GENDER GOES TO SCHOOL

THE INFLUENCE OF GENDER NORMS ON EARLY ADOLESCENTS' SCHOOL FUNCTIONING
Gender goes to school:
The influence of gender norms on early adolescents’ school functioning

Wendelien Vantieghem

Submitted to the faculty of political and social sciences of Ghent University in partial fulfillment of the requirements for the degree of Doctor in Sociology

Promotor: Prof. Dr. Mieke Van Houtte
Co-promotor: Dr. Hans Vermeersch
Academic year: 2015-2016
# Table of Contents

Dankwoord.......................................................................................................................... iv

Chapter 1. Introduction........................................................................................................... 1

Chapter 2. The gender gap in education.................................................................................. 5

Chapter 3. Theoretical approaches to gender......................................................................... 9

Chapter 4. Why “gender” disappeared from the gender gap:
(Re-)introducing gender identity theory to educational gender gap research.................... 27

Chapter 5. An integrated model of gender identity and school functioning............................. 53

Chapter 6. Methodology.......................................................................................................... 63

Chapter 7. Transcending the gender dichotomy in educational gender gap research: The association between gender identity and academic self-efficacy.................................................. 75

Chapter 8. Differences in study motivation within and between genders:
An examination by gender typicality among early adolescents............................................. 99

Chapter 9. Are girls more resilient to gender-conformity pressure?
The association between gender-conformity pressure and academic self-efficacy.............. 123

Chapter 10. The impact of gender divergence on adolescents’ well-being:
Does the school context matter?........................................................................................... 149

Chapter 11. Discussion........................................................................................................... 181

References.............................................................................................................................. 203

Summary................................................................................................................................. 225

Samenvatting.......................................................................................................................... 229
List of Tables

Table 1. Gender gap in education in Flanders.................................................................6
Table 2. Comparison of sample and population characteristics...........65
Table 3. Sample characteristics.........................................................................................66
Table 4. School characteristics at Time 1 of all participating schools ....68
Table 5. School characteristics at Time 1 of longitudinally participating schools.................................................................................................................................68
Table 6. Univariate characteristics of gender and school functioning variables........................................................................................................................................70
Table 7. Descriptive statistics for variables: Frequencies (%), means, standard deviation (SD), Cronbach’s alpha and N. .........................................................89
Table 8. Zero-order correlations among mains and control variables ...90
Table 9. Multilevel analysis: Unstandardized coefficients and standard errors in parentheses.................................................................................................................92
Table 10. Descriptive statistics........................................................................................110
Table 11. Zero-order correlations divided by gender.................................111
Table 12. Results of multilevel analysis on autonomous motivation.....113
Table 13. Results of multilevel analysis on controlled motivation........114
Table 14. Demographic characteristics of the participants in the final sample by gender. .......................................................................................................................134
Table 15. Descriptive statistics.........................................................................................136
Table 16. Zero-order correlations at Time 1 divided by gender.............138
Table 17. Zero-order correlations at Time 2 divided by gender.............138
Table 18. Results of multilevel analysis at Time 1: Unstandardized coefficients and standard errors in parentheses .........................................................140
Table 19. Results of multilevel analysis at Time 2: Unstandardized coefficients and standard errors in parentheses .........................................................142
Table 20. Descriptive statistics on the school level. ........................................157

Table 21. Descriptive statistics on the individual level.........................................159

Table 22. Zero-order correlations on the school-level...........................................165

Table 23. Multilevel analyses of gender typicality on well-being: Unstandardized coefficients and standard errors in parentheses ..........168

Table 24. Multilevel analyses of pressure for gender conformity on well-being: Unstandardized coefficients and standard errors in parentheses170
Dankwoord

Mensen zeggen soms dat een doctoraat is als een bergbeklimming. En alhoewel ik niet graag in clichés verval, mag dit in een dankwoord misschien wel eens. Zodus…: Een doctoraat is als een bergbeklimming! We zijn allemaal op weg naar dezelfde top, maar ieder van ons legt een verschillend parcours af. We vertrekken op verschillende plaatsen onderaan de berg en worden geconfronteerd met specifieke hindernissen en uitdagingen. Voor sommigen is de weg langer dan voor anderen. Sommigen worden tijdens de beklimming bijgestaan door een uitgebreid team, terwijl anderen de tocht alleen aanvatten. Maar voor ieder van ons geldt dat bepaalde mensen ons onderweg de nodige hulp, aanmoediging en ondersteuning geboden hebben. Soms worden we eventjes door hen gedragen, soms is deze hulp eerder een schop onder de kont. Hoe het ook zij, dit dankwoord is opgedragen aan al deze mensen: Merci, merci, …MERCI!

Eerst en vooral zou ik graag Promotor Mieke bedanken. De zelf-determinatie theorie stelt dat mensen 3 basisnoden hebben: competentie, autonomie en verbondendheid. Mieke, je hebt me altijd competent laten voelen met een complimentje op tijd en stond, en door me toe te staan bij te leren en me te ontplooien (nóg een cursus statistiek? Waarom niet!). Verder wil ik je ook graag bedanken voor alle autonomie die je me gegeven hebt. Je hebt me gedurende mijn doctoraat steeds de kans gegeven om mijn eigen ding te doen en mijn eigen weg te zoeken, zelfs wanneer die over de oceaan naar het verre Santa Cruz leidde. En het moet gezegd, ondanks dat hoog niveau aan autonomie was er geen gebrek aan verbondendheid. Jouw deur heeft altijd —letterlijk— voor me open gestaan, en je had steeds tijd voor een babbeltje of overleg. Kortom, merci om zo’n goeie promotor te zijn.

I would also like to thank the members of the jury: professor Carolyn Jackson, professor Campbell Leaper, professor Piet Bracke, Professor John Lievens, doctor Jessy Siongers and doctor Hans Vermeersch. Thank you for taking the time to thoughtfully read and judge this dissertation (all 230 pages of it!). It is a relief to know that after years of hard work, this dissertation will be read by at least six people!

A specific thank you as well to professor Campbell Leaper. Cam, thank you for welcoming me to UCSC and Leaper lab. Working with you has been one of the — sunny— highlights of my Ph.D. I had expected to be largely left to my own devices
or to be thrust upon some unexpected Ph.D student, but never in my wildest dreams would I have imagined going on hikes, having breakfast, watching movies and visiting musea with the great professor Leaper! Thank you so much for making me feel at home in Santa Cruz.

Timi, Toni, Veronica & Christy, thank you for the thoughtful feedback in Leaper Lab, but more than anything, thank you for taking that Belgian girl with you to parties and karaoke. Also, thank you to the other UCSC-students: Ben, Graciela, Bryan, Rob, Ketty, Sean, Max, Kyle, Juan, Alicia and Joey, for your kindness, hospitality and general awesomeness. What I had anticipated to be a lonely three months turned out to be an amazing time thanks to y’all. I’ve been California dreaming ever since.

Ook een dikke merci aan de Procrustes-collega’s: Dimi, Ellen, Els, Marian, Nathalie, Sofie, Ankelien en Valérie (alsook, once upon a time, Elke, Veronique en Pieter). Mijn compagnons de route, merci om samen met mij de weg naar de top aan te vatten. Het was een verademing om met zo’n goed team data te mogen verzamelen, onderzoek te mogen doen, en valorisatie-materialen te creëren. Natuurlijk ook enorm bedankt voor de zotte momenten tussendoor, inclusief gendernomiekends, TILF’s, OT’s gevuld met handcrèmes, horoscopen, sexy vodka-posters, roze hemden en Casper & Hobbes. En ik onthoud: Tips & Tricks werken gewoon niet!

Ook bedankt aan de directies, leerkrachten en leerlingen van de scholen die deelnamen aan het Procrustes-project. Merci om ons toe te laten in jullie scholen en u door talloze enquêtepagina’s te worstelen (al dan niet met het nodige gezucht). Zonder jullie was dit onderzoek nooit mogelijk geweest.

Bedankt aan de secretariaatsmedewerkers van de Vakgroep Sociologie om als ware sherpa’s mij te begeleiden op de tocht. Weten dat ik allerhande praktische beslommeringen aan jullie kon toevertrouwen was echt een last van mijn schouders.

Merci aan de vroegere en huidige collega’s van de Vakgroep Sociologie voor de ongedwongenheid, ontspannende koffiepauzes en open-door policy. Dankzij jullie kwam ik telkens graag op bureau, en heb ik nooit de drang gevoeld om Mieke’s regels rond thuiswerk te overtreden.
Natuurlijk ook een hele dikke merci aan mijn bureaugenoten: Fanny, Koen en Sarah! Koen, mijn eeuwige hulplijn bij iedere mogelijke statistische of computergerelateerde vraag. Merci om altijd je geduldige, sympathieke, droog-humoristische en behulpzame zelf te zijn. Zonder jou was mijn doctoraat vast met veel meer frustratie — en mogelijks gênante statistische fiasco’s — verlopen. Sarah, of liever, professor Vandevelde, wat ben ik toch steeds onder de indruk van de manier waarop je internationaal fenomeen & academische hoogvlieger combineert met laid-back mama, vriendin en collega. We hebben enkele van onze gekste/plezantste bureaumomenten aan jou te danken (olympische spelen op google, posters van Ronald Reagan en genocide, de benaming van Truus, Blossom, Fatima & Junior, chippendale-peter vergelijkingen culminerend in een waar erotisch kortverhaal,…). Je wordt gemist!
Fanny, wij brengen al bijna een decennium in elkaars nabijheid door, en daar heeft nog geen dag van verveeld! Merci voor je immer spontane, enthousiaste, blozende aanwezigheid. Merci om altijd tijd te hebben voor een woordvertaling, een snelle afchecking, een onnozel youtube-filmpje en een babbel wanneer de bureau alleen van ons is.

Els, Ad, Karen, Emilien, Laura, Sarah, Rachel, Josephine, Stijn & Robbe: de mensen waarvoor de grens tussen collegialiteit en vriendschap al lang geleden vervaagd is. Merci voor de zomerse ijsjes op de Graslei, de filmavonden, Halloween-vieringen, dansfeesten, ontspannende saunabezoeken, pyjama-party’s en pretpark-adrenaline. Bij zoveel fantastische mensen terecht kunnen voor een ontspannend moment, ongedwongen babbel, steunende knuffel of gierende lachbui heeft mijn werkplek getransformeerd in een veilige haven. Ook al slaat de deur van de Korte Meer nu achter me dicht, ik twijfel er niet aan dat we onze zotte momenten verderzetten!

Roselien, bij veel medewerkers is de grens tussen collega & vriend vervaagd, maar in jouw geval heeft die zelfs nooit bestaan. Merci om al jarenlang mijn partner in (party-)crime te zijn, altij klaar te staan voor een vluggertje over de middag, en simpelweg om zo’n fantastische vriendin te zijn.

Susan, ik ben toegekomen op de Korte Meer als iemand die geen koffie lustte, maar mede dankzij jou heb ik er een (cafeïne-)verslaving bij (yep, I'm blaming you!). Merci voor de talloze koffiepauzes bij ieder etablissement met een espressomachine in de wijde omtrek van de Korte Meer. Inspirerende momenten van sociologische verbeelding, zeker en vast (dat vertellen we in ieder geval steevast aan onze promotoren), maar telkens ook het lichtpunt van mijn namiddagdip! Merci!

Anouk, mi cariña, ik ben nog steeds blij dat we allebei Salamanca als Erasmus-locatie gekozen hebben. En meer dan wat dan ook, ben ik dankbaar dat Mieke ons allebei heeft aangenomen in januari 2012 (Merci, Mieke!). Ik ben zeker dat wij elkaar nooit meer uit het oog verliezen, zelfs al jaag jij je dromen na aan de andere kant van de wereld. Ik zou hier een hele opsomming kunnen maken van de dingen die we samen mee- en doorgemaakt hebben, maar laat ik het kort houden: Merci om te zijn wie je bent.

Kindtie, Greet, Jeroen, Tim, Bart, Sam & Jonas, merci voor alle momenten weg van het doctoraat. De relaxte brunches, lunches, dinertjes en feestjes; de cava’s, wijntjes, ciders en whiskies; de (al dan niet foute) humor. Mijn leven is zoveel kleurrijker met jullie erbij, merci!

Suisse, Mi & Joke, waar moet ik in hemelsnaam beginnen bij jullie? Jaren van vriendschap laten zich niet samenvatten in een alinea. Van de schoolbanken van het SJB tot de nachtelijke straten van Gent, van Salamanca tot Santa Cruz, van chocoladebar naar restaurant, van twijfelachtige hostel tot overnachting in de auto of openlucht… door dik en dun, jullie zijn er altijd geweest. Merci voor altijd te luisteren naar mijn gezaag, te lachen met mijn dubieus gevoel voor (nonkel)humor, mijn onzekerheden te sussen, de schouder te zijn waarop ik kon uithuilen, mij aan te porren tot actie en mijn overwinningen te vieren. Merci voor de vriendschap.

Mama, Papa, Breg en meter Els. Merci om mij te tonen wat doorzetten en hard werk is. Jullie twijfelden tijdens een van mijn relaxte examenperiodes waarschijnlijk wel eens of ik die waarden doorgekregen had, maar door jullie lichtend voorbeeld en door mij altijd te stimuleren, sta ik hier nu vandaag. Merci. De voorbije vier jaar zijn ook niet altijd makkelijk geweest: we hebben slecht nieuws ontvangen en tegenslag te verwerken gekregen. Maar altijd zijn we er geweest voor elkaar en zijn we er samen uitgekomen. Die onvoorwaardelijke steun is niet in woorden te vatten. Merci.
Gender goes to school:  
The influence of gender norms on early adolescents’ school functioning
Chapter 1.
Introduction

Differences between men and women have always garnered attention from both popular media and researchers. This is especially true when the field of divergence is something as central to the life course as education. So, when research in the nineties discovered that girls were outperforming boys on most educational parameters, the “boy problem” captured the attention of popular media, policy makers and researchers alike (Warrington & Younger, 2000). A lot of research has considered aspects that could possibly explain the observed gender-differential achievement patterns, such as gender-specific traits, study behaviors, attitudes or interactions with teachers.

While it has added immensely to our understanding of the gender gap in education, this research is flawed in several respects. Because the research originated from the observation that boys were underachieving, the focus was predominantly on boy issues, while girls were much less considered (Schippers, 2007). By focusing on overall boy-girl differences, intrasexual variation was neglected as well. Nevertheless, average differences in school functioning mask a wide range of variation within the sexes, with some boys excelling and some girls doing rather poorly (Epstein, Elwood, Hey, & Maw, 1998; Gilbert & Gilbert, 1998; Mac an Ghaill, 1994). Furthermore, the functioning of gender is reduced to differences between the sexes. That is, the influence of cultural conventions concerning gender norms is then supposedly reflected in average boy-girl differences, but the more subtle workings of gender in attitudes, ideology, identity, expectations, and actual behavior have been swept under the carpet. This is especially true for quantitative studies that rarely include more nuanced gender aspects besides the sex of the participants.

Hence, the central focus of this dissertation is on uncovering inter- and intrasexual variations in the school functioning of boys and girls by using a gender focus. To do this, we will employ the concepts of gender identity and pressure for gender conformity. First, gender identity—which is operationalized as gender typicality in this research (for reasons discussed in chapter 4)—taps the extent to which people perceive themselves as masculine or feminine, given what it means to be masculine or feminine in a specific society (Perry & Pauletti, 2011; Tobin et al., 2010). According to social identity theory, salient social identities influence what behavior comes to mind in a specific context, and increase the likelihood of
identity-congruent behavior (Oyserman & Destin, 2010). Gender identity is a core social identity, because it is established early in childhood; and because society considers sex an important marker of difference, it is salient in many diverse contexts (Elmore & Oyserman, 2012; West & Zimmerman, 1987). Because of the differences in school functioning between boys and girls, gender identity seems a pertinent concept to consider. Second, we will also consider pressure for gender conformity. That is, despite the fact that identities are key motivators for the enactment of behavior, it would be wrong to think of individuals as being free agents of will, not impacted by societal expectations or disapproval. Indeed, research has shown that many boys face anti-school pressures from their peer group (Epstein, 1997, 1998), resulting in boys trying to balance peer expectations with getting good grades. As such, it is equally pertinent to consider social pressures when considering the enactment of gendered behavior.

The concept of gender identity is tightly interwoven with the concepts of masculinity and femininity (Perry & Pauletti, 2011; Tobin et al., 2010). It refers to the way individuals compare themselves to normative expectations for their gender category, and is therefore located predominantly on the social psychological level. A push towards gender conformity, conversely, refers to socializing pressures from the environment that may be, to a greater or lesser extent, internalized (Adler, Kless, & Adler, 1992; Egan & Perry, 2001), which makes this concept more akin to sociology. By employing these concepts, we integrate ideas more central in social psychological research with those from sociology. This dissertation thus places the individual firmly within its social context, and accounts for the norms and expectations that go with this environment and influence the individual.

A lack of consideration for these concepts in previous—especially quantitative—research stems partly from the lack of a validated, reliable and accepted instrument (see chapter 4). Recently, the self-concept questionnaire by Egan & Perry (2001) has become widely accepted in scientific circles, opening up the possibility to assess the concepts of gender typicality and pressure for gender conformity through survey research. In this light, the central goal of this research is to study the inter- and intrasexual variation in the school functioning of boys and girls by considering the influence of gender typicality and pressure for gender conformity. In line with the expectancy-value model by Eccles and colleagues (Wigfield & Eccles, 2000), we consider several indicators of school functioning. More specifically, we assess the influence on academic self-efficacy (i.e., the expectancy-component), study motivation (i.e., the value-component) and well-being (i.e., the affect-component). By using several indicators, we can give a more extensive and robust overview of the impact of gender aspects on school
functioning, and tap both academic and affective aspects of school functioning. By using this approach, this dissertation contributes to previous scientific research by opening up the gender categories and examining intragender variation, rather than resting upon simple boy-girl differences. This research similarly adds to public and policy concerns by shedding further light on the educational gender gap. Only by opening up the black box of boys’ and girls’ different achievement patterns, can we find ways to combat this gender inequality in our societies.

In what follows, I discuss the history and cross-national evidence of the educational gender gap. I then provide an overview of theories central in explaining gender differences, and more specifically, those often used in an educational context, such as masculinities theory (Connell, 1995) and the identity-based motivation model (Oyserman & Destin, 2010). First, more distal gender theories will be discussed, such as feminist thought concerning patriarchy and heteronormativity (de Beauvoir, 1949; Lorber, 1994; Oakley, 1972), as well as West & Zimmerman’s notion of “doing gender” (1987). These theories have been less directly applied to educational research, but have proven vital in furthering our understanding of gender in society. They provide the framework in which the more proximal theories function, namely masculinities theory, the identity-based motivation model and multifactorial theory (Spence, 1993). These proximal theories are more closely related to educational research and have framed the gender gap by considering gender performances at school. I end the first section by developing a conceptual model that will be applied in the empirical research discussed in the second part of this dissertation.
Chapter 2.
The gender gap in education

The differential educational achievement of boys and girls has been the focus of both popular and scientific attention for decades. Originally, the focus was on the underachievement of girls and women in specific courses, such as science, technology and mathematics (STeM), and the female underrepresentation in different levels of education (Byrne, 1978). In Belgium, for instance, 14- to 15-year-old girls were not equally represented in education until 1969, while it took until 1974 before there was equal school attendance between 16- to 17-year-old girls and boys (Derks & Vermeersch, 2001). The higher the level of education, the longer it took for gender equality to be reached. For example, in Belgium it took until 1980 before boys and girls attended higher education in equal measure.

