S. (2014). Gender and educational achievement. *Educational Research*, *56*(2), 117–125. https://doi.org/10.1080/00131881.2014.898908

Hadjar, A., & Lupatsch, J. (2010). Der Schul(miss)erfolg der Jungen: Die Bedeutung von sozialen Ressourcen, Schulentfremdung und Geschlechterrollen [The lower educational success of boys—The impact of social resources, school alienation and gender role patterns]. *Kölner Zeitschrift Für Soziologie Und Sozialpsychologie*, *62*(4), 599–622. https://doi.org/10.1007/s11577-010-0116-z

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(8), 2395–2400. https://doi.org/10.1073/pnas.1416587112

Hegarty, P., Ansara, Y. G., & Barker, M.-J. (2018). Nonbinary gender identities. In N. K. Dess, J. Marecek, & L. C. Bell (Eds.), *Gender, sex, and sexualities: Psychological perspectives* (pp. 53–76). Oxford University Press.

Herrmann, J., Koeppen, K., & Kessels, U. (2019). Do girls take school too seriously? Investigating gender differences in school burnout from a self-worth perspective. *Learning and Individual Differences*, *69*, 150–161. https://doi.org/10.1016/j.lindif.2018.11.011

Heyder, A., van Hek, M., & van Houtte, M. (2021). When gender stereotypes get male adolescents into trouble: A longitudinal study on gender conformity pressure as a predictor of school misconduct. *Sex Roles*, *84*, 61–75. https://doi.org/https://doi.org/10.1007/s11199-020-01147-9

Heyder, A., & Kessels, U. (2013). Is school feminine? Implicit gender stereotyping of school as a predictor of academic achievement. *Sex Roles*, *69*(11–12), 605–617. https://doi.org/10.1007/s11199-013-0309-9

Heyder, A., & Kessels, U. (2015). Do teachers equate male and masculine with lower academic engagement? How students' gender enactment triggers gender stereotypes at school. *Social Psychology of Education*, *18*(3), 467–485. https://doi.org/10.1007/s11218-015-9303-0

Heyder, A., & Kessels, U. (2017). Boys don't work? On the psychological benefits of showing low effort in high school. *Sex Roles*, *77*(1–2), 72–85. https://doi.org/10.1007/s11199-016-0683-1

Hine, B. A. (2017). Identifying the male prosocial niche: the gender-typing of prosocial behaviour across childhood and adolescence. *European Journal of Developmental Psychology*, *14*(2), 206–220. https://doi.org/10.1080/17405629.2016.1186009

Hinojosa, M. (2021). The descriptive representation of women in politics. *Oxford Research Encyclopedia of Politics*, 1–24. https://doi.org/10.1093/acrefore/9780190228637.013.1978

Huyge, E., Van Maele, D., & Van Houtte, M. (2015). Does students' machismo fit in school? Clarifying the implications of traditional gender role ideology for school belonging. *Gender and Education*, *27*(1), 1–18. https://doi.org/10.1080/09540253.2014.972921

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, *74*(2), 171–193. https://doi.org/10.1037/amp0000307

Ipsos. (2023). *LGBT+ Pride 2023: A 30-Country Ipsos Global Advisor Survey*. https://www.ipsos.com/sites/default/files/ct/news/documents/2023-05/Ipsos LGBT%2B Pride 2023 Global Survey Report - rev.pdf

IPU Parline. (2024). *Global and regional averages of women in national parliaments*. https://data.ipu.org/women-averages/

Jackson, C., & Dempster, S. (2009). 'I sat back on my computer … with a bottle of whisky next to me': Constructing 'cool' masculinity through 'effortless' achievement in secondary and higher education. *Journal of Gender Studies*, *18*, 341–356. https ://doi.org/10.1080/09589 23090 32600 19.

Jackson, C., & Nyström, A. S. (2014). 'Smart students get perfect scores in tests without studying much': Why is an effortless achiever identity attractive, and for whom is it possible? *Research Papers in Education*, *30*(4), 393–410. https://doi.org/10.1080/02671522.2014.970226

Jacobs, J. E. (1991). Influence of gender stereotypes on parent and child mathematics attitudes. *Journal of Educational Psychology*, *83*(4), 518–527.