By the 1990s, however, research revealed that boys were no longer outperforming girls. Two interlocking processes are important in explaining this transformation. First, education expanded worldwide on all levels of schooling during the second half of the 20th century (Meyer, Ramirez, Rubinson, & Boli-Bennett, 1977). Through this educational expansion, women were increasingly able to finish secondary education and enroll in higher education. Hence, women were gradually catching up to boys, who historically had more chances as regards education (Buchmann & DiPrete, 2006; Buchmann, DiPrete, & McDaniel, 2008; Byrne, 1978). Second, a decline in gender discrimination during this period (Brooks & Bolzendahl, 2004; McHugh & Frieze, 1997) reinforced the effects of the educational expansion (Buchmann & DiPrete, 2006; Buchmann et al., 2008). These egalitarian processes not only ensured women’s increased participation in education, but also opened up more options that were previously considered improper or inappropriate. Women were no longer limited to home economics, but increasingly able to study A-levels or “masculine” courses such as mathematics or sciences (Byrne, 1978).

Nowadays in industrialized countries, girls’ historical disadvantage in educational participation has been largely resolved (Buchmann & DiPrete, 2006). In Belgium, for instance, more girls than boys have been enrolled in universities since 1998 (Derks & Vermeersch, 2001). Since girls are now participating in equal, or indeed even greater measure than boys, attention has shifted to gender differences during the school career (Derks & Vermeersch, 2001). This shift prompted investigation into grade point averages (GPA), classroom behavior, school-related
attitudes and grade retention, which revealed that boys tend to be worse off on most of these indicators. In Belgium, boys repeat grades more often than girls (Van Landeghem, Goos, & Van Damme, 2010), have lower marks in class (Van de Gaer, Pustjens, Van Damme, & De Munter, 2006a), are enrolled less often in academic-oriented tracks (Van Landeghem & Van Damme, 2007; Van Woensel, 2007), drop out more often without qualification (Van Landeghem et al., 2010), have lower enrollment in higher education (Van Woensel, 2007) and are overrepresented in special education services (Van Landeghem & Van Damme, 2007) (See Table 1 for detailed figures). Belgian girls also report higher school well-being, less deviant school behavior (Engels, Aelterman, Petegem, & Schepens, 2004), and more positive school attitudes than boys (Van Houtte, 2004b). Notably, these differences tend to be small in elementary school, but gender-differential achievement gaps tend to widen throughout secondary education (Derks & Vermeersch, 2002).

Table 1.

<table>
<thead>
<tr>
<th>Gender gap in education in Flanders</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade retention by end of secondary education, birth cohort of 1984-85(^1)</td>
<td>28.8%</td>
<td>43.1%</td>
</tr>
<tr>
<td>Enrollment in academic tracks in 9th grade, school year 2005-06(^2)</td>
<td>51%</td>
<td>41%</td>
</tr>
<tr>
<td>Enrollment in higher education, school year 2005-06(^2)</td>
<td>67%</td>
<td>55%</td>
</tr>
<tr>
<td>Enrollment in special education at age 11, school year 2006-2007(^3)</td>
<td>8.9%</td>
<td>5.7%</td>
</tr>
<tr>
<td>Drop-out without qualification, birth cohort of 1984-85(^1)</td>
<td>10.7%</td>
<td>17.9%</td>
</tr>
</tbody>
</table>


Belgium is a case in point for a widespread cross-national phenomenon. That is, most industrialized western countries tend to experience similar gender gaps in education that, while small in elementary education, tend to increase throughout secondary education (Buchmann et al., 2008; Fergusson & Horwood, 1997; Voyer & Voyer, 2014). Girls, for instance, do more homework than boys in all subject areas, have higher graduation rates in general programs in secondary education in all OECD-countries, drop out less often without qualification in all OECD-
countries except Turkey, and have higher graduation rates from higher education in all OECD-countries except Japan, Switzerland and Turkey (PISA, 2009). Boys are also overrepresented in special education services in all OECD-countries (Benjamin, 2003). Even in more traditional masculine courses, girls’ disadvantage in western industrialized countries seems to gradually disappear. Where STeM achievement is concerned, PISA research (2009) revealed that gender differences in science achievement tend to be small, and even insignificant in most OECD-countries. While boys often scored higher on mathematics than girls, the difference was smaller than the gender difference in reading and was also not consistently encountered in all OECD-countries. Conversely, girls outperformed boys in reading in all participating countries. What is more, when school marks are considered instead of ability tests (as used in PISA research), the differences are even more pronounced. A cross-national meta-analysis of gender differences in school marks revealed a small but consistent female advantage in all school courses, even mathematics (Voyer & Voyer, 2014). Such differences between ability tests and school marks are often encountered and explained through the consistent effort school achievement requires versus the one-time application of standardized tests (Duckworth & Seligman, 2006). Furthermore, in contrast to standardized tests, school grades reflect the social context of the school and are influenced by pupils’ classroom behavior, test anxiety and teacher bias (Voyer & Voyer, 2014). Overall, it seems that the female advantage in education presents itself as persistent and wide-ranging, encountered across courses, grades and nations.

It is hard to pinpoint a historical moment where boys started to underperform, since the focus on girls’ educational underrepresentation concealed underlying differences in classroom behavior and attitudes. Nevertheless, some research suggests that the “boy problem” is not a recent manifestation, but rather the discovery of an age old process. For instance, Belgian boys have been confronted with higher rates of grade retention since at least the 1950s (Derks & Vermeersch, 2002). Similarly, the meta-analysis of Voyer and Voyer (2014) revealed that differences in school marks have remained stable since, at least, 1914. Furthermore, Cohen (1998) argues that boys have been doing worse in languages and conversational skills than girls since the 17th century. This disparity was not perceived as a problem or indicator of male underachievement, however, because of specific conceptions of masculinity. That is, the ideal male was a taciturn, naturally intelligent, strong and silent gentleman. Conversely, the hard-working, communicative or clever achievements of women or lower-class males were construed as inferior expressions of shallowness or “trying too hard”.

7
Cohen’s (1998) historical account drew attention to how a specific construction of masculinity influences behavior and problem perception. In line with such research, attention shifted from noting the gender disparities in educational achievement, to trying to find the roots of these patterns. More specifically, an important contribution to the study of the educational gender gap has been masculinities theory (Connell, 1995), which investigates how conceptions of masculinity and femininity influence gendered behavior at school. Consequently, in the next section, we will discuss gender theories which have proven influential and central in the study of the educational gender gap.
Chapter 3.
Theoretical approaches to gender

In what follows, we first discuss influential strains of thought in gender theory, before we turn our attention to theories more specifically geared towards framing the educational gender gap. While these overarching theories are less applicable to the research in question, they have profoundly altered the way social scientists think about gender. Consequently, they provide a broad framework for research concerning gender differences, and as such, are fundamental in understanding these proximal theories. Hence, we first discuss more distal theories concerning gender order and heteronormativity in society and how these link to the “doing of gender” in everyday life. Next, we discuss more proximal theories such as masculinities theory, social identity theory, and theories concerning gender identity, which provide insight into the educational gender gap. Note that these theories map onto the distinct levels where gender processes can be analyzed (Lorber, 1994; Risman, 2004; Unger & Crawford, 1993). That is, theories concerning patriarchy and heteronormativity are situated at the institutional level, analyzing processes of power and stratification in societies at large. Conversely, the theory of West and Zimmerman (1987) concerning doing gender is situated at the interactional level, while identity theories are located on the individual level and consider intra-personal processes of self-categorization, social comparison and evaluation. Importantly, a consideration of all three levels is necessary to fully grasp the functioning and influence of gender in society (Risman, 2004).

UNVEILING THE GENDER ORDER IN SOCIETY

Gender differences have historically been explained through a biological framework. This framework posits that evolutionary changes sparked the differentiation in physiology, hormones, and genetics between women and men, leading to notable differences in traits and behavior (Hyde, 2014; Mertens, 2006; Pedersen, Putcha-Bhagavatula, & Miller, 2011). More specifically, the processes of parental investment and sexual selection would have been central in prompting the current distinctness in, for instance, childcare, physique, aggressiveness, promiscuity and nurturance. The idea that differences between men and women, including their social roles and status, would be invariably caused by underlying
biological dimorphism is implicit in such a framework. Consequently, the biological framework is often perceived as advocating that such differences are not only natural, but self-evident, appropriate and beyond social change (Huber, 2007). So, when feminists started to raise questions about women’s place in society, such a biological approach was not compatible with their appeal for change. Moreover, thinkers such as Simone de Beauvoir, Judith Butler and Ann Oakley argued that the biological basis of the sexes is not sufficient to explain the diverse ways masculinity, femininity and the larger gender order have been constructed across time and space (Butler, 1990; de Beauvoir, 1949; Oakley, 1972). Such arguments led to the distinction between the concepts of sex and gender. According to this distinction, sex is a biological basis upon which culture builds a superstructure of meaning and signifiers of difference that can be labeled gender. To explain how such gendered behavior is learned, feminist thinkers built upon socialization theory.

According to socialization theory, the base of gender socialization is laid in the family context during early childhood. This so-called primary socialization works through several processes, one of which is modeling (Mead, 1934; Mead, 1955). Modeling implies that children learn gendered performances by imitating the behavior of role models, such as their mother and father. This behavior then evokes environmental responses of reinforcement and punishment, which further shape gendered enactments. People are rewarded for displaying proper gender behavior by appreciation, recognition and respect, prompting them to reinforce this behavior. When gender norms are violated, however, people are discouraged from repeating the behavior by social sanctions that include disapproval, ridicule, exclusion, or even violence. It should be noted that these socialization processes are not limited to the family context, but continue throughout people’s lives. Principal forces in this so-called secondary socialization are the educational environment and peers (Harris, 1995). Research has extensively investigated, for instance, reinforcement and punishment processes concerning gendered behavior in school contexts (e.g., Collier, Bos, & Sandfort, 2013; Epstein, 1998; Kimmel, 2007). Importantly, gender norms are internalized over the course of time, resulting in people displaying “appropriate” behavior without the spur of external forces (Harris, 1995; Mead, 1934; Mead, 1955). Once internalized, gender role performances become self-evident and feel right and natural, not limiting.

Such insights into the learning process behind gendered behavior allowed feminist scholars to challenge the biological framework, leading Simone de Beauvoir to say: “on ne naît pas femme, on le devient” (1949, p. 245). Other scholars continued down this path by further developing the concept of gender. Notable in feminists’ conceptualization is that gender is seen as both binary and
hierarchical (de Beauvoir, 1949; Lorber, 1994; Oakley, 1972). That is, the gender concept is binary in that masculinity and femininity are defined in relation and opposition to each other. For instance, while femininity is associated with emotionality, masculinity is connected to its antonym rationality. The notion of hierarchy then points out that gender constructions do not just entail difference, but also subjection. That is, not only are gender differences constructed in opposition to each other, they are also valued differently, with the masculine aspect superior to the feminine one. Returning to our example, in western societies, “masculine” rationality is valued as a sign of intellect, sensibleness, brightness and effectiveness. “Feminine” emotionality, conversely, is viewed more negatively as something hysterical, impulsive and exaggerated. According to feminist scholars, such conceptualizations are indicative of the overall subjugation of women in society, reflected in the concept of patriarchy. Patriarchy is understood as a social and symbolic structure of male dominance and the oppression of women (Van Klinken, 2011). Feminist thinkers analyzed the historical construction and significance of this structure, and its ongoing influence on contemporary society (e.g., de Beauvoir, 1949; Scott, 1999). Scholars point out that in such a context of male domination, masculinity is not only established as desirable and superior, but as normative, as well (de Beauvoir, 1949; Kimmel, 1997; Lorber, 1994). As Simone de Beauvoir puts it: “In practice, the relationship between the sexes is not that of two equivalent electric poles. The man represents, in effect, both the positive and the neutral element” (de Beauvoir, 1949, p. 11). Such processes of normalization contribute to making the masculine gender “invisible”, and rendering it objective, true and natural. Such notions are also fundamental to Foucault’s conceptualization of power.

According to Foucault (1976), power should not be regarded as a monolithic entity situated at the organizational level of society, possessed more by some people than others, and functioning mostly through prohibition and law. Rather, power should be conceptualized as diffuse, prevalent in all domains and members of society, observing, regulating and normalizing some types of behavior, while silencing and pathologizing others. In line with feminist thinkers’ ideas concerning the male gender, Foucault notes that such power processes function optimally when they are hidden. “Power is tolerated only on the condition that it masks a significant part of itself. The extent to which it is effective, is proportional to the degree to which it manages to hide its mechanisms.” (Foucault, 1976, p. 88). Hence, the gender order and differences between women and men are all the more accepted and influential when they are seen as valid, natural, and unchangeable. Such conceptualizations also underlie Foucault’s ideas concerning
heteronormativity, which involves the celebration of heterosexuality in the culture and organization of a society (Foucault, 1976). This involves a binary vision on both gender and sexuality that is presented as natural. This way, heteronormativity enforces ideas about what is “normal” and normative, and pathologizes that which is not (Chesir-Teran, 2003; Toomey, McGuire, & Russell, 2012; Wilkinson & Pearson, 2009). Consequently, in heteronormative societies, the differences between men and women are highlighted, heterosexual relationships are celebrated, and gender transgressions sanctioned. Heterosexuality and the construction of gender difference are thereby entwined, and people who display cross-gender behavior are often assumed to be homosexual and vice versa (Pascoe, 2007; Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010). By conflating gender behavior with sexual orientation, the marginalization of both homosexuality and gender non-conformity is enforced. Such processes of stigmatization function as a deterrent to gender-deviant behavior and ensure appropriate conduct in most members of society, leaving the existing gender order intact (Nielsen, Walden, & Kunkel, 2000).

While social sanctions for gender transgressions and a non-normative sexual orientation go both ways (e.g., Jackson & Tinkler, 2007), it should be noted that the backlash towards homosexuals and men displaying cross-gender behavior is usually worse than towards women or lesbians (Herek, 1987; Horn, 2007; Skidmore, Linsenmeier, & Bailey, 2006). The source of this bias is theorized to lie in the lower status of femininity in society. Consequently, when women perform stereotypically masculine behaviors, they enact traits and behaviors that are valued. However, when men are seen to enact stereotypically feminine behaviors, they are displaying gender-incongruent conduct as well as inferior traits, earning further contempt.

With the development of the concept of heteronormativity, the initial focus on female subjugation in patriarchy is widened to include sexual orientation. Similarly, our understanding of the underlying power processes is deepened, while the complexity and nuances in processes of subjugation are recognized. That is, heteronormativity casts light on the way several indicators (such as gender and sexual orientation) intersect to create specific forms of stigmatization, even among men. While such developments have increased our understanding of the power mechanisms maintaining the gender order at an institutional level, the individual and interactional experience tends to disappear in these macro descriptions. This gap has been filled by West and Zimmerman (1987), who explicitly focused on “doing gender” in everyday interactions.
DOING GENDER

The social constructionist view of West and Zimmerman (1987) on gender considers the ways people infuse their everyday behavior and social interactions with gendered symbolic behavior and signifiers. In contrast to previous research, West and Zimmerman (1987) reconceptualize gender not as a role or as situated within the individual. Rather, they see gender as a dynamic part of the social order, which influences and is influenced by everyday social interactions. Hence, women and men do not enter or relinquish specific gendered roles in each situation, nor do they possess essentially different traits. Rather, men and women contextually perform and enact gender in interactions. According to the theory, gender is in essence a social construct, which is (re)created in everyday social interactions.

Even though this framework focused the lens on an interactionist level, the societal level is not relinquished. That is, the “doing of gender” is not ad hoc invented by individuals, but is inspired by cultural definitions of sex-specific appropriate conduct in which people are socialized. More specifically, gender beliefs, that is, widely shared cultural beliefs about the distinguishing characteristics of men and women, function as cultural rules or instructions for the enactment of gender (Ridgeway & Correll, 2004). Since doing gender is in essence an interactional process, gender enactments usually involve the presence of others, either real or virtual (West & Zimmerman, 1987). In these interactions, people evaluate others and expect to be evaluated according to the reigning gender beliefs. Since gender is a fundamental basis for organization in society, it functions as a master identity. That is, gender cuts across situations and is omni-relevant, because any action can be interpreted as exemplifying it. Consequently, people can always be held accountable for the gender appropriateness of their behavior, whether at school, at home or in the street (West & Fenstermaker, 1995). Realizing that their behavior and appearance will be scrutinized according to the reigning gender beliefs, most people carefully and continuously construct their gender enactment to avoid social sanctions for inappropriate conduct. By appropriately doing gender to ensure favorable evaluations, the social convention is confirmed, sustained, reproduced and legitimized.

Despite the focus on the reproduction of the social order, West and Zimmerman (1987) explicitly include the possibility of resistance and transformation. That is, individuals possess the agency to deviate from normative expectations in their gender enactments. While such behavior is usually socially reprimanded, the possibility of change remains. More specifically, West and Zimmerman (1987) argue that when the institutional level (for instance through
judicial or policy change) or social movements explicitly support transgressions of gender norms, change in gender beliefs and the overall performance of gender are possible. “Social change, then, must be pursued both at the institutional and cultural level of sex category and at the interactional level of gender” (West & Zimmerman, 1987, p. 147). Indeed, such changes have happened several times in the course of history, for instance in the 20th century with the establishment of women’s right to vote.

By including resistance and change in their framework, West and Zimmerman (1987) add to previous conceptualizations of the gender order. That is, the possibility of transformation and change was largely absent from socialization theories and most feminist accounts of patriarchy. By opening it up to individual variation and the “doing of difference”, with simultaneous enactments of race and class (e.g., West & Fenstermaker, 1995), West and colleagues break through the monolithic view on masculinity and femininity. That is, early feminist scholars’ accounts of femininity and masculinity were often criticized for their singular view on gender enactment. While different enactments of femininity were emphasized by black feminism (e.g., Crenshaw, 1991; Hull, Bell Scott, & Smith, 1982), other scholars similarly argued for the recognition of distinct performances of masculinity. They advocated for the acknowledgement that patriarchy included, next to the systematic oppression of women, the simultaneous subjugation of non-privileged men (Connell, 1995; Holter, 2005; Van Klinken, 2011). Such critiques and the insights of “doing gender” were used in the construction of masculinities theory, which focused on the construction of masculinities in society. This theory, and the accompanying ethnographic research in schools, will be discussed in the next section. Before we turn our attention to these proximal theories, we would like to make a note on the concepts of sex and gender.

**SEX VERSUS GENDER: A NOTE FROM THE AUTHOR**

The distinction between sex and gender has been immensely valuable from a theoretical perspective (Hood-Williams, 1996). It has helped to uncover social influences on the differentiation between men and women, which was previously considered immutable. This way, it has also contributed to uncovering the processes of normalization and naturalization in the construction of man-woman differences, which helped to obscure the underlying power processes at play. Hence, the concept of gender transformed gender differences into a social scientific research topic, and simultaneously into a working ground for social policy and change. Notwithstanding these invaluable contributions, issues with the
conceptualization of sex and gender have surfaced and remain largely unresolved to this day.

The conceptualization of sex as the biological basis upon which gender constructed cultural meaning, builds upon certain problematic assumptions. First, underlying the distinction between sex and gender is the idea that the biological sphere can be separated from the cultural one (Hood-Williams, 1996). While this distinction has been tremendously helpful from an analytical point of view, in practice, however, the division is severely flawed. While many studies claim to study them separately and independently, different scholars have unveiled that not only is sex gendered, but gender is sexed as well. The former (i.e., when we say that sex is gendered) refers to the fact that the biological study of sex is necessarily influenced by cultural frames of reference. Historical analysis has revealed a previous one-sex model, instead of the current two-sex model (Hood-Williams, 1996). In this model, the feminine body was seen as an intermediate stage of human development, which was epitomized in the male anatomy. Interestingly, the morphologically androgynous development of the human embryo would still be compatible with a one-sex model, and the development of the two-sex model did not seem prompted by greater knowledge but rather by a paradigm shift. Similarly, Butler (1990) argues that surgical alterations to intersex people are spurred by the current dichotomous view on biological sex, whereas other cultures would see them as an independent, legitimate third sex. These examples illustrate that “[…] the meaning of an object comes not from the object itself but from the interpretations placed upon it” (Hood-Williams, 1996, p. 8). The latter (i.e., when we say that gender is sexed) refers to the fact that the study of gender, which should theoretically be independent from biology, remains very much linked to the body (Francis, 2002; Francis, 2010; Harrison, 2006). For instance, the socialization of children into masculine or feminine roles is inextricably bound to the establishment of the child’s sex. Similarly, gender researchers commonly classify men’s behavior as masculine, whereas girls performing non-feminine behavior are rarely categorized within masculine cultures. In short, it seems that while sex and gender are theoretically independent, in practice they are repeatedly conflated with each other.

A second problematic assumption underlying the sex/gender distinction is the notion that sex refers to the biological, and hence immutable aspects of differences between the sexes. Gender, on the other hand, would then refer to the cultural and changeable aspects (Hood-Williams, 1996). This distinction has contributed to most feminists’ aversion of biological studies, since biological explanations of gender differences are experienced as justifications of a gender
status-quo (Huber, 2007). Nevertheless, research is increasingly revealing that such ideas are flawed, and that biology is more open to change than previously thought. Studies have revealed ways in which biology and environment are intersecting to create new outcomes. There is research that shows, for instance, how the presence of a gene moderates the influence of childhood maltreatment on the development of psychopathology (Kim-Cohen et al., 2006). Similarly, several studies have shown how human brains adapt structurally to disabilities or changes in activities (Doidge, 2007). Such findings illustrate that the idea “biology is destiny” is severely flawed. Rather, it seems that biology and culture, and consequently, sex and gender are not independent, nor easily differentiated. As Hood-Williams (1996, p. 14) puts it: “The world may not be parcelled up in this way [in nature and culture]. […] The real difficulty is in assigning cause or consequence to either, discreet realm. In the sterile debates over nature and nurture what ought to be clear is that a most difficult question is what is which.” Or as Unger and Crawford (1993, p. 124) succinctly but elegantly put it: “[…] biological versus social factors. Complex interactions are the rule, not the exception”.

Despite these issues, the use of the sex-gender distinction is still very much in vogue, albeit often spurred by ideological considerations (Unger & Crawford, 1993). More specifically, researchers who wish to emphasize either the cultural or malleable aspects of male-female differences tend to use “gender”, whereas those who wish to stress the biological facets use ”sex”. Others seem to feel that the use of “sex” has somehow become politically incorrect, and hence, resort to the concept of gender to describe any male-female differentiations. While not as tragic as the study where rats had a “gender” (as described by Unger & Crawford, 1993), a lot of researchers claim to research gender when they simply fall back on the biological sex of the respondent. Considering the nuances, contradictions and complexities in the cultural differentiations of masculinity and femininity, such use seems a crude reduction of the gender concept. Rather, by claiming to do gender research when simply considering the reported sex of the participants, we feel that biology and culture are jumbled together into average girl-boy differences, without considering either sex or gender in a more nuanced and thorough way. Consequently, in this dissertation, we use the word sex when referring to overall comparisons between people based on biological sex, whereas we will employ the gender concept when considering more nuanced aspects of superstructure, culture or socialization.
MASCULINITIES THEORY

Theoretical Development
Connell (1995) is the main proponent and developer of masculinities theory, and advanced a vision on gender relations by highlighting two central issues. First, building upon the patriarchy analyses by feminists, Connell conceptualized external hegemony (Demetriou, 2001). By recognizing the power differentials central in the construction of the gender order, this concept refers to the dominance of men over women in society, and hence the simultaneous dominance of masculinity over femininity. The dominant form of masculinity, called hegemonic masculinity (see below), consists of practices that maintain the subordination of women in society. What is crucial in the dominant conceptualization of masculinity and femininity in society is the recognition that gender is always relational. That is, masculinity is constructed in opposition to femininity, and, as such, traits that are predominantly associated with one gender tend to invoke the opposite association for the other gender. For instance, males are considered to be active, dominant, logical, forceful, self-reliant, whereas women are seen as passive, subordinate, emotional, peaceful, co-operative and so on. This conceptualization harks back to Judith Lorber’s famous paper entitled “night to his day” (Lorber, 1994), which also drew attention to the diametrical relationship between the male and female, as well as the “othering” central in the work of Simone de Beauvoir (1949).

The second process Connell (1995) calls attention to is called internal hegemony (Demetriou, 2001). While external hegemony focuses on a hierarchy in the gender order with masculinity superseding femininity, internal hegemony highlights the existence of a hierarchy within masculinities. This concept builds upon the insights of black feminism (e.g., Crenshaw, 1991), which recognized the intersecting ways in which, among others, class and race contribute to the doing of gender. Consequently, masculinities theory recognized four different types of masculinities, which are in hierarchical placement to each other, with hegemonic masculinity at the top.

Hegemonic masculinity represents the dominant ideal of masculinity in a given culture (Connell & Messerschmidt, 2005). The content of hegemonic masculinity is not static, but influenced by changes in time and place. In contemporary western societies, hegemonic masculinity refers predominantly to authority, dominance, status, control, independence, heterosexuality, competitiveness and aggression. A key insight into the concept of hegemony is that it refers to a process of dominance and subordination. This is not so much reflected in the use of brute force, but in keeping with Foucault’s (1976) ideas on power, works rather through cultural acceptance, institutionalization, discursive
centrality and the marginalization of alternative versions of masculinity. It is similarly important to recognize that hegemonic masculinity as such does not refer to a specific constellation of characteristics most prevalent among males in a particular society. Rather, echoing insights from West and Zimmerman’s (1987) interactional perspective, hegemonic masculinity is a way of doing masculinity that is regarded as the cultural ideal, something to strive for in a certain region. For instance, this type of masculinity is exemplified through certain famous proponents, such as athletes, movie stars, fictional characters, war heroes, and so on (Connell & Messerschmidt, 2005). Masculinities theory thus points at the existence of a certain ideology of masculinity that is configured as the ideal for which to strive within a certain culture. Research, predominantly ethnographic research (e.g., Mac an Ghaill, 1994; Martino, 1999; Renold, 2001), then investigates the way this hegemonic masculine ideal type inspires behavior in boys and men in real-life settings, and similarly leads to mismatches, tensions, or resistance. By recognizing the issue of tension and resistance in doing gender, Connell (1995) distinguishes three other types of masculinities besides hegemonic masculinity. These other three are complicit, subordinate and marginalized masculinity (Connell, 1995).

With complicit masculinity, Connell (1995) parries the critique inherent in the concept of hegemonic masculinity. That is, if hegemonic masculinity refers to an ideal type that is rarely achieved, where does that leave the majority of men? Connell (1995) places these men firmly within complicit masculinity. Their production of masculinity is complicit in that they agree with and benefit from the hegemonic ideal, but do not actively embody hegemonic masculinity. These men do not personify the uncontested display of authority that is prevalent in hegemonic masculinity, but still profit from the patriarchal dividend; that is, the benefits resulting from the overall dominance of men over women in society.

A third form of masculinity is subordinate masculinity. As is clear from the name, subordinate masculinities are inferior embodiments of masculinity placed at the bottom of the male gender hierarchy. This indicates that dominance is not only exalted upon femininity, but among groups of men as well. The most frequently cited case of subordinate masculinity in contemporary European and American society is the homosexual. The oppression of gay men is evident in material practices, such as street violence, cultural abuse, and discrimination in both informal and legal ways. The insights from previous work on heteronormativity and marginalization are integrated in Connell’s ideas on subordinate masculinity. For instance, by conflating subordinate masculinity with femininity — indeed, the stereotype of the effeminate gay is widespread and evident in several homophobic
slurs (Collier et al., 2013)—, these divergent enactments of masculinity are symbolically placed in an inferior position (Connell, 1995; Connell & Messerschmidt, 2005).

The last type of masculinity that Connell distinguishes is marginalized masculinity. Classed and racialized expressions of masculinity are placed within this category. Hence, marginalized masculinity encompasses highly masculine enactments that are nevertheless ostracized based on their powerlessness in overall society. That is, they are unable to fulfill claims to hegemony because of the lack of power and resources inherent in their class or racial position. In other words, they are unable to achieve the cultural validation and legitimization necessary to achieve a hegemonic position in society at large.

**Application to the Educational Setting**

Research into boys’ relative underachievement focused in the first place on the school resistance of working class boys. Willis’ depiction of how “lads” obtain working-class jobs in “Learning to labour” has become a classic in educational sociology and led to the development of resistance theory. Willis proposed that the values of the working class clashed with middle-class norms dominant in most schools, resulting in open school opposition among working-class boys (Willis, 1977). Through a process of reduced educational engagement and confrontations with teachers, the rebellion of these so-called lads tends to hamper their academic aspirations and achievement.

While Willis’ work has been of central importance in the development of resistance theory, it also suffered some sharp critique. One critic posited that Willis had mistaken the central explanatory mechanism of gender for class. More specifically, Davies (1995) argued that gender and the previous school career were more central than class background in explaining reduced aspirations, disaffection and delinquency among students. In this view, school underachievement fuels the development of traditional gender roles in students, because conventional adult roles function as a self-defense mechanism against the loss of a learner identity and the negative influence of school failure on self-esteem. That is, Davies (1995) argued that traditional masculinity and femininity redirected attention from unattainable educational goals to other sources of pride and achievement, with girls emphasizing future roles as a mother and wife, whereas boys turned toward masculine antagonism and physical labor.

Such critiques can be situated within masculinities theory, which posits that working class boys do not oppose the middle-class ethics of schools. Rather, all boys regardless of socio-economic background ostensibly oppose the values of the
school institute (such as obedience, hard work, punctuality, tidiness and discipline), which are perceived as feminine. Various ethnographic studies, based in masculinities theory, have uncovered how school’s gender regimes give rise to specific constellations of masculinities, and how the enactment of these distinct masculinity types relates to educational achievement (e.g., Connell, 1989; Mac an Ghaill, 1994; Martino, 1999; Morris, 2012). A common theme in these field studies is how contesting femininity is quintessential to doing hegemonic masculinity. In school settings, this takes the shape of opposing certain courses (such as language, arts, or home economics; Martino, 1996), study behaviors (such as sitting down to read and study, being silent, tidy and punctual; Epstein, 1998), and even the school institute itself. Heyder and Kessels (2013) showed that 15-year-old German boys who identified as stereotypically masculine, rated school as more feminine. This mismatch between self-concept and environment then resulted in lower scores on language.

In short, masculinities theory argues that, because masculinity is constructed in contrast to femininity and because school is often deemed feminine, boys’ identification and engagement with school is hampered. Ultimately, these processes of doing gender in the school setting bring about the educational gender gap, with boys underperforming when compared to girls.

As has frequently been noted by researchers, however, gender performance is fraught with tensions, contradictions and resistance (e.g., Francis, 2000, 2010). Consequently, there is considerable variation among boys in the extent to which they conform to or strive for the hegemonic ideal. Such non-conformity, however, carries considerable social costs, with children being teased, bullied and ostracized (Epstein, 1997, 1998; Swain, 2005). Because of the conflation between femininity and homosexuality, boys who violate gender norms are often targeted with homonegative slurs, regardless of their underlying sexual orientation. Such homonegative teasing is considered by adolescent boys to be among the most degrading insults (Plummer, 2001). Consequently, most boys try to avoid such harassment by portraying a masculine image and steering clear of activities perceived as feminine. Ethnographic studies have provided several examples of such strategies. For instance, boys tend to compensate getting good grades by engaging in masculinizing practices, such as sport accomplishments (Francis, Skelton, & Read, 2010), acting as the class-clown, being silent and disengaged in the classroom (Jones & Myhill, 2004; Renold, 2001), or by claiming that their good grades result from natural talent, not hard work (Mac an Ghaill, 1994). Other boys completely disengage from school and focus on achieving popularity in the peer
group through risk behavior, challenging school authority and emphasizing sportive and sexual exploits (Jackson, 2003; Martino, 1999).

There are marked similarities between this latter masculine enactment and the lad-culture described by Willis (1977). Indeed, Connell (1995) recognizes that class and race intersect with masculinity performance in her conceptualization of marginalized masculinity, and has distinguished classed masculinity performances in her own ethnographic work (i.e., Connell, 1989). Importantly, Connell (1989) points out that there is no such thing as “a marketplace of masculinities”, indicating that students do not freely choose between different forms of masculinity. Rather, boys navigate doing masculinity within the constraints imposed upon them by their bodies and the power relations in society. For instance, since educational achievement is more easily attained by middle class children, distinct types of masculinity are often constructed around the relation with educational aspirations and accomplishment. Consequently, depending on the type of masculinity a boy enacts, different sources of masculine pride are emphasized (Connell, 1989; Jackson & Nyström, 2014; Mac an Ghaill, 1994). For example, middle class boys are better able to integrate educational achievement into their masculinity enactments by emphasizing mental dominance through effortless achievement, whereas working-class boys might emphasize physical dominance and disengage more openly with school goals and required behaviors.

**Issues with Current Research**

The application of masculinities theory to school settings has provided valuable insights into the processes underlying the educational gender gap. Nonetheless, some issues remain with how the theory has been applied in educational research. For one, as is already evident from the theory’s name, the focus is on boys and masculinities. Conversely, femininities have remained undertheorized and underanalyzed (Connell & Messerschmidt, 2005). Schippers (2007) added to the framework by delineating emphasized femininity and pariah femininities. Emphasized femininity, on the one hand, would be complementary to hegemonic masculinity, and thereby legitimize the hierarchical relationship between the genders. Pariah femininities, on the other hand, would contaminate the hierarchical and complementary relationship between the genders by embodying masculine characteristics. Yet, little applied research has focused on girls’ situations or included Schippers’ insights (for an exception, see Pomerantz & Raby, 2015). Hence, one half of the study into the educational gender gap, that is the school functioning of girls, is compromised.
Second, masculinities theory has been predominantly studied with an ethnographic method. While this has provided rich and detailed descriptions of school life, each account tends to stay firmly linked to a specific time and setting. Furthermore, each field study describes slightly different embodiments of masculinity, and gives each a distinct name (Huyge, Van Maele, & Van Houtte, 2015). Mac an Ghail (1994) distinguished, for instance, the real Englishmen, new entrepreneurs and macho lads, Lyng (2009) described the golden boys, geeks, and nerds, Martino (1999) talked about the squids, poofers and party animals, while Connell herself (1989) wrote about the cool guys, swots and wimps. As Francis argued on several occasions (e.g., Francis 2000, 2010), the use of typologies risks reducing the analysis to different sorts of masculinity, obscuring the possible continuity in gendered practices and concealing differences between boys in each category.

Third, masculinities theory is used as a post-hoc explanation of boys' achievement. That is, research that applies masculinities theory tends to divide the school population into several masculine peer groups connected to specific study or peer cultures. However, the explanation of why a boy is part of a specific study culture is characterized by a circular argument. Boys are categorized into specific masculine enactments because of their study attitudes, achievement, gender ideology, and so forth. But, at the same time, it is assumed that boys acquire these characteristics because they are part of these groups. Hence, researchers tend to classify underachieving boys in the “lads”-category, but are unable to predict beforehand where each boy will go. As such, masculinities research lacks a predictive quality.

We argue that a masculinities theory-based analysis of the educational setting with a quantitative method, focusing equally on boys and girls, and with the use of continuous instead of categorical measures, could prove an interesting addition to the current ethnographic research dominating the field. Since the detail and richness of ethnographic research is hard to transpose to quantitative methods, we propose to focus on certain variables. More specifically, we suggest that the concepts of gender identity and pressure for gender conformity could be fruitful in such an endeavor. On the one hand, pressure for gender conformity captures the strong influence of the peer group on adolescents’ gender enactment, as extensively discussed in (ethnographic) studies of study cultures (e.g., Renold, 2001, 2004; Stoudt, 2006; Swain, 2005). On the other hand, gender identity taps individuals’ assessment of their masculinity and femininity, given the dominant definitions in a specific society (Perry & Pauletti, 2011; Tobin et al., 2010). By acknowledging the societal role in the construction of self-perceived masculinity and femininity, the
concept of gender identity connects to conceptualizations of hegemony prevalent in masculinities theory. In what follows, we discuss social identity theory and the identity-based motivation model. These models provide a framework for quantitatively testing the importance of gender identities in a school setting.

GENDER AS A SOCIAL IDENTITY

Social Identity Theory
While masculinities theory has provided important insights into educational processes on a macro as well as an interactional level, Sheriff (2007) argues that the underlying social psychological processes have largely been ignored. Because this hampers our understanding of the individual variations in masculine enactment, Sheriff (2007) suggests that social identity theory could provide a valuable addition to educational gender gap research. In what follows, we discuss the central tenets of social identity theory and the identity-based motivation model. The application to the study of gender identity in an educational setting is discussed in chapter 4.

Social identity theory focuses on how people’s membership of social groups influences cognitions, emotions and behavior (Stets & Burke, 2000b). The two main processes in social identity theory are self-categorization and social comparison. First, self-categorization refers to classifying oneself — and others — in different social groups (Hogg, Terry, & White, 1995; Stets & Burke, 2000b). When people self-categorize as part of a social group, the defining characteristics of the group are made part of the self-concept and guide ensuing behavior. Individuals are members of several social groups, such as gender, race, class, nationality, music taste, sport teams, and so on. Persons who are similar to the self are labeled in-group, and similarities between in-group members are accentuated. Second, self-categorization is followed by social comparison in which the characteristics of social groups are compared (Stets & Burke, 2000b). While similarities between in-group members are emphasized, differences with the out-group are accentuated. Social comparisons of relevant group characteristics tend to be more favorable for the in-group than the out-group, boosting the self-esteem of in-group members. Importantly, social identities would only be influential on perceptions and behavior when they become salient in a context (Hogg et al., 1995; Stets & Burke, 2000b).