Jacobs, J. E. , Davis-Kean, P. , Bleeker, M. , Eccles, J. S. , & Malanchuk, O. (2005). "I can, but I don't want to": The impact of parents, interests, and activities on gender differences in math. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender difference in mathematics* (pp. 246–263). Cambridge University Press.

Jewell, J. A., & Brown, C. S. (2014). Relations among gender typicality, peer relations, and mental health during early adolescence. *Social Development*, *23*(1), 137–156. https://doi.org/10.1111/sode.12042

Jones, W. H., Chernovetz, M. E., & Hansson, R. O. (1978). The enigma of androgyny: Differential implications for males and females? *Journal of Consulting and Clinical Psychology*, *46*(2), 298–313. https://doi.org/10.1037/0022-006X.46.2.298

Jones, S., & Myhill, D. (2004). 'Troublesome boys' and 'compliant girls': Gender identity and perceptions of achievement and underachievement. *British Journal of Sociology of Education*, *25*(5), 547–561.

Kane, E. W. (2006). "No way my boys are going to be like that!": Parents' responses to children's gender nonconformity. *Gender and Society*, *20*(2), 149–176. https://doi.org/10.1177/0891243205284276

Keener, E., & Kotvas, K. (2023). Beyond he and she: does the singular use of "they, them, their" function generically as inclusive pronouns for cisgender men and women? *Gender Issues*, *40*(1), 23–43. https://doi.org/10.1007/s12147-022-09297-8

Kervyn, N., Bergsieker, H. B., & Fiske, S. T. (2012). The innuendo effect: Hearing the positive but inferring the negative. *Journal of Experimental Social Psychology*, *48*(1), 77–85. https://doi.org/10.1016/j.jesp.2011.08.001

Kervyn, N., Bergsieker, H. B., Grignard, F., & Yzerbyt, V. Y. (2016). An advantage of appearing mean or lazy: Amplified impressions of competence or warmth after mixed descriptions. *Journal of Experimental Social Psychology*, *62*, 17–23. https://doi.org/10.1016/j.jesp.2015.09.004

Kessels, U. (2002). *Undoing Gender in der Schule: Eine empirische Studie über Koedukation und Geschlechtsidentität im Physikunterricht [Undoing gender in school. An empirical study on coeducation and gender in physics classes]*. Juventa Verlag.

Kessels, U. (2005). Fitting into the stereotype: How gender-stereotyped perceptions of prototypic peers relate to liking for school subjects. *European Journal of Psychology of Education*, *20*(3), 309–323. https://doi.org/10.1007/BF03173559

Kessels, U. (2015). Bridging the gap by enhancing the fit: How stereotypes about STEM clash with stereotypes about girls. *International Journal of Gender Science and Technology*, *7*(2), 281–296.

Kessels, U., & Hannover, B. (2008). When being a girl matters less: Accessibility of gender-related self-knowledge in single-sex and coeducational classes and its impact on students' physics-related self-concept of ability. *British Journal of Educational Psychology*, *78*(2), 273–289. https://doi.org/10.1348/000709907X215938

Kessels, U., & Heyder, A. (2020). Not stupid, but lazy: Psychological benefits of disruptive classroom behavior from an attributional perspective. *Social Psychology of Education*, *23*, 583–613. https://doi.org/10.1007/s11218-020-09550-6

Kessels, U., & Van Houtte, M. (2021). Side effects of academic engagement? How boys' and girls' well-being is related to their engagement and motivational regulation. *Gender and Education*, *34*(6), 627-642. https://doi.org/10.1080/09540253.2021.2011840

Kisfalusi, D., Janky, B., & Takács, K. (2019). Double standards or social identity? The role of gender and ethnicity in ability perceptions in the classroom. *Journal of Early Adolescence*, *39*(5), 745–780. https://doi.org/10.1177/0272431618791278

Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin*, *125*(4), 470–500.