When we apply these notions to gender groups, we can see the self-categorization process in the classifications people routinely make in men and women (Goffman, 1977). Indeed, sex functions as a primary group marker (together with age and race) from a very young age (Harris, 1995). Inherent in this
self-categorization as male or female is the concurrent demarcation of which characteristics are typical for the social group (Hogg et al., 1995; Stets & Burke, 2000b). A woman may consider, for instance, fashionableness, caring, talkativeness, and organization skills as defining characteristics of women. These traits then guide not only her own behavior, but also the behavior she expects to see from other women (Hogg et al., 1995; Stets & Burke, 2000b). This process of self-categorization is then followed by a process of social comparison with relevant out-groups, such as men. Differences with the out-group are routinely accentuated and comparisons are usually in favor of the in-group. Consequently, the woman in our example may consider aggressiveness, sloppiness, self-involvement and taciturnity as typical masculine traits, which most men would have in spades, especially when compared to women. Importantly, gender would only become such a notable basis for self-categorization and social comparison when it becomes salient in context. This implies that when other social identities are dominant for group differentiation, sex-typed behavior would be minimized. This has, for instance, been demonstrated for groups where the distinction between children and adults (i.e., age) is more dominant than between boys and girls (i.e., gender) (Harris, 1995), and by the observation that behavioral differences between boys and girls are minimal when observed individually, but become pronounced in sex-segregated playing groups (Maccoby, 1990).

Social identity theory has focused especially on intergroup biases, with studies expanding on different displays of in-group favoritism and out-group derogation (Hogg et al., 1995; Stets & Burke, 2000b). As Sheriff (2007) showed, these mechanisms parallel those described in masculinities theory, with differences between gender groups being exaggerated to construct a more distinct group identity, ascribing more positive characteristics to the in-group than to out-groups (note that these out-groups include the other gender, as well as within-gender groups that deviate from hegemonic norms), and displaying prejudiced attitudes and discriminatory behavior toward out-groups, including verbal taunting and physical abuse of, for instance, effeminate boys.

An interesting extension of social identity theory for the application of masculinities theory to the educational setting, is the identity-based motivation model (IBM). The identity-based motivation model builds upon social identity theory by equally assuming that the traits associated with salient identities guide people’s behavior (Oyserman & Destin, 2010). The model adds to social identity theory by expanding upon intergroup bias to an advanced understanding of identity content and construction. Furthermore, this model is especially interesting for
educational contexts, because it considers behavioral motivations and the influence of perceived task-difficulty. We discuss this model in the next section.

The Identity-based Motivation Model

The identity-based motivation model (IBM) builds on three central postulates: action-readiness, dynamic construction and interpretation of difficulty (Oyserman & Destin, 2010). Action-readiness refers to the core assumption of social identity theory, that identities cue behavioral readiness and understanding of the world in terms of norms associated with the identity. These relevant actions and understanding are dependent on the identity-content, which IBM claims is dynamically constructed. So, even though identities tend to feel stable and context-independent, IBM states that what identities mean to an individual, and hence what the associated behavior is, is constructed within a specific context. The last postulate refers to interpretation of difficulty, indicating that when an action is identity-congruent, difficulties with performing the behavior tend to be interpreted as an indication of importance, not impossibility.

We can apply these postulates to gender performances in educational settings. First, the postulate of action-readiness implies that, to the extent that a feminine gender identity is stereotypically linked to being well-behaved, dutiful and cooperative (Beaman, Wheldall, & Kemp, 2006; Jones & Myhill, 2004), girls’ chances of demonstrating these behaviors at school are raised. Conversely, when masculine gender identities are linked to being an academic underachiever, class clown, or effortless achiever (Francis, 2000; Mac an Ghaill, 1994), boys’ chances of academic engagement are reduced. Second, the postulate of interpretation of difficulty indicates that when boys feel that studying is not “masculine” or typical for boys, they will interpret difficulty with homework as a signal that studying does not come naturally to them. Conversely, if studying is considered to be part of the gender identity, difficulty can be interpreted as a sign that the task is worthwhile, not impossible. Third, the postulate of dynamic construction implies that the traits individuals associate with gender groups are malleable and context-specific. This is in line with masculinities theory-based research, which shows how specific gender regimes at school give rise to different emphases in the enactment of masculinity (Connell, 1996; Swain, 2005). Elmore and Oyserman themselves (2012) demonstrated this malleability through experimental manipulation with priming cues. They showed that by connecting their gender category to success, early adolescents held higher academic aspirations and boys persisted longer at school tasks.
In short, the identity-based motivation model is compatible with masculinities research, and draws attention to micro-level socio-psychological processes in gender enactment. It also adds to the ethnographic research in the field by using more quantitative methods such as experimental manipulation, and testing the viability of interventions. By focusing resolutely on social identities as a whole, in this case gender identity, the scope is broadened to include femininities as well as masculinities. This is an important addition to the previous research, since femininities have received little attention (Schippers, 2007). Nevertheless, some issues remain. First, the model tends to focus on the impact of class and ethnic identities (e.g., Oyserman, Bybee, & Terry, 2006; Oyserman & Destin, 2010), and has rarely been applied to gender identity (for an exception, see Elmore & Oyserman, 2012). Second, a common oversight with masculinities theory is that intrasexual variation has received little attention. As is the case with ethnographic research, this issue stems from from the applied research methods. That is, the identity-based motivation model has been mostly used in experimental research where certain identities are primed, or in interventions which adapt identity contents (e.g., Elmore & Oyserman, 2012; Oyserman et al., 2006; Oyserman & Destin, 2010). In this way, average changes in boys’ versus girls’ school functioning are assessed, while intrasexual variation is ignored. Remarkably, the model does provide ways to study intrasexual variations via the postulate of dynamic construction. This postulate states that what identities mean and which behavior they cue depends on the traits a person associates with the social group. As such, intrasexual variation in identity content is explicitly part of the model, yet it has remained understudied where gender identity is concerned (for an application of intragroup differences in ethnic identity, see Oyserman, Kemmelmeier, Fryberg, Brosh, & Hart-Johnson, 2003). A way to remedy this gap could be to study gender identity through survey research, since surveys provide data with which both average trends and individual variations can be assessed. In the next section, we discuss historical evolutions in relevant theories and survey measures concerning gender identity.
Chapter 4.
Why “gender” disappeared from the gender gap:
(Re-)introducing gender identity theory to educational gender gap research.

Wendelien Vantieghem, Hans Vermeersch, & Mieke Van Houtte
Published in Social Psychology of Education, 17, 357-381.
Doi: 10.1007/s11218-014-9248-8

Educational gender gap research tries to explain the differential achievement of boys and girls at secondary school, which manifests in many western countries. Several explanatory frameworks are used for this purpose, such as masculinities theory. In this review article, the history of educational gender gap research in Anglo-Saxon literature and problems with the contemporary approach are discussed. It is argued that gender identity theory could prove valuable both in furthering educational gender gap research and mitigating several problems with masculinities theory. To this end, an overview of the history and recent developments in gender identity theory is given. The scarce research combining educational gender gap research with gender identity theory is reviewed. Possible contributions, assets and research questions from gender identity theory to educational gender gap research are discussed.
INTRODUCTION

Ever since boys and girls shared a single classroom, their differences in interests and achievement have been the focus of both popular discussion and scientific research. While the original emphasis was on the low achievement of girls and their underrepresentation in mathematics and science (Byrne, 1978; Foster, Kimmel, & Skelton, 2001), the current focus has shifted to the “boy problem”. This shift occurred in the nineties when studies in western industrialized countries showed that girls had started to outperform boys in several domains, including lower dropout, better test scores, and more enrolment in higher education (Connell, 1996; Fergusson & Horwood, 1997; Kleinfeld, 1999; Younger & Warrington, 1996).

Educational scientists have tried to explain the gender gap in education through several theories and frameworks, such as innate traits (Cole, Martin, Peeke, Seroczynski, & Fier, 1999; Duckworth & Seligman, 2006), a tendency for “laddish” behavior (Fergusson & Horwood, 1997) and an overall masculinity culture (Connell, 1989; Francis, 2000; Jackson, 2003). Nevertheless, in the past decades too little research has attempted to transcend the gender dichotomy and too many explanations have stayed firmly linked to sex categorizations and distinctions, thus ignoring unifying mechanisms across the sexes and reducing intra-sex differences.

As a new way of approaching the educational gender gap, this review article would like to propose gender identity theory. On the one hand, it is quite clear that the gender gap in education is an important problem. Too many boys become demotivated and uninterested in the course of their school career, thereby throwing away chances of self-development and acquiring diplomas. Only by understanding how this gender gap originates can we develop ways to combat this waste of talent in our societies. We believe that gender identity theory can be an extra pathway to explaining and therefore reducing the gender gap. Next to this community-oriented incentive, we believe that gender identity theory has been isolated from other research areas for too long and that the time has come for gender identity theory to contribute to applied research. After all, there are several strongpoints to this theory. For example, it classifies people according to how “masculine” or “feminine” they are, thereby paying attention to both sides of the gender order, while being able to separate this from purely biological sex. It can be used not only as a qualitative method, but a quantitative one as well. Furthermore, gender identity does not have to be utilized solely as a post-hoc clarification, but can have predictive and explanatory power as well. Where masculinity theories tend to
explain processes on the social and cultural group level, gender identity theory can be used to discern inherent traits on the individual and personality level.

The goal of this review article is to show that gender identity theory could prove to be a valuable contribution to educational gender gap research. To this end, we will provide an overview of relevant scientific literature to support this argument. We start by giving a historical overview of the educational gender gap and its accompanying scientific literature. We will then discuss the history and recent developments in gender identity theory. Ultimately, we will make a case for linking both research fields by reviewing and unifying the scarce literature that exists up to this point and by indicating possible contributions, assets and research questions.

The reader should note that this review will primarily consider secondary school students, unless stated otherwise. Even though aspects of the educational gender gap are noticeable in higher education (Buchmann et al., 2008; Marrs & Sigler, 2012; Van Woensel, 2007) and as early as primary education (Buchmann et al., 2008; Derks & Vermeersch, 2001; Fergusson & Horwood, 1997), most research has focused on gender differences in secondary educational achievement. Furthermore, although aspects of the educational gender gap manifest throughout the western world (Martínez, Julia, Mari-Klose, & Mari-Klose, 2012), we have decided to limit our scope for practical and consistency reasons. So, unless stated otherwise, we will mainly consider Anglo-Saxon literature, with studies based in Canada, Australia, New-Zealand, the UK and the USA. We will also cite several Belgian-based studies. Belgium occupies place 12 on the gender inequality index, whereas the Anglo-Saxon countries occupy places 18 to 42 (UNDP, 2011). Belgium is thus more gender egalitarian than the Anglo-Saxon countries and provides an example demonstrating the pervasiveness of the educational gender gap, despite progressive gender beliefs. Since we focus on literature from these countries, the reader should thus always keep in mind that effects may differ in other countries and cultures.

GENDER GAP IN EDUCATION

Historical Overview

For several decades, differences between boys and girls in educational achievement have been the focus of scientific attention. In the seventies, the emphasis was on the low achievement of girls. Girls had lower scores and participation in science and mathematics when compared to boys (Byrne, 1978; Foster et al., 2001). Researchers posited several explanations for these findings, which could be divided
into two categories. Firstly, girls were positioned within a deficit framework (Anyon, 1983; Foster et al., 2001; Hodgetts, 2008; Spence, Helmreich, & Stapp, 1975). Their poor performances were attributed to a lack of certain qualities, such as low confidence, high anxiety and fear of success. Furthermore, girls’ tendency to be compliant in class, work hard and hand in homework that was neat and on time, was taken as a sign of a passive, compliant and malleable learning style (Hodgetts, 2008). This was contrasted with boys’ active and curious learning style, which would be focused on understanding rather than on achievement or teacher expectations. This way, girls’ attitudes and behavior in school were taken to be a sign of an inferior learning style (Byrne, 1978).

Secondly, several researchers pointed to factors other than innate traits that impeded girls’ academic success. They suggested studying the impact of societal norms and expectations of gendered behavior (Byrne, 1978; Skelton & Francis, 2011). For instance, research pointed out that textbooks lacked positive role models for girls, and that boys dominated the classroom and teacher attention. These researchers declared that it was therefore no surprise that girls would be less confident than boys. Policy recommendations based on this literature focused on changing learning material and raising teachers’ awareness of inequalities in their classroom management by letting girls answer questions as often as boys, debunking gender stereotypes concerning mathematics and science, and so on.

However, from the nineties on the attention shifted from girls’ low achievement to that of boys. Research revealed that, contrary to expectations, girls were outperforming boys in several domains of schooling, with boys repeating grades more often (Fergusson & Horwood, 1997; Van Landeghem et al., 2010), having lower grades (Duckworth & Seligman, 2006; Epstein et al., 1998; Fergusson & Horwood, 1997; Jackson, 1998; Van de Gaer et al., 2006a; Younger & Warrington, 1996), dropping out more often (Buchmann et al., 2008; Fergusson & Horwood, 1997; Van Landeghem et al., 2010), having lower enrolment in higher education (Buchmann et al., 2008; Van Woensel, 2007) and being overrepresented in special education services and remedial classes in all OECD-countries (Benjamin, 2003). Even in the traditional “masculine” fields of mathematics and science, girls were gradually catching up to boys, resulting in very small or even insignificant gender differences in these courses (Fergusson & Horwood, 1997; Gunderson, Ramirez, Levine, & Beilock, 2012; Jackson, 1998; Younger & Warrington, 1996). These revelations instantly turned “the boy problem” into a hot topic for both policymakers and scientists.
**Explanations for the Gender Gap**

Just as was the case when trying to explain why girls underperformed, researchers have investigated both individual-level and societal factors. When considering the societal factors, two important but interlocking processes appear to be important. Firstly, there is the expansion of education, which has taken place worldwide on all levels of schooling during the second half of the last century (Meyer et al., 1977). Through this expansion of education, women were increasingly able to finish secondary education and enroll in higher education. Hence, women were gradually catching up to boys (Buchmann & DiPrete, 2006; Derks & Vermeersch, 2001), who historically have had more chances of education (Buchmann & DiPrete, 2006; Buchmann et al., 2008; Byrne, 1978). Secondly, a decline in gender discrimination during this period (Brooks & Bolzendahl, 2004; McHugh & Frieze, 1997) reinforced the effects of the educational expansion (Buchmann & DiPrete, 2006; Buchmann et al., 2008). These egalitarian processes not only ensured women’s increased participation in education, but also opened up more options that were previously considered improper or inappropriate. Women were no longer limited to home economics, but increasingly able to study A-levels or “masculine” courses such as mathematics or sciences (Byrne, 1978). Nowadays in industrialized countries, girls’ historical disadvantage in educational participation has been largely resolved (Buchmann & DiPrete, 2006; Derks & Vermeersch, 2001). Thus, attention has shifted to gender differences during the school career (Derks & Vermeersch, 2001), which has prompted investigation into grades, drop-out, and school track and individual-level factors that could explain these differences.

When considering individual level-factors, researchers have investigated certain traits, qualities and behaviors of boys that could lead to their low achievement. Since most researchers would agree that intelligence does not vary between the sexes in a manner that is sufficiently consistent to influence academic performance (Duckworth & Seligman, 2006; Fergusson & Horwood, 1997), research has focused on non-cognitive skills that would explain the gender gap in school performances. For instance, Belgian and British research found that boys had more negative attitudes towards school and were less motivated than girls (Van de Gaer et al., 2006a; Van Houtte, 2004b; Warrington, Younger, & Williams, 2000). Furthermore, Anglo-Saxon research has consistently shown that boys were more inattentive and exhibited more disruptive behavior in the classroom (Fergusson & Horwood, 1997; Francis, 2000; Warrington et al., 2000; Younger, Warrington, & Williams, 1999). They were also less self-disciplined when it came to homework (Duckworth & Seligman, 2006), overestimated their own abilities and tended to attribute their successes to talent and ability, rather than hard work (Cole et al.,
1999; Gunderson et al., 2012; Meece, Glienke, & Burg, 2006). This demotivation, disruptive behavior, overestimation of own abilities and lack of self-discipline would then all contribute to the gender gap in school performances.

While this research does provide an answer to why boys are underachieving, the focus on traits and behavior tends to keep the explanation on the descriptive level and does not explain the origins of these behaviors and attitudes. Masculinity theories, however, put the underachievement of boys in a broader context by situating their behavior and attitudes in a culture of masculinity. The famous starting point hereof is the British ethnography by Willis (1977) that showed how the rebellious and anti-academic attitude of working class boys, which impeded their success at school, fit within a larger working class culture. This work inspired various investigations into the “lads-culture” by authors like Connell (1989) and Mac an Ghaill (1994). They demonstrated how these boys asserted their masculinity through several practices, one of which was defying the authority and goals of school. Other common practices were “having a laugh”, placing importance on the physical (most prominently on sexuality and sports), acting tough, and displaying a sexist and homophobic attitude (Francis, 2000; Jackson, 2003; Swain, 2005). These practices would impact negatively upon boys’ school-related attitudes and behavior. For instance, “having a laugh” by being the class-clown diverts attention from the content of the lessons (Francis, 2000). It also results in more negative interactions with teachers, who need to intervene to get students back on task (Younger et al., 1999). Furthermore, The emphasis on physicality, such as being a great athlete, is in contrast with necessary behavior for academic excellence, such as sitting down to read and study (Martino, 1999). Moreover, challenging the authority of schools and displaying a tough demeanor impede positive relationships with teachers and school staff.

Since it soon became clear to researchers that not only working class boys enacted “laddish” behavior at school, the original focus on working class culture broadened to a more general masculinity culture which influenced all boys (Van Houtte, 2004b). In these theories, the concept of hegemonic masculinity takes central stage. Hegemonic masculinity refers to the dominant form of masculinity in a certain context, which is superior in the gender order (Connell, 1996). In order for hegemonic masculinity to exist, it must define itself against what it is not, what it is superior to. Therefore, masculinity is constructed within this gender order against subordinated “others” (Brutsaert, 2006; Connell, 1996; Epstein, 1997; Herek, 1987; Schippers, 2007). These others include not only femininity, but also marginalized and subordinated masculinities, such as homosexuality, thus creating a hierarchy of masculinities in the process. Often, these masculinities would be
conflated with femininity, thereby firmly ensuring their subordinate position (Connell, 1996; Schippers, 2007).

These theories have proven valuable in understanding the gender gap, since they show how the culture of masculinity has a profound impact on the way boys enact masculinity and “do gender” in their everyday lives at school. For instance, certain subjects are considered to be “feminine”, such as language, arts or home economics (Connell, 1996; Martino, 1996). Consequently, boys do not want to be caught paying attention or even liking these courses for fear of being associated with femininity. Moreover, studying itself is construed as passive and therefore as a devalued and feminine activity (Epstein, 1998; Martino, 1999). Boys who do work hard at school or cannot compensate good grades with “appropriate masculine” behavior, such as excelling in sports or being the class clown, get ridiculed and called names such as “poofster” or “fag” (demonstrating how homosexuality is construed as both subordinate and akin to femininity) (Epstein, 1998; Martino, 1999; Stoudt, 2006; Swain, 2005; Warrington et al., 2000). The impact of this culture of masculinity on boys’ lives should not be underestimated. Various ethnographic studies show how boys actively try to negotiate maintaining a masculine image with getting good grades. Boys who fail to do so get harassed, while others ultimately place more importance on peer acceptance and popularity and thus end up neglecting their studies. One clear illustration of this mechanism is how the classroom behavior of British high-achieving boys’ changes remarkably when they get older. At first, all through primary school, their classroom behavior resembles most closely that of high-achieving girls: being enthusiastic, providing answers and having positive interactions with the teacher (Jones & Myhill, 2004). By age 14, however, this behavior changes dramatically. High-achieving boys are the least likely to answer questions in class, even less likely than low achieving boys. Researchers attribute this change to the emerging male culture, in which being seen as hard-working or enthusiastic about school is not cool.

**Criticisms of masculinity theory.** Regardless of how enlightening this hegemonic masculinity theory proved to be, it still received criticism. For instance, several researchers stated that the concept of hegemonic masculinity was too static and ignored the real-world variance in masculinities in classes, regions and cultures (Connell, 1996; Connell & Messerschmidt, 2005; Smiler, 2004). They felt that the theory would benefit from acknowledging several possible masculinities, which would still be in a hierarchical gender order in reference to each other as well as to multiple femininities.
These critiques were swiftly incorporated into the masculinity theory and proved beneficial for research into the gender gap in school as well. Researchers were then able to show how several types of masculinity existed next to each other in schools, such as the golden boy, macho boy, geek or nerd (Lyng, 2009) or the real Englishmen, macho lads and new enterprisers (Mac an Ghaill, 1994). Acknowledging these different types of masculinity helped to explain why some groups of boys were more or less able to incorporate good grades into their masculine identity, thus recognizing the fact that some boys were in fact high-achievers and that not all boys were so-called lads who failed at school. Especially illuminating in these masculinities-typologies is the way class, race and gender intersect. For instance, the working-class macho lads from Mac an Ghaill’s ethnography (1994) prided themselves on physical dominance. The real Englishmen on the other hand, who had a predominantly upper middle-class background, emphasized mental dominance and effortless achievement. These intersectional masculinities are an improvement to research because it acknowledges the fact that some boys, especially those from higher socio-economic background, tend to do well in school (Epstein et al., 1998). This, however, does not imply that laddish masculinity conceptions and anti-school cultures are a marginal phenomenon, impacting little on the overall achievement of boys and thus not explaining the educational gender gap. Despite variations in masculinities, Francis (2000) has shown that core masculinity-conceptions have changed little over the past decades and that the laddish construction of masculinity continues to be the most accepted form among secondary school pupils.

Nevertheless, masculinity theories remain insufficient for educational gender gap research on several accounts. For instance, there is a regrettable lack of attention for femininities in both research and theory (Connell & Messerschmidt, 2005; Schippers, 2007), thereby handicapping one half of the educational gender gap research (that is, the research into the achievement of girls). Furthermore, masculinity theory tends to stay firmly linked to biological sex (Francis, 2000, 2010). In the framework of the theory, it is impossible for a boy to be part of feminine culture or for a girl to be part of masculine culture. In such cases, researchers tend to invent a marginalized masculinity or femininity in order to fit these people into a category which remains within their biological sex. What is more, even though the theory acknowledges multiple masculinities, this does not remedy the fact that differences between boys within each category tend to go unrecognized. As Francis argues on several occasions (e.g., Francis 2000, 2010), the use of typologies risks diminishing gender analysis to different sorts of masculinity or femininity, thus ignoring the fact that people tend to do gender in opposite ways.
and that all performances of gender are characterized by contradictions and tensions. For example, people may behave differently depending on the context and the persons present (Francis, 2000, 2010; Swain, 2006). Just imagine a man sitting in the pub with his “mates”, versus that same man alone with his spouse. Categorizing this person as a lad based on his interactions in the pub, would be to ignore the variety of ways he can behave, think and feel on other occasions and the ways in which he differs from other men in the lads category. Moreover, masculinity theory is used as a post-hoc explanation of boys’ achievement, whereby the researcher classifies those boys that underachieve into the lad-group, but is unable to predict beforehand where each boy will go. Last but not least, masculinity theory has stayed firmly rooted in qualitative investigations, thereby limiting the findings to each specific time and setting.

An approach that might overcome these limitations is the gender identity theory. This theory and its history will be discussed in the following section.

**GENDER IDENTITY THEORY**

Gender identity refers to the degree to which a person perceives the self to be masculine or feminine, given what it means to be masculine or feminine in a given society (Perry & Pauletti, 2011; Stets & Burke, 2000a; Tobin et al., 2010; Wood & Eagly, 2009). This concept clearly links to the “doing gender” theory of sociology. Doing gender refers to the ways people infuse their everyday behavior and social interactions with gendered symbolic behavior and signifiers (West & Zimmerman, 1987). According to this theory, gender is a master identity. This means that gender cuts across situations and is omni-relevant, since any action can be interpreted as exemplifying it. As such, people can always be held accountable for the gender appropriateness of their behavior, whether at work, at home or in the street (West & Fenstermaker, 1995). The differences between the fields of sociology and social psychology clearly come to play here. The sociological theory of doing gender focuses on interpersonal interaction and symbolic behavior in the social sphere, whereas the gender identity concept from social psychology starts off at the intrapersonal level as a self-evaluation of masculinity or femininity. However, rather than being at odds with each other, these concepts complement each other and have a reciprocal influence. On the one hand, societal norms concerning ideal masculine and feminine natures may inform people’s gender identity (Tobin et al., 2010; West & Zimmerman, 1987), through a comparison of own characteristics with those from a gender category (for a more thorough discussion of this identity construction, we refer to Tobin et al. [2010]). On the other hand, gender identity
can form the rationale for the specific gendered behavior people display in the social sphere (Tobin et al., 2010; West & Zimmerman, 1987) (for a more thorough discussion of this stereotype emulation, we refer to Tobin et al. [2010]). Via this way, gender identity influences how people perceive the world around them and how they behave.

Historically, most authors have based their gender identity research around the concepts of masculinity and femininity. Since masculinity and femininity have been dubbed “one of the muddiest concepts in social sciences” (Constantinople, 1973), it will come as no surprise that the last half century of research is characterized by a wavelike motion, where periods of consensus on theory and methodology alternate with moments of widespread discussion and disagreement. At the moment, we are at the crest of a wave, where discussion seems to be making way for a new consensus.

### Unidimensionality before the 1970s

The first ones to develop a masculinity-femininity measure were Terman and Miles (Lippa, 2001; Smiler, 2004). Their instrument was called “The Attitude Interest Analysis Test” and they assumed that masculinity and femininity were a unidimensional and bipolar construct. This means that masculinity-femininity form a single continuum with masculinity on the one end of the continuum and femininity on the other. They also presumed that people’s scores should be in accordance with their biological sex and that extreme scores were ideal (Constantinople, 1973; Lippa, 2001; Stets & Burke, 2000a). Therefore, to score “atypically” was considered to be a sign of mental maladjustment. Their work inspired subsequent authors to develop their own instruments, which also assumed a unidimensional bipolar masculinity-femininity construct. Examples of these are the Guilford-Zimmerman Temperament Survey, the California Psychological Inventory, and the Minnesota Multiphasic Personality Inventory (see Beere, 1990).

**Criticism.** While having been popular for quite some time, these instruments got criticized quite heavily. One of the most notable critiques came from Constantinople in 1973. She stated that research supported neither the assumed bipolarity of masculinity and femininity as opposites, nor the supposed unidimensionality of the construct (Constantinople, 1973). Furthermore, the instruments were a mishmash of items that differentiated between the sexes without any theoretical grounding. One could therefore wonder why certain items were chosen in favor of others that distinguished men from women just as well. Finally, several instruments that were supposedly measuring the same construct
barely correlated, demonstrating that the construct was badly defined and/or multidimensional. This was combined with a rising feminist criticism, stating that the construct exaggerated the differences between men and women (Lippa, 2001; Stets & Burke, 2000a). Furthermore, it was posited that the instruments were based on cultural stereotypes and were essentially sexist, since the feminine items often carried negative connotations. People usually feel the need to conform to these cultural stereotypes and expectations, leading to socially desirable answers. While this issue is hardly limited to gender research, and rather is a pervasive concern for any survey study (Billiet & Waege, 2003), it is hard to ascertain to what degree it distorts the answering patterns to gender identity scales. Or as Constantinople (1973, p. 403) puts it: “While it is clear […] that item content, sex role stereotypy and social desirability interact in measures of M-F, making it difficult to obtain a relatively pure measure, it is not yet clear how (and how much) to control for their effects.”

**Androgyny in the Seventies**

In response to the earlier-discussed criticisms on gender identity measures, instruments that adopted a two-dimensional approach to masculinity and femininity were developed. The most famous ones are the Bem Sex Role Inventory (BSRI) and Spence’s Personal Attributes Questionnaire (PAQ) (Lippa, 2001; Smiler, 2004; Stets & Burke, 2000a). While earlier measures usually took (vocational) interests into account, these new instruments focused resolutely on psychological traits. Masculinity was assumed to cover traits such as independence, assertiveness and dominance, while femininity was thought to be represented by sensitivity, kindness and empathy (Bem, 1974; Spence et al., 1975). Respondents indicated to what degree these traits described themselves. Since masculinity and femininity were thought to be independent, orthogonal dimensions, scores on both dimensions were combined. People could consequently score high on masculinity and low on femininity; high on femininity and low on masculinity; low on both; or high on both. With this last category, the concept of androgyny was born. In an attempt to resolve the social desirability issues of gender scales, Bem (1974) also included a social desirability scale consisting of gender neutral items. Her research demonstrated that the correlations with this scale were near zero, indicating that the BSRI did not measure a tendency to answer socially desirable.

While extreme scores in accordance to biological sex were considered to be ideal before the seventies, the present approach took androgyny for the summum (Lippa, 2001; Smiler, 2004; Stets & Burke, 2000a). Especially Bem (1974, p. 162) considered androgyny to be the new “standard of psychological health”. She
believed that the BSRI measured not only masculinity and femininity, but also the
degree to which people were gender schematic. This meant that people with
extreme scores on either masculinity or femininity thought and acted in a sex-typed
way: they would see the world through a gendered lens and act accordingly.
Androgynous people, on the other hand, who possessed both masculine and
feminine traits in equal measure, had a broad range of possible behaviors and could
therefore adapt their behavior to what the situation required. According to Bem
(1974, p. 162), this ability was of the utmost importance in “a society where rigid
sex-role differentiation had already outlived its utility”. Spence (1993; Spence et al.,
1975), however, did not support the gender schema theory, and considered the
PAQ to be measuring one of the psychological representations of masculinity and
femininity, rather than global masculinity and femininity. Nevertheless, with the rise
of feminism and the growing criticism on earlier measures, the concept of
androgyny rapidly became popular (Lippa, 2001) and several researchers jumped on
the bandwagon (e.g., Boldizar, 1991; Chusmir & Koberg, 1988; Orlofsky, 1977;
Schiedel & Marcia, 1985). Consequently, androgyny and the corresponding
instruments were responsible for a rise in gender identity research.

**Criticism.** The BSRI and PAQ remain popular even to this day (e.g.,
Brown, Garavalia, Fritts, & Olson, 2006; Harter, Waters, Whitesell, & Kastelic,
1998; Robison-Awana et al., 2002; Smith, Noll, & Bryant, 1999). Nevertheless,
criticism of these measures and the accompanying ideas was formulated quite
quickly. Research showed, for instance, that some of the basic assumptions on
which the measures were built were faulty. Firstly, masculinity and femininity, as
measured by both PAQ and BSRI, failed to relate to gender-related behavior,
attributes or attitudes (Lippa & Connelly, 1990; Signorella, 1999; Spence, 1993).
This spoke against gender schema theory, since sex-typed persons would be acting
according to either a “masculine” or “feminine” behavioral pattern. Secondly,
masculinity and femininity were supposed to be independent dimensions, while
investigations indicated they actually correlated mildly with each other (Marsh &
Myers, 1986; Spence et al., 1975). Thirdly, while androgyny was in fact positively
associated with wellbeing and self-esteem (Spence et al., 1975), hereby confirming
one of Bem’s central tenets, researchers also showed that this was explained solely
by the presence of high scores on the masculinity scale (Gill, Stockard, Johnson, &
Williams, 1987; Lippa, 2001; Marsh & Myers, 1986). Certain items on the
masculinity and wellbeing scales were similar and would thus explain the
association. Furthermore, researchers began to question whether the masculinity
and femininity scales actually consisted of one dimension. Some research revealed
that several underlying factors existed (Bernard, 1981; Choi & Fuqua, 2003). Most studies found one feminine factor, two “masculine” dimensions (one could be described as an instrumental orientation, the other as self-sufficient), and one bipolar factor consisting of the adjectives “feminine” and “masculine” (Choi & Fuqua, 2003).

The operationalization of the masculinity and femininity concept was criticized as well. The traits which were said to describe a “masculine” personality actually depicted an instrumental-oriented personality, whereas the “feminine” traits depicted an expressive personality (Gill et al., 1987; Lippa, 2001; Marsh & Myers, 1986). As such, these measures failed to grasp global masculinity-femininity and could be considered to measure solely a sub-aspect. After all, one would be hard put to defend the view that masculinity and femininity entail only these characteristics.

These remarks on the operationalization of the masculinity-femininity construct was further expanded with new feminist criticism. Feminists stated that the traits chosen to represent masculinity and femininity were once again replications of cultural stereotypes, which were sexist in nature (Lippa, 2001). After all, many women nowadays would say that they are in fact assertive, independent, logical or individualistic — characteristics all taken by PAQ and BSRI to describe a “male” personality. This view was backed by later American research, which showed that women tended to score higher on masculinity over the years and that so-called masculine traits had become more normative and desirable for women (Auster & Ohm, 2000; Gill et al., 1987; Palan, Areni, & Kiecker, 1999). These findings displayed a societal change in which it had become more acceptable for women to be independent and work. At the same time, it showed that these measures were firmly linked to a cultural frame of reference and thus failed to grasp the abstract reality of masculinity and femininity.

Some researchers concluded that, although the PAQ and BSRI could still be used to measure an instrumental versus expressive personality, they were by no means able to measure the complexity of global masculinity or femininity (Choi & Fuqua, 2003; Marsh & Myers, 1986; Spence, 1993).

**The Trough of the Wave: Crisis in Gender Identity Theory**
When doubt was cast on the validity of the BSRI and PAQ, it was clear that gender identity theory needed a new impulse. It was, however, unclear what this new impulse should be. From the eighties well up into the nineties we can speak of a trough in the wavelike motion of gender identity theory. Discussion and disagreement was widespread. Some authors tried to advance theory (Marsh &
Myers, 1986; Spence, 1993), others developed new instruments. Since a lot of research happened without considering previous results by peers, there was a great deal of confusion, overlap and influence between authors, theories and methodology.

These discussions were not limited to social psychologists, since economists theorized about gender identity as well. Akerlof and Kranton, (Akerlof & Kranton, 2000, 2002) for instance, adapted the economic utility function to include the identity processes. One of the possible identities they theorized about, next to ethnic identity, is gender. However, their conceptualization of gender identity is quite different from the one used here. They do not consider a continuum from masculine to feminine identity, but rather how central an identity as a man or a woman is for decisions and behaviors.

When considering gender identity as a continuum, Bernard uttered as early as 1981 that a multidimensional approach might be favorable and suggested, by analyzing previous measures, as many as six categories: aesthetic interests, manual and physical interests, timidity and sentimentality, temerity, power, and empathy. Lippa (1990) also supported multidimensionality, but used a different approach that was based on factual differences between men and women. Initially, he proposed the empiric-based gender diagnosticity approach in which a probability is computed for every person. This probability reflects the degree to which a respondent is male-like or female-like in a certain field compared to a local peer group. After years of research, Lippa (2001) concluded that it would suffice to consider three dimensions: vocational interests, instrumentality and expressiveness. Wood and Eagly (2009) also proposed a three-dimensional approach to cover the multidimensionality of gender identity, even though their categories were quite different from those of Lippa. They suggested considering traits and interests, whether a person had independent versus collective-oriented relationships, and to what degree a person felt they belonged to the social category of man or woman. Wood and Eagly (2009) also included a short overview of research, including both direct and indirect measures. Indirect measures usually employ implicit attitude tests or priming tasks, which are not subject to conscious control. Interestingly, the majority of the discussed findings seem to accord, regardless of research method. So, even though effect sizes may certainly vary, this overview seems to suggest that the problems with social desirable answers are not as large as to distort general directions and tendencies in the field of gender identity.

Storms (1979), on the other hand, did not develop a multidimensional method. Instead, he developed a new gender identity measure in which people had to answer three face-valid questions about their masculinity and femininity. Storms’
data even suggested a return to the original bipolar continuum with masculinity and femininity as opposites. He stated that despite all scientists’ discussions, people continued to see masculinity-femininity as diametric extremes.

**Questions and criticism.** Notwithstanding all these new ideas and research, several researchers started to doubt the actual existence of such a thing as masculinity and femininity other than as cultural fictions (Lippa, 2001; Perry & Pauletti, 2011; Spence, 1984). Their questions included: Is masculinity and femininity just a personal trait? What is masculinity or femininity for that matter? Is it unchanging through time and space or is it firmly linked to a specific expression of culture? And if the latter is so, what is the use of trying to investigate this ever-changing concept?

Next to this very fundamental criticism, more specific ones were uttered as well. One of these is what we call the identity versus identification discussion. Past measures had employed an identity approach, in which researchers use disguised measures to infer someone’s masculinity or femininity (Stern, Barak, & Gould, 1987; Tobin et al., 2010). This meant that respondents were asked to answer certain questions or indicate to what degree certain items described themselves. The respondents were, however, unaware that what was actually being scrutinized was their masculinity or femininity. The researchers themselves inferred gender identity by comparing the respondents’ score to a self-formulated standard or normative score. By using this approach, several respondents might consider the result they ended up with as not reflecting their felt gender identity. After all, most people consider several aspects of their personality, behavior, interests or appearance as influential for their overall felt masculinity or femininity. The catch is that what is most central for felt gender identity tends to differ from person to person (Perry & Pauletti, 2011; Spence, 1993; Tobin et al., 2010) and would rarely coincide with what the researcher places most importance on (Egan & Perry, 2001). For instance, a woman may score high on the instrumentality of the BSRI, but may still feel very feminine because she is interested in fashion and dance. While another might have no interest at all in a so-called feminine appearance, but may consider herself to be very feminine because she is empathic and relationship-oriented. With this kind of criticism, researchers realized that masculinity and femininity are part of several domains in life, and are not just psychological traits. They therefore doubted that masculinity-femininity could be inferred from any single instrument, whether it measured interests, psychological traits or appearance. They concluded consequently that no existing instrument was able to measure such a broad concept.
as masculinity-femininity (Egan & Perry, 2001; Perry & Pauletti, 2011; Spence, 1984, 1993; Stern et al., 1987).

This train of thought fitted nicely into the commencing understanding of the multifactorial nature of the masculinity-femininity construct. While such thoughts had been uttered early on by, for instance, Constantinople (1973) and Spence (1984), it had never been a popular notion as it was outshone by the rise of androgyny. However, once this concept and its measures did not prove to be up to par, attention started to shift towards the idea of a multifactorial masculinity-femininity construct.