Klysing, A. (2023). Prototypicality at the intersection of gender and sexual orientation. *British Journal of Social Psychology*, *62*(3), 1506–1533. https://doi.org/10.1111/bjso.12645

Koenig, A. M. (2018). Comparing prescriptive and descriptive gender stereotypes about children, adults, and the elderly. *Frontiers in Psychology*, *9*, 381468. https://doi.org/10.3389/fpsyg.2018.01086

Korlat, S., Holzer, J., Schultes, M.-T., Buerger, S., Schober, B., Spiel, C., & Kollmayer, M. (2022). Benefits of psychological androgyny in adolescence: The role of gender role self-concept in school-related well-being. *Frontiers in Psychology*, *13*, 856758. https://doi.org/10.3389/fpsyg.2022.856758

Lohbeck, A., Grube, D., & Moschner, B. (2017). Academic self-concept and causal attributions for success and failure amongst elementary school children. *International Journal of Early Years Education*, *25*(2), 190–203. https://doi.org/10.1080/09669760.2017.1301806

Lyng, S. T. (2009). Is there more to "Antischoolishness" than masculinity? On multiple student styles, gender, and educational self-exclusion in secondary school. *Men and Masculinities*, *11*(4), 462–487. https://doi.org/10.1177/1097184X06298780

Ma, H. K. (2005). The relation of gender-role classifications to the prosocial and antisocial behavior of Chinese adolescents. *Journal of Genetic Psychology*, *166*(2), 189–202. https://doi.org/10.3200/GNTP.166.2.189-202

Martin, A. E., & Slepian, M. L. (2021). The primacy of gender: Gendered cognition underlies the big two dimensions of social cognition. *Perspectives on Psychological Science*, *16*(6), 1143–1158. https://doi.org/10.1177/1745691620904961

Matteucci, M. C., & Gosling, P. (2004). Italian and French teachers faced with pupil's academic failure: The "norm of effort." *European Journal of Psychology of Education*, *19*(2), 147–166. https://doi.org/10.1007/BF03173229

Meimoun, E., Bonnot, V., Berenguer, J., & Aelenei, C. (2023). How are gender counter-stereotypical adolescents evaluated by their peers? Investigating the backlash effect in French schools. *Social Psychology of Education*, 1-25. https://doi.org/10.1007/s11218-023-09807-w

Meyer, W. U., Bachmann, M., Biermann, U., Hempelmann, M., Ploger, F.-O., & Spiller, H. (1979). The informational value of evaluative behavior: Influences of praise and blame on perceptions of ability. *Journal of Educational Psychology*, *71*(2), 259–268. https://doi.org/10.1037/0022-0663.71.2.259

Miller, J. (1996). *School for women.* Virago Press.

Mok, M. M. C., Kennedy, K. J., & Moore, P. J. (2011). Academic attribution of secondary students: Gender, year level and achievement level. *Educational Psychology*, *31*(1), 87–104. https://doi.org/10.1080/01443410.2010.518596

Möller, J. (2005). Paradoxical effects of praise and criticism: Social, dimensional and temporal comparisons. *British Journal of Educational Psychology*, *75*(2), 275–295. https://doi.org/10.1348/000709904X24744

Morris, E. W. (2008). “Rednecks,” “rutters,” and ’rithmetic: Social class, masculinity, and schooling in a rural context. *Gender and Society*, *22*(6), 728–751. https://doi.org/10.1177/0891243208325163

Muntoni, F., & Retelsdorf, J. (2019). At their children’s expense: How parents’ gender stereotypes affect their children’s reading outcomes. *Learning and Instruction*, *60*, 95–103. https://doi.org/10.1016/j.learninstruc.2018.12.002

Myhill, D. (2002). Bad boys and good girls? Patterns of interaction and response in whole class teaching. *British Educational Research Journal*, *28*(3), 339–352.