**New Wave or the Rise of Multifactorial Theory**

*The theory.* By the end of the nineties, a new consensus started to form on multifactorial theory. This theory states that not only is masculinity-femininity multidimensional, it is multifactorial as well. This means that masculinity-femininity is evident in several fields of life, such as interests, behavior, appearance, traits, attitudes, and so forth. However, the scores in each of these fields do not necessarily correlate much with a score in another field. Furthermore, the correlations between these factors are not fixed, but may vary from individual to individual (Spence, 1993, 1999; Perry & Pauletti, 2011). Spence stated concerning this relation that “categories of gender-related attributes, beliefs and behaviors typically contribute to separate factors whose relationships to other factors are variable in magnitude, even though often close to zero, and are often complex” (Spence & Hall, 1996, p. 686). It was therefore not surprising that the PAQ and BSRI failed to have clear associations with other gender-related measures such as gender attitudes (Spence, 1993, 1999).

Spence, however, did suggest that there was an overarching masculinity-femininity construct that was the summary of the scores on all of these fields, which could be called gender identity (Spence, 1984). This gender identity-score would be calculated according to the importance each individual placed on every one of these gender-related fields for feeling masculine/feminine. How much each of the fields contributes to this overall felt masculinity-femininity would differ for each person (Spence, 1984, 1993; Tobin et al., 2010). For instance, one person would place more importance on interests, while another would find personality characteristics more decisive.

This theory explains why people who are so very different in characteristics, behavior, interests or appearance might all still feel equally masculine or feminine. This system functions as a sort of defense-mechanism that allows people to feel secure in their gender identity, no matter how different they are to the reigning
gender-stereotypes. That is, people tend to place importance on their gender-congruent characteristics, while discounting a lack of certain gender-congruent traits or possession of gender-incongruent characteristics (Perry & Pauletti, 2011; Spence, 1993; Tobin et al., 2010). What is interesting is that this overarching femininity-masculinity construct would, in fact, be bipolar; thereby once again returning to the earlier models where masculinity and femininity form the endpoints of one continuum. The rationale here is that regardless of the academic discussions about one-dimensionality, two-dimensionality or multidimensionality, in people’s heads masculinity and femininity have been, are, and will continue to be each other’s opposites. This has been backed by several investigations which show that people, in fact, see overall masculinity and femininity as each other’s opposites, that both constructs are defined as what the other is not and that which is one can therefore not be the other (Lippa, 2001; Spence, 1984, 1993; Spence et al., 1975).

This does not have to be a problem or a step backwards in scientific research, since multifactorial theory does break with earlier constructions of masculinity and femininity. Where earlier models assumed that overall masculinity-femininity could be inferred from scores on one sex-differentiating quality, multifactorial theory clearly separates gender identity from any single domain-score. The overarching masculinity-femininity is people’s felt gender identity, a continuum where every person can pinpoint his/her own place on the line as to how masculine or feminine he/she feels him/herself to be. This gender identity is a phenomenological sense of self and, since most persons feel secure in their gender identity, people tend to situate themselves towards the endpoints of the continuum. Clearly separated from this identity concept are the gender-related characteristics and behaviors that people possess. Multifactorial theory states that all of these features contribute to people’s overall felt masculinity and femininity through a subconscious calculus, whose idiosyncratic result is gender identity. The specific constellations of gender-related features tend to vary widely from one person to another. This does not pose a problem to a safe sense of gender identity, since most people assign extra importance to those gender-congruent qualities they do possess, while regarding those they do not possess as inconsequential. Acknowledging that every gender-related characteristic contributes in a different way to people’s felt gender identity, recognizes that what defines somebody as masculine or feminine is different for each person. Hence, it is impossible for a researcher to infer someone’s overall gender identity based on a score in a single field (Spence, 1984, 1993; Tobin et al., 2010). This way, the proposed gender identity theory succeeds in explaining two core characteristics of gender-related behavior that have been puzzling researchers for years: (1) how people within each
gender category are vastly different in behavior, attitudes and interests (Tobin et al., 2010) —so much so that the vast majority of behaviors differs more between people of one gender category than between the sexes (Perry & Pauletti, 2011; Stets & Burke, 2000a). (2) At the same time, despite this huge variance within the sexes, most people feel very confident in their own masculine or feminine identity.

The accompanying methodology. Ever since multifactorial theory has become accepted, research into gender identity is starting to rise. However, research questions cannot be answered as long as there is no valid and reliable instrument that measures gender identity correctly. There have been a few instruments that are in accordance with the basic principles of multifactorial theory. For instance, the Sexual Identity Scale (SIS) is a multidimensional, bipolar and undisguised measurement of people’s self-assessed gender identity (Palan et al., 1999; Stern et al., 1987). Respondents are asked to rate how feminine or masculine they feel on a 5-point scale concerning their looks, interests, behavior and feelings. This way, the SIS taps the physical, cognitive, societal, and emotional dimensions of gender identity. While the SIS is elegantly simple and provides face-validity, it appears that scores tend to be quite extreme. That is, the instrument correlates very highly with biological sex (Stern et al., 1987) and sex accounts for 87% of the variance in responses to the SIS (Palan et al., 1999). This means that the SIS cannot provide a surplus value to gender identity research, simply because most people look, act and feel like their own sex. It should be mentioned, however, that these investigations have been with an American-based representative sample. We speculate that the instrument might provide more variance with certain populations, such as with people who identify as transgender.

For research with representative samples, and more specifically with children, the Self-Perception Profile of Egan and Perry has proved to be valuable. According to Egan and Perry (2001), gender identity is a construct that should be measured multidimensionally. They differentiate several aspects of gender identity, such as membership knowledge, gender compatibility, pressure for gender-conforming behavior and attitudes toward gender groups (Egan & Perry, 2001; Tobin et al., 2010). Membership knowledge, more commonly referred to as gender constancy, indicates to what degree children realize they are of a certain sex and that this will not change of its own accord. Since most children reach gender constancy around age six or seven, the researchers agreed that it was not opportune to include this in the measure. Gender compatibility, which refers to a sense of belonging and contentment with one’s gender category, is divided into two dimensions: gender typicality and gender contentedness. Gender contentedness
originally referred to how happy one is with one’s sex. Nowadays, researchers tend to agree that the scale actually measures how happy one is with one’s gender-role (Leaper, Farkas, & Brown, 2012). Secondly, gender typicality refers to the degree to which one feels similar to others in the gender category (Egan & Perry, 2001; Tobin et al., 2010). In accordance with multifactorial theory, it is supposed that different children will feel gender typical for different reasons. Pressure for gender-conformity refers to the degree to which one experiences pressure from parents, peers and self to not exhibit cross-gender behavior. Finally, “attitudes toward gender groups” refers to the degree of in-group bias. This instrument has proved to be both reliable and valid with American adolescents (Egan & Perry, 2001; Tobin et al., 2010). Consequently, it has become more widely used throughout the last decade.

While most research in gender identity the past decades has been about validating or refuting the reigning theory of the time, there is at last a dawning consensus on both theory and instrument. This means that the time has come for gender identity to step back into the world of applied research.

(Re-)Introducing Gender Identity Theory to Gender Gap Research

In the following, we will give a brief overview of research that connects gender identity theory with research into the educational gender gap. One should keep in mind that the first part focuses on the use of PAQ or BSRI, which should not be considered a correct measure of gender identity. We include this, however, for a sense of completeness. Moreover, androgyny-research has been such an important part of gender identity research that it cannot be ignored. Furthermore, this research can provide valuable inspiration for research questions or explanatory mechanisms. In the second part, we give an overview of the recent research that has employed the multifactorial theory and its instruments. While these investigations are still scarce, they seem a promising new avenue for research into the educational gender gap.

The First Investigations using Androgyny

When everybody jumped on the androgyny-bandwagon, the measures were applied in several research areas. One of these areas was the explanation of differential academic achievement by boys and girls. Boldizar (1991) found for instance that masculinity, as measured by the Children’s Sex Role Inventory (CSRI), connected
to self-perceived scholastic competence in American nine to thirteen-year olds of both sexes and boys’ mathematic achievement scores. Femininity, on the other hand, negatively related to girls’ mathematic achievement scores. Robinson-Awana (2002) found, in a similar vein, that American ten to twelve year-old girls with high academic competence had high masculinity scores, as high as those of the average boy. By considering attribution style, self-esteem and masculinity-femininity scores, the researchers were able to correctly classify 83.3% of the above-average achieving girls. This led them to conclude that certain “masculine” characteristics were beneficial for girls’ achievement. They were, however, unable to satisfactorily classify the boys according to these variables. This led Robinson-Awana to conclude that what determined girls’ academic achievement was not necessarily similar for boys.

While the above-mentioned investigations found that masculinity-femininity related to academic achievement, other authors, however, did not. For instance, Brown et al. (2006) did not find a link between the PAQ and academic achievement among American computer science students. Similarly, Zand and Thomson (2005) found no direct link between gender identity and academic achievement among Afro-American eleven to 14-year olds. Nevertheless, they did find an indirect connection. Gender identity explained self-worth, which was associated with school bonding, which in turn related to school grades. Such an indirect connection was found as well by Eisele, Zand and Thomson in a similar sample (2009). Gender identity explained differences in acceptance by peers, which related to behavioral conduct, which in turn was associated with school bonding, which once again related to academic achievement. In both investigations, it was the independence and leadership aspects, the so-called masculine orientation of gender identity, which explained both self-worth and acceptance by peers. And in both investigations, once controlled for gender identity, sex itself was no longer relevant.

We can conclude this section by stating that research into academic achievement using the BSRI and PAQ has provided some inconclusive results. While some researchers found direct connections (Boldizar, 1991; Robison-Awana et al., 2002), others only found indirect links (Eisele et al., 2009; Zand & Thomson, 2005) and still others found no connections at all (Brown et al., 2006). Furthermore, when considering those who do find a connection between gender orientation and academic achievement, it is quite notable that only the masculinity scale provides positive associations. This runs counter to the current findings in gender gap research, where girls are outperforming boys. However, we should keep in mind that the measures used here should not be considered correct interpretations of an overarching gender identity construct. As we have stated
before, the PAQ and BSRI only measure instrumental versus expressive traits. It is therefore wrong for researchers to portray their results as the influence of “masculinity” or “femininity” on academic achievement. Only a few researchers, such as Zand and Thomson (2005), explicitly state they measure leadership and independence qualities, not gender identity or masculinity. Once we keep this in mind, certain puzzling or astounding results are quite easily explained. For instance, it has been demonstrated that instrumental-oriented people have higher self-esteem (Marsh & Myers, 1986; Orlofsky, 1977; Spence et al., 1975). It follows quite easily that these confident people who feel they are independent, logical and intelligent, have higher self-perceived scholastic competence and tend to score high in mathematics (since this is a school subject in which logical thinking is key).

**Riding the Wave: Using the Self-Perception Profile**

Even though the Self-Perception Profile has started to become more accepted and widespread in the academic world, its use has been limited to mostly psychological adjustment research with adolescents and children. Consequently, research that utilizes the Self-Perception Profile for the explanation of educational achievement is still in its infancy. Therefore, the research overview we can present here is limited.

Research by Leaper, Farkas and Brown (2012) showed that felt pressure for gender-role conforming behavior was related to American adolescent girls’ motivation in English, mathematics and science. Felt pressure for gender-conforming behavior from parents, for instance, was positively related to motivation for English and negatively related to science and mathematics motivation. Gender typicality and contentedness, on the other hand, were not related to girls’ motivation. The researchers explained this unforeseen finding by suggesting that being interested in mathematics or science is in our present day and age no longer incompatible with feeling like a typical girl. Nonetheless, an investigation by Leaper and Van (2008) showed that gender typicality was related with American college men’s self-efficacy and interest in traditional versus non-traditional majors. Men who were high in masculinity ideology and gender typicality were most likely to hold traditional interests (such as engineering, physics, economy and mathematics), while men who were low in gender typicality and covert sexism had the highest self-efficacy in non-traditional fields (such as social sciences and linguistics). Consequently, men’s scores on covert sexism and gender typicality fully explained the effective choice for traditional versus non-traditional majors. Thus, this research showed that men who felt to be gender typical or endorsed traditional gender attitudes were less interested in feminine majors, and felt traditional majors
to be more compatible with their self-concept. As a consequence, the likelihood that these men took up these non-traditional majors was lower in favor of pursuing traditional majors.

Patterson and Pahlke (2011) investigated American ten to thirteen year-old girls’ achievement predictors in a single sex school and took gender typicality into account. It should be noted here that the measure they used was not the measure developed by Egan and Perry. However, for a sense of completeness of the research overview, we include this investigation. They found no connection between gender typicality of these girls and their academic achievement one year later. The authors did remark that this might have been due to the fact that gender typicality was measured before the girls entered the single-sex school. The authors hint that a student’s felt gender typicality may have changed as a consequence of the single sex environment and that their —lack of— results may be due to this.

After this short review, we can conclude that the application of the new gender identity measures on gender gap research is still minimal. However, the cited investigations here do offer some first valuable insights and, even more so, raise questions by pointing out gaps in knowledge. For instance, it is an interesting finding that while gender typicality is related to college men’s interest, self-efficacy and choice for academic fields, it is not related to girls’ motivation in English, mathematics or science. We could wonder whether this is an age-effect since Leaper and Van’s study (2008) concerned college-aged persons, while Leaper, Farkas and Brown’s research (2012) concerned 13 to 18-year olds. However, it might be more plausible that this is a gender effect, since the first study concerns men and the second concerns girls. It might be possible that different variables explain girls’ versus boys’ motivation for gender-appropriate versus gender-atypical study subjects. That girls’ motivation was not related to gender typicality while men’s was, might be a sign that gender-atypical behavior for women has become much more accepted than men’s in the current western society. This is a suggestion that has been put forward by several other investigations in different fields (Egan & Perry, 2001; Kimmel, 2007; Palan et al., 1999; Stern et al., 1987). It is stated that in our western society, with the influence of feminism and emancipation, women have gained the freedom to behave more like men. However, men have not gained the freedom to behave more like women without suffering the ridicule of peers and society. We suggest such statements to be the focus of new research into the relation between gender typicality and academic motivation.

Furthermore, it is notable that the investigations cited here have focused their research on either males or females. So far, we are unaware of any investigation that includes both sexes simultaneously in its sample. However, this is
necessary in order to make legitimate claims as to differing effects of gender typicality on men or women. Nevertheless, we suggest at the same time to break through the sex dichotomy that has been pervading gender identity research since its beginnings. Too much research has focused on differences between men and women. With the multidimensional measure of Egan and Perry, it would be most interesting to contrast the experiences of gender typical with atypical persons, gender-role content with gender-role discontent people and individuals who experience pressure for gender conformity with those who do not. We propose that, rather than blindly focusing on sex differences, there might be mechanisms that unify people across the sexes according to their sense of typicality or felt pressure. It is therefore important for researchers to pay attention to effects of gender identity, while keeping the possibility of interactions between sex and gender identity in mind.

How Gender Identity Theory can contribute to Gender Gap Research
To summarize, we believe that gender identity theory could provide a valuable contribution to the already-existing gender gap research in two ways. Firstly, gender identity theory can be a direct factor in the explanation of the gender gap in education. Gender identity could, for instance, explain achievement, motivation and interest differences between boys and girls by considering how gender typicality connects to these outcomes. As became clear in the previously discussed investigations, it could be that high gender-typicality prevents boys, but not girls, from being successful in gender-atypical study fields, thus creating a gender gap in, for instance, language achievement. We could further speculate that high pressure for gender-conformity associates with lower motivation and interest in gender-atypical courses for both boys and girls, while at the same time positively relating to achievement in gender-appropriate studies. We could also wonder whether overall felt “masculinity” or “femininity” is connected to a certain study method or classroom behavior, which could be either characterized by high self-discipline, conscientious and calm demeanor versus a more rebellious, active, or fun-oriented pragmatism.

Secondly, gender identity theory can contribute to gender gap research in an indirect fashion by ameliorating the explanatory power of masculinity theories. We have mentioned before that masculinity theories offer post-hoc, group-based and time and place-specific explanations which remain firmly linked to sex. Gender identity theory’s strength, on the other hand, is that it deals with an individual-based characteristic, which can be considered independently from specific peer
cultures or biological sex. We suggest that, by connecting the insights of gender identity theory to masculinity theory research, more profound explanations of the impact of masculinity culture on the gender gap in education can be achieved. For instance, research that uses masculinity theories to explain academic achievement tends to divide the school population in several study or peer cultures. However, the explanation of why any person is a member of this or that study culture tends to be characterized by a circular argument. A certain boy is part of a certain group because of his gender ideology, study attitudes, and so forth. But at the same time, it is assumed that boys acquire these characteristics because they are part of these groups. We propose that gender identity can provide a possible explanation for group-membership. For instance, it is plausible that children who feel extremely gender typical or experience strong pressure for gender-conforming behavior have a higher likelihood of becoming members of “macho-groups”. Whereas children who feel atypical and do not experience much pressure might feel free to become part of less rebellious, less “cool” or even pro-school groups. In this way, by connecting the findings of gender identity research with those of masculinity theory, we can better understand the influences on both the individual and the social level. As a result, we might develop theoretical models that are better able to take into account the complexity and nuances of social realities. Hence, this might enable us to go from explanations that necessarily remain post-hoc and categorical to explanations that can be nuanced and predictive.

CONCLUSION

In this article, we discussed the gender gap in education. Even though this phenomenon occurs in many western countries, we have focused on Anglo-Saxon literature. The reader should thus be careful in extrapolating the presented findings, effects and suggestions beyond this geographical scope. Several scientific explanations for why either girls or boys underachieve have been examined. There is, for instance, attention for traits and behavior, such as self-discipline or lack thereof, school attitudes, and classroom behavior. Other theories go beyond these rather descriptive explanations and aim to explain why such behavior occurs. One of the most notable theories herein is the masculinity theory, which discusses how the culture surrounding masculinity and macho behavior influences boys’ school behavior and achievement.

While this theory has been very illuminating, we formulated several criticisms. Firstly, masculinity theory tends to conflate gender with sex, since people cannot be part of a gender culture that contrasts with their biological sex (Francis,
Secondly, femininity remains under-theorized and under-investigated when compared to its counterpart masculinity (Connell & Messerschmidt, 2005; Schippers, 2007). Thirdly, nuance and contradictions go unrecognized because of masculinity theory’s group-based approach. Several people get lumped into a single masculinity-category, which leads to overlooking people’s individual behavior and the contradictions and differences between people in this category (Francis, 2000, 2010; Swain, 2006). Finally, masculinity theory is almost exclusively used in qualitative investigations, thereby limiting its findings to post-hoc explanations and specific places and time.

We would like to suggest that gender identity theory could provide an answer to most of these criticisms, and hence help to further gender gap research. It is remarkable, however, that gender identity theory has scarcely been used in applied educational research up to this day. We have discussed the history of gender identity theory and have shown how both theoretical and methodological ambiguity has impeded its applied research. Nevertheless, in the recent decade important advances have been made in gender identity theory, providing us with a theoretical framework and an accompanying instrument that has received more and more support in scientific circles.

Hence, we would like to propose that the time has come to re-introduce gender identity theory to gender gap research. While the first investigations combining both fields have been rather scarce, they do provide some promising new pathways and food for thought. For instance, it would be interesting to investigate how gender identity differently relates to boys’ and girls’ motivation and interest for gender appropriate versus atypical school subjects. Next to these possible sex differences, there might be mechanisms that unify people across the sexes according to their sense of typicality or felt pressure. For instance, gender typical boys and girls may feel and behave similarly at school, while their experiences may contrast with those of gender atypical boys and girls. In this way, gender identity might be linked to certain classroom behaviors or study methods.