Namy, S., Heilman, B., Stich, S., Crownover, J., Leka, B., & Edmeades, J. (2015). Changing what it means to ‘become a man’: participants’ reflections on a school-based programme to redefine masculinity in the Balkans. *Culture, Health and Sexuality*, *17*, 206–222. https://doi.org/10.1080/13691058.2015.1070434

Nicholls, J. G. (1976). Effort is virtuous, but it’s better to have ability: Evaluative responses to perceptions of effort and ability. *Journal of Research in Personality*, *10*(3), 306–315. https://doi.org/10.1016/0092-6566(76)90020-9

O’Dea, R. E., Lagisz, M., Jennions, M. D., & Nakagawa, S. (2018). Gender differences in individual variation in academic grades fail to fit expected patterns for STEM. *Nature Communications*, *9*(1), 3777. https://doi.org/10.1038/s41467-018-06292-0

Orbach, R. (2017). Bringing talent off the bench and into the game: the underrepresentation of women in the boardroom. *Fordham Journal of Corporate & Financial Law,* 22(2), 203.

Orlofsky, J. L., & Stake, J. E. (1981). Psychological masculinity and femininity: Relationship to striving and self-concept in the achievement and interpersonal domains. *Psychology of Women Quarterly*, *6*(2), 218–233. https://doi.org/10.1111/j.1471-6402.1981.tb00409.x

Orr, E., & Ben-Eliahu, E. (1993). Gender differences in idiosyncratic sex-typed self-images and self-esteem. *Sex Roles*, *29*(3–4), 271–296. https://doi.org/10.1007/BF00289939

Pahlke, E., Bigler, R. S., & Martin, C. L. (2014). Can fostering children’s ability to challenge sexism improve critical analysis, internalization, and enactment of inclusive, egalitarian peer relationships? *Journal of Social Issues*, *70*(1), 115–133. https://doi.org/10.1111/josi.12050

Park, G. R., & Kim, J. (2023). Parental involvement and academic achievement of adolescents: The mediating roles of school adjustment. *School Psychology International*. https://doi.org/10.1177/01430343231207134

Pauletti, R. E., Menon, M., Cooper, P. J., Aults, C. D., & Perry, D. G. (2017). Psychological androgyny and children’s mental health: A new look with new measures. *Sex Roles*, *76*(11–12), 705–718. https://doi.org/10.1007/s11199-016-0627-9

Pedhazur, E. J., & Tetenbaum, T. J. (1979). Bem Sex Role Inventory: A theoretical and methodological critique. *Journal of Personality and Social Psychology*, 37(6), 966-1016.

Piché, C., & Plante, C. (1991). Perceived masculinity, femininity and androgyny among primary school boys. *European Journal of Psychology of Education*, *6*(4), 423–435. https://doi.org/10.1007/BF03172775

Proudfoot, D., Kay, A. C., & Koval, C. Z. (2015). A gender bias in the attribution of creativity: Archival and experimental evidence for the perceived association between masculinity and creative thinking. *Psychological Science*, *26*(11), 1751–1761. https://doi.org/10.1177/0956797615598739

Puglisi, J. T. (1980). Equating the social desirability of Bem Sex-Rote Inventory masculinity and femininity subscales. *Journal of Personality Assessment*, *44*(3), 272–276. https://doi.org/10.1207/s15327752jpa4403_9

Quenneville, S., Talwar, V., & Bosacki, S. (2022). Teacher ratings and adolescent students' perceived social behaviours and gender-role orientations. *Journal of Gender Studies*, *31*(4), 444–456. https://doi.org/10.1080/09589236.2021.1988530

Rashotte, L. S., & Webster, M. (2005). Gender status beliefs. *Social Science Research*, *34*(3), 618–633. https://doi.org/10.1016/j.ssresearch.2004.05.004

Räty, H., Vänskä, J., Kasanen, K., & Kärkkäinen, R. (2002). Parents' explanations of their child's performance in mathematics and reading: A replication and extension of Yee and Eccles. *Sex Roles*, *46*(3–4), 121–128. https://doi.org/10.1023/A:1016573627828

Renold, E., & Allan, A. (2006). Bright and beautiful: High achieving girls, ambivalent femininities, and the feminization of success in the primary school. *Discourse*, *27*(4), 457–473. https://doi.org/10.1080/01596300600988606

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*, *107*(1), 186–194. https://doi.org/10.1037/a0037107

Rudman, L. A., & Glick, P. (2010). *The social psychology of gender: How power and intimacy shape gender relations*. Guilford Press.

Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, *48*(1), 165–179. https://doi.org/10.1016/j.jesp.2011.10.008

Saidah, B., Louvet, E., & Pansu, P. (2019). Are students who make an effort perceived as successful or just liked by their teachers? *Social Psychology of Education*, *22*(2), 405–419. https://doi.org/10.1007/s11218-019-09481-x

Sanborn-Overby, U. A., & Powlishta, K. K. (2020). Differential reactions to male and female gender-role violations: Testing the sexual orientation hypothesis. *Archives of Sexual Behavior*, *49*(7), 2531–2546. https://doi.org/10.1007/s10508-020-01803-x

Schach, A. (2023). Gendergerechte Sprache [Gender-fair language]. In *Diversity & Inclusion in Strategie und Kommunikation* (pp. 265–304). Springer: Gabler

Schoneveld, E., & Brummelman, E. (2023). "You did incredibly well!": teachers' inflated praise can make children from low-SES backgrounds seem less smart (but more hardworking). *Science of Learning*, *31*, 1–12. https://doi.org/10.1038/s41539-023-00183-w

Schöne, C., & Stiensmeier-Pelster, J. (2016). *Selbstwertinventar für Kinder und Jugendliche – SEKJ [Self-Esteem Inventory for Children and Adolescents]*. Hogrefe. https://doi.org/10.13140/RG.2.1.3583.5607.

Schöne, C., Tandler, S. S., & Stiensmeier-Pelster, J. (2015). Contingent self-esteem and vulnerability to depression: Academic contingent self-esteem predicts depressive symptoms in students. *Frontiers in Psychology*, *6*, 1573. https://doi.org/10.3389/fpsyg.2015.01573

Shutts, K., Kenward, B., Falk, H., Ivegran, A., & Fawcett, C. (2017). Early preschool environments and gender: Effects of gender pedagogy in Sweden. *Journal of Experimental Child Psychology*, *162*, 1–17. https://doi.org/10.1016/j.jecp.2017.04.014

Sidanius, J., Levin, S., Liu, J., & Pratto, F. (2000). Social dominance orientation, anti-egalitarianism and the political psychology of gender: An extension and cross-cultural replication. *European Journal of Social Psychology*, *30*(1), 41–67. https://doi.org/10.1002/(SICI)1099-0992(200001/02)30:1<41::AID-EJSP976>3.0.CO;2-O

Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles' expectancy - value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology*, *48*(4), 1019–1032. https://doi.org/10.1037/a0027468

Skinner, O. D., Kurtz-Costes, B., Wood, D., & Rowley, S. J. (2018). Gender typicality, felt pressure for gender conformity, racial centrality, and self-esteem in African American adolescents. *Journal of Black Psychology*, *44*(3), 195–218. https://doi.org/10.1177/0095798418764244

Spence, J. T. (1991). Do the BSRI and PAQ measure the same or different concepts? *Psychology of Women Quarterly*, *15*, 141–165.

Spinelli, E., Chevrot, J. P., & Varnet, L. (2023). Neutral is not fair enough: Testing the efficiency of different language gender-fair strategies. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1256779

Starr, C. R., & Zurbriggen, E. L. (2017). Sandra Bem's gender schema theory after 34 years: A review of its reach and impact. *Sex Roles*, *76*(9–10), 566–578. https://doi.org/10.1007/s11199-016-0591-4

Steinmayr, R., & Kessels, U. (2017). Good at school = successful on the job? Explaining gender differences in scholastic and vocational success. *Personality and Individual Differences*, *105*, 107–115. https://doi.org/10.1016/j.paid.2016.09.032

Steinmayr, R., & Spinath, B. (2008). Sex differences in school achievement: What are the roles of personality and achievement motivation? *European Journal of Personality*, *22*, 185–209. https://doi.org/10.1002/per

Steinmayr, R., & Spinath, B. (2009). What explains boys' stronger confidence in their intelligence? *Sex Roles*, *61*(9–10), 736–749. https://doi.org/10.1007/s11199-009-9675-8

Stoeber, J., & Rambow, A. (2007). Perfectionism in adolescent school students: Relations with motivation, achievement, and well-being. *Personality and Individual Differences*, *42*(7), 1379–1389. https://doi.org/10.1016/j.paid.2006.10.015

Sullivan, J., Ciociolo, A., & Moss-Racusin, C. A. (2022). Establishing the content of gender stereotypes across development. *PLoS ONE*, *17*(7), e0263217. https://doi.org/10.1371/journal.pone.0263217

Sullivan, J., Moss-Racusin, C., Lopez, M., & Williams, K. (2018). Backlash against gender stereotype-violating preschool children. *PLoS ONE*, *13*(4), e0195503. https://doi.org/10.1371/journal.pone.0195503

Thomas, R. N., & Blakemore, J. E. O. (2013). Adults' attitudes about gender nonconformity in childhood. *Archives of Sexual Behavior*, *42*(3), 399–412. https://doi.org/10.1007/s10508-012-0023-7

Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics*, *41*, 191–207.

Tobin, D. D., Menon, M., Menon, M., Spatta, B. C., Hodges, E. V., & Perry, D. G. (2010). The intrapsychics of gender: A model of self-socialization. *Psychological Review*, *117*(2), 601–622.

Treviño, L. J., Gomez-Mejia, L. R., Balkin, D. B., & Mixon, F. G. (2018). Meritocracies or masculinities? The differential allocation of named professorships by gender in the academy. *Journal of Management*, *44*(3), 972–1000. https://doi.org/10.1177/0149206315599216

Twenge, J. M., & Campbell, W. K. (2001). Age and birth cohort differences in self-esteem: A cross-temporal meta-analysis. *Personality and Social Psychology Review*, *5*(4), 321–344. https://doi.org/10.1207/S15327957PSPR0504_3

Van der Graaff, J., Carlo, G., Crocetti, E., Koot, H. M., & Branje, S. (2018). Prosocial behavior in adolescence: Gender differences in development and links with empathy. *Journal of Youth and Adolescence*, *47*(5), 1086–1099. https://doi.org/10.1007/s10964-017-0786-1

van der Kaap-Deeder, J., Wouters, S., Verschueren, K., Briers, V., Deeren, B., & Vansteenkiste, M. (2016). The pursuit of self-esteem and its motivational implications. *Psychologica Belgica*, *56*(2), 143–168. https://doi.org/10.5334/pb.277

Van Houtte, M. (2023). Understanding the gender gap in school (dis)engagement from three gender dimensions: The individual, the interactional and the institutional. *Educational Studies*, *49*(2), 260–278. https://doi.org/10.1080/03055698.2020.1842722

Vandello, J. A., & Bosson, J. K. (2013). Hard won and easily lost: A review and synthesis of theory and research on precarious manhood. *Psychology of Men and Masculinity*, *14*(2), 101–113. https://doi.org/10.1037/a0029826

Vandello, J. A., Bosson, J. K., Cohen, D., Burnaford, R. M., & Weaver, J. R. (2008). Precarious manhood. *Journal of Personality and Social Psychology*, *95*(6), 1325–1339. https://doi.org/10.1037/a0012453

Vantieghem, W., & Van Houtte, M. (2020). The impact of gender variance on adolescents' wellbeing: does the school context matter? *Journal of Homosexuality*, *67*(1), 1–34. https://doi.org/10.1080/00918369.2018.1522813

Verniers, C., Martinot, D., & Dompnier, B. (2016). The feminization of school hypothesis called into question among junior and high school students. *The British Journal of Educational Psychology*, *86*(3), 369–381. https://doi.org/10.1111/bjep.12111

Vervecken, D., Hannover, B., & Wolter, I. (2013). Changing (s)expectations: How gender fair job descriptions impact children's perceptions and interest regarding traditionally male occupations. *Journal of Vocational Behavior*, *82*(3), 208–220. https://doi.org/10.1016/j.jvb.2013.01.008