All in all, there is still much to discover regarding the impact of gender identity on school achievement. We posit therefore that this is a promising new area of interest for both gender researchers and educational researchers alike.
Chapter 5.
An integrated model of gender identity and school functioning

In the previous chapters, we discussed how feminist thinkers drew attention to the subjugation of women, as represented in the concept of patriarchy. Simultaneously, the concept of gender was developed, resulting in growing interest from social science researchers. We discussed how this new focus unveiled underlying processes of power, both on the institutional level in the construction of heteronormativity, as well as on the interactional level in “doing gender”. At the same time, the exclusive attention on the plight of women expanded to crosscutting mechanisms of oppression through race, class and sexual orientation. Scholars also drew attention to masculinities by demonstrating variety in masculine enactment and the subjugation of non-privileged men.

When it was revealed in the 1990s that boys were underachieving compared to girls, these theoretical developments were utilized in educational research. Initially, this research tended to be dominated by sociology-based theories, such as masculinities theory. Later, insights from social-psychology, such as the identity-based motivation model, were added. Through this combination, the opportunity arose to include more individual-based quantitative studies, and hence a more nuanced exploration of intrasexual variations. However, problems with the theoretical and methodological development of the gender identity concept impeded this integration (see chapter 4). Nonetheless, with the growing consensus on multifactorial theory and the validation of the self-concept questionnaire developed by Egan and Perry (2001), new avenues for research were opened up. These avenues will be explored in this dissertation, by considering the associations between the school functioning of boys and girls and their self-perceived gender typicality and pressure for gender conformity. In this conceptualization, gender typicality functions as the contemporary approximation of self-assessed masculinity-femininity. Since this concept is associated with identity and located within the self, gender typicality can be situated within a social-psychological tradition. Conversely, the concept of pressure for gender conformity is associated with societal expectations and pressures. As such, it taps social aspects of gender enactments, and is therefore part of sociological tradition.
In previous chapters, we have focused on the gender aspect of this story. In what follows, we discuss our operationalization of school functioning, before we derive the concrete research questions that will be assessed in the empirical chapters.

**SCHOOL FUNCTIONING**

In this research, we consider several indicators of school functioning to strengthen the robustness of our findings. We rely on the expectancy-value model developed by Eccles and colleagues (Wigfield & Eccles, 2000) for the delineation of these indicators. Note that using this model offers several benefits. First, this model can be used not only to explain performance, but is also applicable to choice, persistence, and effort in activities. This has often been applied to areas of gender difference, such as the interest in STeM or language (Fredricks & Eccles, 2002; Lupart, Cannon, & Telfer, 2004; Robnett & Leaper, 2013), or career-related decisions (Eccles, 1987; Leaper, 2015). Second, the expectancy-value model has contributed to opening up the black box for why we find differences in achievement, rather than simply observing these differences. Consequently, this model provides more insight than if one were to solely rely upon indicators of achievement.

The expectancy-value model suggests that three factors are key in affecting students’ learning experiences: expectations, value and affect (Pintrich & De Groot, 1990; Wigfield & Eccles, 2000). First, the expectancy component refers to students’ beliefs about their abilities, and hence, taps into a sense of competence. This factor has been operationalized in many ways, such as control beliefs or attributional styles, but also as academic self-efficacy (Pintrich & De Groot, 1990). Second, the value component refers to students’ beliefs about the importance and interest of studying. This has once again been operationalized in a variety of ways, such as attainment value or usefulness, but has often included a distinction between autonomous and controlled motivation (Wigfield & Eccles, 2000). Third, the affective component of the model includes students’ emotional states. While this has frequently been assessed as task-specific emotions, such as test anxiety (Pintrich & De Groot, 1990), we will consider a more global assessment of affect. Because students spend most of their waking hours at school, the school environment constitutes an important social and developmental context (Eccles, 2007; Roeser, Eccles, & Sameroﬀ, 2000), impacting on school functioning above and beyond simple task-related emotions. Consequently, we will consider students’ global sense of well-being as an affective indicator.
In short, we will consider three indicators of school functioning derived from the expectancy-value model by Eccles and colleagues (Wigfield & Eccles, 2000). This includes subjective well-being as an indicator of global affective functioning; and two indicators of academic school functioning with study motivation (as the value component) and academic self-efficacy (as the expectancy component).

First, study motivation will be conceptualized in line with insights from self-determination theory (Deci & Ryan, 1985), which revealed that the quality of motivation is more influential than the quantity. Consequently, we distinguish two subtypes of study motivation: autonomous and controlled study motivation (Deci, Vallerand, Pelletier, & Ryan, 1991). Autonomous motivation is characterized by enjoyment and personal relevance, while controlled motivation is prompted by external pressures—such as force—, or internal ones—such as shame or guilt. According to this conceptualization, autonomous motivation is of better quality than controlled motivation. Consequently, autonomous motivation is linked with superior study behaviors and higher academic achievement, while controlled motivation is associated with maladaptive learning strategies and reduced educational performance (Ryan & Connell, 1989; Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009).

Second, academic self-efficacy refers to people’s beliefs in their own capability to perform academic tasks (Bandura, 1997). Such a sense of confidence ensures that people exploit their capabilities to the fullest, and is especially powerful in the face of adversity—for instance, in order to persevere when faced with difficulties or failure. Meta-analyses have pointed out that both study motivation and academic self-efficacy are among the strongest self-measures contributing to achievement (self-efficacy $d = .43$; study motivation $d = .48$) (Hattie, 2008). Additionally, these indicators function largely independently and additively on achievement (Pintrich & De Groot, 1990), making it relevant—and not redundant—to study them alongside each other. Hence, we will consider the influence of gender on academic self-efficacy in chapters 7 and 9, and the association with study motivation in chapter 8.

Third, well-being has been shown to influence academic school functioning in a variety of ways. For instance, a meta-analysis has shown that happy people tend to be more creative and efficient problem solvers, culminating in notably large effects ($d = .54$) (Lyubomirsky, King, & Diener, 2005). With regards to the indicators of academic school functioning considered in this research, research has shown that well-being impacts upon self-efficacy through recollection-bias...
(Bandura, 1997). That is, people who feel down tend to remember previous failures rather than accomplishments, which in turn negatively impacts their sense of efficacy. Similarly, well-being also affects study motivation with, for instance, negative emotions such as anger and anxiety leading to an increase in controlled motivation (Assor, Kaplan, Kanat-Maymon, & Roth, 2005).

Next to the influence of well-being on academic school functioning, well-being itself is impacted by gender-divergence. As discussed above, gender-incongruent behavior is often associated with social sanctions, such as disapproval, ostracism or bullying. Besides these processes of social disparagement, individual self-evaluations are important as well. That is, people do not live in a social vacuum and are well aware of gender beliefs in society (Ridgeway & Correll, 2004). When feeling atypical or performing gender-incongruent behavior, people realize they are not living up to societal expectations that are often deemed normal and desirable. This notion is confirmed in empirical research by Menon (2011), who showed that gender non-conforming adolescents reported higher levels of psychological distress partly because they felt a sense of incompatibility with their gender category. Unfortunately, gender-divergent individuals are trapped in a catch-22: even though gender divergence is associated both with social sanctions and negative self-evaluations, disavowing gender-nonconformity would not necessarily lead to more positive outcomes. Being able to behave in ways that feel authentic and true to the self are paramount to a person’s welfare (Deci & Ryan, 1985; Uysal, Lin, & Knee, 2010). Indeed, Good and Sanchez (2010) showed that voluntarily behaving in gender-appropriate ways was associated with higher self-esteem. However, when the same behavior was enacted because of a sense of pressure or obligation, this was associated with lower self-esteem.

In short, well-being is a factor of central importance in this research, being related to gender performances, as well as academic school functioning. Consequently, in all upcoming empirical chapters, well-being will be considered as either a mediator (see chapter 7, 8 and 9) or an outcome variable (chapter 10).

All in all, this research can be represented in a conceptual model (see Figure 1). In what follows, we will infer specific research questions from this model, that will be discussed in detail in the empirical chapters.
**RESEARCH QUESTIONS**

To consider intrasexual differences from a gender perspective and to bring a gender focus back into educational gender gap research, we introduce two central explanatory variables: gender typicality and pressure for gender conformity. We link these variables to several indicators of school functioning in the empirical chapters.

First, we will focus on the associations between gender typicality and academic school functioning. Social identity theory and the identity-based motivation model suggest that social identities are vital motivators for the specific courses of action people follow (Elmore & Oyserman, 2012; Oyserman, Bybee, & Terry, 2006). When certain behaviors are congruent with contextually salient identities, they are more likely to come to mind and to be performed. Furthermore, research has shown that performing identity-congruent behavior results in a higher sense of autonomy and well-being (Uysal et al., 2010). Previous research has only
rarely considered the influence of gender identity on school parameters, and especially intrasexual variation has been overlooked (Elmore & Oyserman, 2012). Consequently, we apply these notions to the connection between gender identity, operationalized here as gender typicality, and school functioning. Because femininity is thought to be more compatible with the school setting and study-related behaviors than masculinity (Francis, 2000), we expect girls to display better academic school functioning than boys. When considering intrasexual variation, we could expect girls who perceive themselves to be typical for their gender to show more advantageous academic functioning than girls who feel they deviate from feminine norms. For boys, we might expect opposite connections. Boys who perceive themselves to be highly typical for their gender category might be less concerned with applying themselves to school, which would then hamper their academic school functioning. Conversely, such concerns might be less prevalent among self-perceived atypical boys, leading them to score higher on academic functioning than gender-typical boys.

These hypotheses will be assessed independently for academic self-efficacy in chapter 7, and for study motivation in chapter 8. The concrete models assessed in these chapters are graphically presented in Figure 2 and Figure 3.

**Figure 2.**
Gender typicality and academic self-efficacy (see Chapter 7)
Second, we consider how pressure for gender conformity relates to adolescents’ school functioning. The concept of gender-conformity pressure seems especially relevant in adolescence, when the gender intensification process is in full swing (Galambos, Almeida, & Petersen, 1990). Indeed, gender plays a central role in the developmental tasks of early adolescence, when youth strive for the emulation and consolidation of adult male and female roles—not only in terms of physical and sexual maturation, but also the planning of an academic and occupational future (Perry & Pauletta, 2011; Tobin et al., 2010). Combined with the fact that the approval and acceptance of peers takes center stage (Eccles, 2007), it seems pertinent to consider concepts that measure an internalized sense of pressure for gender conformity among early adolescents. Previous research concerning gender-conformity pressures has mostly been ethnographic, where researchers observe interactional patterns between youth (e.g., Francis et al., 2010; Jackson, 2003; Pascoe, 2007; Skelton & Francis, 2011). While very illuminating, this research has focused predominantly on boys, and individual variations in students’ susceptibility to such pressures has been hard to assess. A benefit of assessing gender-conformity pressure on the individual level is exactly that we can consider how its influence varies from one individual to the next. While it is certainly true that some environments subject their members to more conformity-pressures than others...
(Swain, 2005), it would be short-sighted to think this influences everybody in the same way. That is, some people might be more susceptible to the internalization of such pressures, while others might have peer groups or home environments that can buffer the conformity-pressures from the environment. By using the concept of pressure for gender conformity, we are able to provide a more nuanced analysis of these individual variations.

Hence, in chapter 9, we consider the association between gender-conformity pressure and academic self-efficacy. Consistent with conceptualizations of femininity and masculinity in school settings (Francis, 2000), and the accompanying pressures in peer cultures to conform to these notions (Pascoe, 2007), we expect no positive effects of the experience of gender-conformity pressure on boys’ academic self-efficacy. Conversely, girls, who experience strong pressure to conform to feminine norms, might display higher levels of academic self-efficacy than girls who do not experience such pressures.

Risman (2004) underscored the ways in which gender is simultaneously produced and maintained on the institutional, interactional and individual levels. Consequently, in chapter 10, we zoom out from the individual and interactional level that were the focus of previous chapters. In this chapter, we recognize that
gender performances do not happen in a social vacuum, but are always contextually situated. To do this, we consider the impact of schools’ heteronormativity levels on the gender experiences of students. While previous ethnographic research has considered the influence of gender regimes on students’ school lives, comparing several schools simultaneously is a difficult scope to attain in most qualitative research (e.g., Morris, 2012; Pascoe, 2007; Swain, 2006). Conversely, quantitative research considering schools’ heteronormativity has rarely considered students’ gender-divergence besides sexual orientation (e.g., Chesir-Teran & Hughes, 2009; Craig & Smith, 2014; Goodenow, Szalacha, & Westheimer, 2006; Szalacha, 2003). We therefore consider the way different school contexts influence the association between students’ well-being and their experienced gender-conformity pressure and gender typicality. Since research suggests that deviating from gender norms is associated with both social disapproval and self-derogation (Menon, 2011), and that heteronormative environments place stronger restrictions on gender expression (Toomey et al., 2012; Wilkinson & Pearson, 2009), we expect that gender-divergent students will experience a stronger impact of heteronormative school contexts on their well-being.

Figure 5.
Student’s well-being according to school’s heteronormativity (see Chapter 10)
Chapter 6.
Methodology

The “Teaching in the bed of Procrustes”-project, financed by the Agency for Innovation by Science and Technology (IWT), investigated the educational gender gap in Flanders from several angles, including student-teacher interactions, peer cultures and identity development. The present research, considering central aspects of gender, constituted one of these research angels and was thus embedded in the Procrustes-project.

Early adolescence is characterized by a process of gender intensification and differentiation (Galambos et al., 1990). This impacts upon the school trajectories of boys and girls, with gender-differential achievement patterns becoming more pronounced during secondary education (Derks & Vermeersch, 2002; Fergusson & Horwood, 1997; Voyer & Voyer, 2014) (see chapter 2). To investigate the origins and consequences of this intensification process, the Procrustes project focused on early adolescents. Furthermore, to better assess these developmental changes, the project employed a longitudinal design with three data collections in the course of two school years. The study design and the data, which have been collected in the course of the project and which are used in this research, will be discussed more thoroughly below.

DATA

Data Collection Strategy
Based on information from the Department of Education, all schools that offer 7th grade education in Flanders were inventoried. A disproportionally stratified sampling method was used, meaning that certain criteria were used to delineate subpopulations, after which random samples were drawn. To adequately represent the Flemish educational context, we advanced three criteria for the construal of these subpopulations: region, school denomination and location.

First, we aimed for a fair representation of all five provinces and the municipal territory of Brussels. Second, based on the census of 1991, we classified schools as either urban or rural-based. Third, a realistic representation of school denomination was based on records from the Department of Education, showing 69% catholic versus 31% public schools in Flanders. The aim was to have 63 schools participate in the study. Consequently, for each region, we strived for the
cooperation of 12 schools, with an even representation of rural versus urban schools, and a 2/3 representation of catholic versus 1/3 public schools. The Municipal territory of Brussels constitutes a special case because of its size and level of urbanization. Consequently, we aimed at the cooperation of three schools from Brussels, with two catholic versus one public school, which would all be classified as urban.

Within these parameters, three random samples were drawn. When a school from the first sample refused after both written and telephonic contact, a matched school from the second random sample was contacted, and so on. By the start of the data collection period in October 2012, 59 schools had agreed to participate in the study. We contacted 124 schools to achieve the participation of 59 schools, translating to a response rate of 47.6%. So, even though the original aim of 63 participating schools was not realized, a response rate of 47.6% is quite good, since schools in Flanders are swamped with research requests. School principals indicated that their refusal to participate in the study was mostly fuelled by either their involvement in a different research project, or the existent heavy workload of the staff. Only one school refused participation because the research topic did not interest them. This implies that refusal to participate was largely at random. Analyses in which we compared our sample to the Flemish school population, based on the statistical yearbook 2012-13 from the Educational Department, seem to confirm this. As can be seen in Table 2, there are some small divergences between the sample and the population. Especially the small overrepresentation of boys stands out. This is probably caused by the stepwise nature of the sample, since schools were selected rather than students an sich. Some schools in Flanders offer, for instance, predominantly technical and vocational tracks, which tend to appeal more to boys. Consequently, some schools in the sample had an almost exclusively male student population, which could explain the small overrepresentation of boys in the overall sample.

In short, while there are some small divergences in the sample from the population, it seems unlikely that these pose a problematic distortion of the data. This suggests that no systematic biases occurred and that the schools in the first wave were largely representative of the population.
Table 2.
Comparison of sample and population characteristics

<table>
<thead>
<tr>
<th>Student variables</th>
<th>Sample</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Girls</td>
<td>46.2%</td>
<td>49.2%</td>
</tr>
<tr>
<td>Boys</td>
<td>53.8%</td>
<td>50.8%</td>
</tr>
<tr>
<td>On track-students</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>76.6%</td>
<td>77.6%</td>
</tr>
<tr>
<td>Students who do not speak Dutch</td>
<td>11.2%</td>
<td>10.9%</td>
</tr>
<tr>
<td>Non-Belgians</td>
<td>8.2%</td>
<td>7.2%</td>
</tr>
</tbody>
</table>

As mentioned above, the Procrustes-project did not only collect cross-sectional data, but employed a longitudinal design to allow the study of changes in students’ attitudes and achievements. Consequently, the students who were part of the first wave, collected in the fall of 2012, were surveyed two more times. Once at the end of 7th grade, in the spring of 2013, and once more at the end of 8th grade, in the spring of 2014. The central measures in the research were part of the surveys at each time-point. However, to conserve time and to reduce the number of pages in the survey, demographic data was only asked at the start and the end of data collection (that is: the first and third wave).

While 59 schools participated in the first wave, one school dropped out of the study in the second wave. This was due to a change in the position of school principal, and a resulting issue with upholding appointments. By the third wave, another school had ceased to exist. In the end, 57 schools participated in all three waves. From a cross-sectional point of view, 6380 students were questioned in the first wave, 6234 students in the second wave, and 6163 students in the third wave. When these waves were matched, 4987 students participated in all three waves. This reduction in the number of cases is due to the disappearance of two schools from the study, students not being present during one of the survey-moments (for instance, due to sickness), or because students transferred to a different school in the course of the study.

**Survey Collection And Data Processing**

Longitudinal design implies that students’ answers from the first wave are matched to their responses on the second and third wave. Because of this, completely anonymous surveys were not an option. Rather, coded questionnaires were used, meaning that every individual receives a unique code that remains the same throughout the data collection. A separate, non-accessible file exists where the codes are linked to the names of the students. However, to protect the privacy of the respondents, the dataset used for analyses consists solely of anonymous data.
In each school, all 7th graders were asked to complete the questionnaire. A waiver of parental consent and the use of child assent were approved by the school and the Belgian Commission for the Protection of Privacy, based on the minimal risk and confidential nature of the study. Students had 50 minutes to complete the paper and pencil survey. Researchers were always present during the completion of the survey to explain the procedure and answer questions. Pupils were told that the survey was not a test and were assured that it was completely confidential.