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Wach, F. S., Spengler, M., Gottschling, J., & Spinath, F. M. (2015). Sex differences in secondary school achievement: The contribution of self-perceived abilities and fear of failure. *Learning and Instruction*, *36*, 104–112. https://doi.org/10.1016/j.learninstruc.2015.01.005

Wallner, L., & Eriksson Barajas, K. (2023). 'It's not her, it's hen'–situated classroom use of the Swedish gender-neutral pronoun hen. *Classroom Discourse*, *14*(4), 327–343. https://doi.org/10.1080/19463014.2022.2084427

Warner, L. R. (2008). A best practices guide to intersectional approaches in psychological research. *Sex Roles,* 59(5–6), 454–463. https://doi.org/10.1007/s11199-008-9504-5

Whitley, B. E. (1983). Sex role orientation and self-esteem: A critical meta-analytic review. *Journal of Personality and Social Psychology*, *44*(4), 765–778. https://doi.org/10.1037/0022-3514.44.4.765

Whitley, B. E. (1984). Sex-role orientation and psychological well-being: Two meta-analyses. *Sex Roles*, *12*(1–2), 207–225. https://doi.org/10.1007/BF00288048

Whitley, B. E. (1988). Masculinity, femininity, and self-esteem: A multitrait-multimethod analysis. *Sex Roles*, *18*(7–8), 419–431. https://doi.org/10.1007/BF00288393

Wood, W., & Eagly, A. H. (2015). Two traditions of research on gender identity. *Sex Roles*, *73*(11–12), 461–473. https://doi.org/10.1007/s11199-015-0480-2

Xiao, H., Strickland, B., & Peperkamp, S. (2023). How fair is gender-fair language? insights from gender ratio estimations in French. *Journal of Language and Social Psychology*, *42*(1), 82–106. https://doi.org/10.1177/0261927X221084643

Xiao, S. X., Hashi, E. C., Korous, K. M., & Eisenberg, N. (2019). Gender differences across multiple types of prosocial behavior in adolescence: A meta-analysis of the prosocial tendency measure-revised (PTM-R). *Journal of Adolescence*, *77*, 41–58. https://doi.org/10.1016/j.adolescence.2019.09.003

Yang, X., & Gao, C. (2021). Missing women in STEM in China: An empirical study from the viewpoint of achievement motivation and gender socialization. *Research in Science Education*, *51*, 1705–1723. https://doi.org/10.1007/s11165-019-9833-0

Yavorsky, J. E., & Buchmann, C. (2019). Gender typicality and academic achievement among American high school students. *Sociological Science*, *6*, 661–683. https://doi.org/10.15195/v6.a25

Yu, J., McLellan, R., & Winter, L. (2020). Which boys and which girls are falling behind? Linking adolescents' gender role profiles to motivation, engagement, and achievement. *Journal of Youth and Adolescence, 50*, 336–352. https://doi.org/10.1007/s10964-020-01293-z

Zeigler-Hill, V., & Myers, E. M. (2012). A review of gender differences in self-esteem. In S. P. McGeown (Ed.), *Psychology of gender differences* (pp. 131–143). Nova.

Zhang, S., Zang, X., Zhang, S., & Zhang, F. (2022). Social class priming effect on prosociality: evidence from explicit and implicit measures. *International Journal of Environmental Research and Public Health*, *19*(7), 3984. https://doi.org/10.3390/ijerph19073984

Zuckerman, M., Li, C., & Hall, J. A. (2016). When men and women differ in self-esteem and when they don't: A meta-analysis. *Journal of Research in Personality*, *64*, 34–51. https://doi.org/10.1016/j.jrp.2016.07.007

IV. Supplementary Documents

Eigenständigkeitserklärung

Hiermit erkläre ich,

dass ich die vorliegende Dissertation selbstständig verfasst und ohne unerlaubte Hilfe angefertigt habe.

dass ich die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken (dazu zählen auch Internetquellen und KI-basierte Tools) entnommen sind, unter Angabe der Quelle kenntlich gemacht habe.

Alle Hilfsmittel, die verwendet wurden, habe ich angegeben.

Die Dissertation ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, 04.07.2024

Hannah Streck