Response-rates were high during the data collection, since only students who were not present (for example, due to sickness) did not complete the survey. The response rate for the first wave, for instance, is 96.9%. Nevertheless, some of these questionnaires were not usable. That is, while most children were more than willing to answer the survey questions, some children were not. In these cases, researchers talked to these children one-on-one, to explain to them the goal and importance of the study, why their co-operation was essential, to stress the confidentiality of the study and/or to motivate them to read and answer all questions. This approach was effective in that most students then proceeded to fill out the questionnaire and only a small minority continued to refuse. Of those who refused, most had nevertheless still filled out the background information (such as, sex, age, school class, and so on) and had only given up when confronted with the likert-scales. Hence, these students are not added to the non-response statistics, but rather tend to have a lot of missings throughout the questionnaire. Since less than .5% of all longitudinal questionnaires constituted cases where students had completed less than half of the survey scales, this amount seems unproblematic.

Table 3.
Sample characteristics

<table>
<thead>
<tr>
<th></th>
<th>Range</th>
<th>Wave 1</th>
<th></th>
<th></th>
<th>Wave 1</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
<td>%</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Sex</td>
<td>0 = Boy</td>
<td>53.8</td>
<td>3435</td>
<td></td>
<td>52.5</td>
<td>2620</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 = Girls</td>
<td>46.2</td>
<td>2945</td>
<td></td>
<td>47.5</td>
<td>2367</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Track</td>
<td>0 = Academic</td>
<td>88.3</td>
<td>5633</td>
<td></td>
<td>90.5</td>
<td>4513</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 = Vocational</td>
<td>11.7</td>
<td>747</td>
<td></td>
<td>9.5</td>
<td>474</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>0 = White-collar</td>
<td>75.7</td>
<td>4831</td>
<td></td>
<td>79.6</td>
<td>3969</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 = Blue-collar</td>
<td>21.4</td>
<td>1363</td>
<td></td>
<td>18.5</td>
<td>923</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>0 = Western-European</td>
<td>77.3</td>
<td>4929</td>
<td></td>
<td>81.9</td>
<td>4083</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 = Non-West-European</td>
<td>19.5</td>
<td>1241</td>
<td></td>
<td>15.4</td>
<td>767</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>10-15</td>
<td>12.23</td>
<td>.51</td>
<td>6367</td>
<td>12.17</td>
<td>.46</td>
<td>4978</td>
<td></td>
</tr>
</tbody>
</table>

Note. Longitudinal = Students who filled out questionnaires during all three waves.
Sample Characteristics

In the first wave, a total of 6380 students filled out the questionnaire. As has already been shortly discussed when comparing the sample to the population (see Table 2), little over half of the sample were boys (53.8%) (see Table 3). A minority of the students (11.7%) attended the vocational-oriented track. Since students from 7th grade were targeted, the mean age was 12.23 years. Nevertheless, there was a range from 10 to 15 years old, indicating that some pupils had skipped years, while others had repeated years. Nevertheless, 76.5% was twelve years old, indicating that the majority of students were on track. The socio-economic background (SES) of the students was assessed by considering the employment situation of the parents. About one fifth of the sample were students whose parents performed manual labor (labeled blue-collar). Most students (77.3%) were from Western-European descent, as indicated by the birth place of their maternal grandmother. Students who indicated that their grandmother was born outside Western Europe, had predominantly Moroccan (5.4%), Turkish (2.9%), or Eastern-European (3.4%) roots, reflecting the migration history of Belgium.

We see some small differences when comparing the respondents from the first wave \(N = 6380\) with those who completed questionnaires during all three waves \(N = 4987\) (see Table 3). Students who were present during all three waves of the study were somewhat more often from Western-European descent, with a white-collar background and attended the academic track more frequently. This was to be expected, since only students who provided answers during all three data collections are present in this longitudinal dataset. Hence, students who were sick during one of the survey-moments, or who transferred schools are no longer present in the longitudinal dataset, resulting in a dataset that is somewhat more privileged than the original cross-sectional dataset.

When considering school characteristics (see Table 4), the average number of students in first grade was 108. There was a pretty large standard deviation of 79.9, indicating a wide variety in school size across the sample (as is also evident from the range in school size). There was also considerable variation in student composition across the schools in the samples, with some schools consisting solely of boys, and almost entirely of students with a Western-European descent or white-collar background. Note that even though two schools dropped out in the course of the study, this had little impact on the overall demographical make-up of the schools in the study (see Table 5).
Table 4.
School characteristics at Time 1 of all participating schools

<table>
<thead>
<tr>
<th>School-level Variables (N = 59)</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>School size</td>
<td>108.14</td>
<td>79.86</td>
<td>11</td>
<td>316</td>
</tr>
<tr>
<td>Gender composition (% of girls)</td>
<td>.438</td>
<td>.223</td>
<td>.0</td>
<td>.85</td>
</tr>
<tr>
<td>Ethnic composition (% of ethnic minority)</td>
<td>.251</td>
<td>.257</td>
<td>.01</td>
<td>.95</td>
</tr>
<tr>
<td>SES composition (% of students with blue-collar background)</td>
<td>.29</td>
<td>.192</td>
<td>.01</td>
<td>.87</td>
</tr>
</tbody>
</table>

Table 5.
School characteristics at Time 1 of longitudinally participating schools

<table>
<thead>
<tr>
<th>School-level Variables (N = 57)</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>School size</td>
<td>110.47</td>
<td>80.25</td>
<td>11</td>
<td>316</td>
</tr>
<tr>
<td>Gender composition (% of girls)</td>
<td>.441</td>
<td>.227</td>
<td>.0</td>
<td>.85</td>
</tr>
<tr>
<td>Ethnic composition (% of ethnic minority)</td>
<td>.246</td>
<td>.258</td>
<td>.01</td>
<td>.95</td>
</tr>
<tr>
<td>SES composition (% of students with blue-collar background)</td>
<td>.282</td>
<td>.186</td>
<td>.01</td>
<td>.87</td>
</tr>
</tbody>
</table>

VARIABLES

It should be noted that a consistent approach was used for the construction of all variables in this dissertation. That is, scales were only computed when the number of missings on items was 25% or less. When respondents had more than 25% missings, they received a missing on the scale. No imputation techniques were used to artificially reduce the number of missings; and all analyses employed a list-wise deletion approach in dealing with missing data.

Independent Variables: Aspects of Gender

Gender typicality. Gender typicality refers to the extent to which somebody feels to be a good example of their gender category. Gender typicality was measured by a subscale from the Self-Concept questionnaire by Egan and Perry (2001). The Dutch translation and likert-scale answering format are based on the work of Bos and Sandfort (2010). The scale has separate versions for boys and girls, each consisting of six items with a five-point likert scale (0 = completely disagree, 4 = completely agree). Examples of items are: “I feel that my skills and interests are similar to those of other [girls/boys]”, “I feel that I am a good example of a typical [boy/girl]”). The gender typicality score was computed using the mean, and was
only calculated when respondents had maximally one missing. The scale displayed good internal reliability at all time-points (see Table 6). Gender typicality scores increased somewhat throughout the study, demonstrating that gender identity is still assessed and constructed during early adolescence.

As has been discussed in chapter 4, this scale is an identification scale. This means that researchers do not force a definition of masculinity or femininity upon the respondent. Instead, respondents are free to consider the characteristics they find most pertinent for their felt gender typicality. We have decided to employ an identification scale because multifactorial theory states and research has confirmed that what is most central for felt gender identity tends to differ from person to person (Perry & Pauletti, 2011; Spence, 1993; Tobin et al., 2010), and rarely coincides with the traits that are considered to define masculinity or femininity in identity-measures, such as the Bem Sex Role Inventory (Egan & Perry, 2001).

**Pressure for gender conformity.** Pressure for gender conformity reflects the degree to which one experiences strong expectations to uphold gender-role norms. To measure pressure for gender-conformity, we used a subscale from the Self-Concept questionnaire by Egan and Perry (2001). The scale has separate versions for boys and girls. The Dutch translation and answering format with a 4-point likert scale (1 = completely disagree, 2 = disagree, 3 = agree, 4 = completely agree) are based on the work of Bos and Sandfort (2010). This scale consists of four items related to pressure from peers and four items related to pressure from oneself. Sample items are: “My friends would be upset if I wanted to play with [boys’/girls’] toys”, “I think it is important to act just like other [girls/boys]”. This scale was constructed using mean, with high scores indicating more pressure for gender conformity. The score was only computed when respondents had maximally two missings. The scale displayed good internal reliability at all time-points (see Table 6), and scores remained relatively stable throughout the study.
Table 6.
Univariate characteristics of gender and school functioning variables

<table>
<thead>
<tr>
<th>Individual-level Variables</th>
<th>Wave 1</th>
<th></th>
<th></th>
<th></th>
<th>Wave 2</th>
<th></th>
<th></th>
<th></th>
<th>Wave 3</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cronbach's Alpha</td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
<td>Cronbach's Alpha</td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
<td>Cronbach's Alpha</td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.77</td>
<td>2.48</td>
<td>.67</td>
<td>4930</td>
<td>.86</td>
<td>2.73</td>
<td>.69</td>
<td>4898</td>
<td>.84</td>
<td>2.79</td>
<td>.66</td>
<td>4973</td>
</tr>
<tr>
<td>Gender-conformity pressure</td>
<td>.77</td>
<td>1.46</td>
<td>.56</td>
<td>4909</td>
<td>.81</td>
<td>1.48</td>
<td>.55</td>
<td>4814</td>
<td>.82</td>
<td>1.47</td>
<td>.58</td>
<td>4913</td>
</tr>
<tr>
<td>Well-being</td>
<td>.81</td>
<td>2.87</td>
<td>.5</td>
<td>4953</td>
<td>.86</td>
<td>2.88</td>
<td>.57</td>
<td>4927</td>
<td>.86</td>
<td>2.74</td>
<td>.62</td>
<td>4950</td>
</tr>
<tr>
<td>Autonomous study motivation</td>
<td>.89</td>
<td>2.03</td>
<td>.81</td>
<td>4638</td>
<td>.89</td>
<td>1.97</td>
<td>.75</td>
<td>4966</td>
<td>.89</td>
<td>1.81</td>
<td>.74</td>
<td>4964</td>
</tr>
<tr>
<td>Controlled study motivation</td>
<td>.79</td>
<td>2.08</td>
<td>.70</td>
<td>4641</td>
<td>.74</td>
<td>2.32</td>
<td>.70</td>
<td>4976</td>
<td>.77</td>
<td>2.33</td>
<td>.73</td>
<td>4965</td>
</tr>
<tr>
<td>Academic self-efficacy</td>
<td>.84</td>
<td>2.64</td>
<td>.64</td>
<td>4888</td>
<td>.87</td>
<td>2.48</td>
<td>.70</td>
<td>4879</td>
<td>.85</td>
<td>2.37</td>
<td>.70</td>
<td>4968</td>
</tr>
</tbody>
</table>
Outcome Variables: School Functioning

Subjective well-being. Subjective well-being was measured by asking students how often they had felt certain emotions in the past 30 days. They indicated the frequency of positive and negative feelings on a 5-point scale ranging from 0 = never to 4 = very often (Keyes, Shmotkin, & Ryff, 2002). The scale consisted of 12 items, equally divided between positive (e.g., happy, peaceful) and negative items (e.g., nervous, worthless). Based on principal component analysis, we constructed a single scale from these items. The items pertaining to negative affect were reverse coded so higher scores on the scale indicate a higher well-being. Using mean, the scale was constructed when respondents had maximally three missings. Students’ well-being was stable during 7th grade, but had declined somewhat by the end of 8th grade. The scale displayed good internal reliability at all time-points.

Academic Self-Efficacy. To assess academic self-efficacy, we used an eight-item scale based on Caprara and colleagues (2011). These items concern self-efficacy for self-regulated learning activities, such as the capacity to plan and organize academic activities, to structure environments conducive to learning and to motivate oneself for schoolwork. The scale presents a stem followed by a question. For instance: “How well can you…?”, “finish homework assignments by deadlines?”, “Arrange a place to study without distractions?”. Students answered on a 5-point likert scale, ranging from 0 = cannot do at all, to 4 = can do very well. The scale was construed using mean, so higher scores constitute a higher sense of academic self-efficacy. The scale was only computed when students had two missings or less, and displayed good internal reliability on all time-points (see Table 6). The academic self-efficacy of students declined during the course of the study.

Study motivation. To assess study motivation, we used the Academic Self-regulation Scale (Ryan & Connell, 1989). This scale consists of four subscales, each containing four five-point likert items (0 = completely disagree, 4 = completely agree), measuring external, introjected, identified and intrinsic study motivation (Deci et al., 1991). The scale presents a stem followed by rationales, for instance “I am motivated to study because…”: “others force me to do it” (external), “I would feel ashamed if I didn’t” (introjected), “this is an important goal in my life” (identified), “studying is fun” (intrinsic). Based on principal component analysis and previous research (e.g., Black & Deci, 2000; Haerens, Kirk, Cardon, De Bourdeaudhuij, & Vansteenkiste, 2010; Vansteenkiste et al., 2004), we created two composite scores. First, autonomous motivation was created by averaging the subscales of intrinsic and identified regulation. Second, controlled motivation was created by averaging
the subscales of introjected and external regulation. The scales were only computed when respondents had two missings or less, and the scales displayed good internal reliability throughout all time-points (see Table 6). Students’ sense of autonomous motivation declined throughout the study, while their level of controlled motivation increased in the course of 7th grade, but remained stable throughout 8th grade.

**DESIGN**

To account for the clustering of the data, with students nested in schools, we employed multi-level analysis. The benefit of this method is, in contrast to simple regression analysis, that it correctly accounts for the successive sampling method (Hox, 2010). That is, by first selecting schools and then selecting students within these schools, the sample is no longer random and the observations are no longer independent from each other. Since simple regression analysis cannot account for the fact that students within a single school tend to be more alike than students from different schools, multilevel analysis is used to avoid this problem of autocorrelation.

Additionally, we consider the influence of school characteristics on students’ well-being in chapter 10. The best way to correctly account for the influence of the organizational level on the individual level is by using multi-level analysis. Another benefit is that multilevel analysis can tease out the amount of variance on each level. That is, we can ascertain how much variance is located at the individual versus the school level. This is done by first running a so-called zero-model, where no variables are included, and calculating the variance at each level. In later models, the explanatory and control variables are added. In all upcoming empirical chapters, we first add the explanatory variables before adding the control variables. This way, we can assess whether or not the associations under study change when the control variables are added. Individual-level controls systematically added to the empirical models are age, socio-economic background, ethnicity, school track and initial academic ability. All these variables have been shown to connect to school functioning in international and Flemish research (Bandura, 1997; Hattie, 2008; Meece et al., 2006; Phalet, Andriessen, & Lens, 2004; Schunk & Pajares, 2001; Van Houtte & Stevens, 2010; Vansteenkiste et al., 2009). School-level controls were only added to the model in chapter 10, because this is the only chapter explicitly focusing on the influence of school characteristics.

Because the focus of this research is to explicate the subtle workings of gender and uncovering intrasexual variation, every chapter considers Sex X Gender
interactions. That is, chapter 7 and 8 consider sex and gender typicality interactions, and chapter 9 focuses on sex and pressure for gender conformity interactions. Chapter 10 considers inter- and intrasexual variations by analyzing the associations between well-being, gender typicality and pressure for gender conformity separately for boys and girls. This has been done to avoid three-way interactions, which tend to hamper the understanding of the reader because of their complexity.

Also, as this research was part of the Procrustes-project, different datasets were used for the different chapters. That is, because data collection was ongoing while this research was carried out, later chapters could use more waves than earlier chapters. More specifically, the first two empirical chapters involve cross-sectional analyses on the first wave. Chapter 9 compares in a cross-sectional way associations on the first and the second wave. The last chapter considers all three waves simultaneously, by including time as an extra level in the multi-level analysis (that is, time clustered within individuals, who are, in turn, clustered within schools).
Chapter 7.
Transcending the gender dichotomy in educational gender gap research:
The association between gender identity and academic self-efficacy.

Wendelien Vantieghem, Hans Vermeersch, & Mieke Van Houtte
Published in Contemporary Educational Psychology, 39, 369-378.
Doi: 10.1016/j.cedpsych.2014.10.001

Educational gender gap research tends to focus on differences between boys’ and girls’ achievement. However, substantial variation exists within the sexes concerning this educational achievement. In this study, we investigate the inter- and intrasexual differences in academic self-efficacy in a sample of 6380 Flemish 7th graders collected in the school year 2012-2013. To adequately consider masculinity and femininity on a micro level, we employ the concept of gender identity. Results show that considerable inter- and intrasexual differences exist. In line with the educational gender gap, girls scoring high on self-perceived femininity are at the top of the pack. Self-perceived masculine boys score considerably lower on academic self-efficacy, however, they still do much better than cross-gender boys and girls. The poor results of boys and girls who perceive themselves to be gender atypical are partly explained by the students’ lower well-being. The results are discussed in light of masculinity and femininity.
In recent decades, the gender gap in education has become a problem that has captured the minds of policymakers and researchers alike. It became clear in the nineties that girls had started to outperform boys on several parameters and these trends continue to this day: boys repeat grades more often (Fergusson & Horwood, 1997; Van Landeghem, Goos, & Van Damme, 2010), have lower marks in class (Duckworth & Seligman, 2006; Epstein, Elwood, Hey, & Maw, 1998; Fergusson & Horwood, 1997; Jackson, 1998; Van de Gaer et al., 2006a; Younger & Warrington, 1996), drop out more often (Buchmann, DiPrete, & McDaniel, 2008; Fergusson & Horwood, 1997; Van Landeghem et al., 2010), have lower enrollment in higher education (Buchmann et al., 2008; Van Woensel, 2007) and are overrepresented in special education services and remedial classes (Benjamin, 2003). These findings are not limited to one country, but are encountered all over the western world (Benjamin, 2003; Martínez, Julia, Mari-Klose, & Mari-Klose, 2012) and thus constitute a pervasive challenge.

In order to remedy this underachievement of boys, it is important to understand the processes through which achievement is determined. Academic self-efficacy has proven to be one of the most central predictors of educational achievement. Self-efficacy refers to people’s beliefs in their capability to perform certain tasks (Bandura, 1997; Schunk & Pajares, 2001; Zimmerman, 2000) and has been shown to connect both directly (Bandura, 1997; Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Multon, Brown, & Lent, 1991) and indirectly to academic achievement (Bandura, 1997; Multon et al., 1991; Zimmerman, 2000).

Gender differences in self-efficacy are rather complex and nuanced. For instance, gender differences in self-efficacy only start to occur in adolescence and tend to be mostly subject-specific, with girls having higher self-efficacy in language arts and boys in mathematics (Huang, 2013). These multifaceted findings can be clarified through the framework of masculinities theory. Masculinities theory takes gendered peer cultures into account and hence clarifies how students “do gender” in their everyday lives at school. For instance, certain subjects are considered to be masculine or feminine (Connell, 1996; Martino, 1996). Since students tend to be more sure of themselves in subjects which fit within their gender-category (Pajares, 2002; Pajares & Valiante, 2001; Schunk & Pajares, 2001), this raises boys’ self-efficacy for “masculine” subjects, such as mathematics, and girls’ self-efficacy for language arts.

One of the largest problems with masculinities theory, however, is that it has a group-based approach and therefore neglects intrasexual differences (Francis,
2000, 2010). Despite the general tendency for girls to outperform boys, research has shown there are high-performing boys and low-achieving girls as well. Because of the theory’s focus on general intersexual differences, these female low-achievers and male high-achievers are rendered invisible (Warrington, Younger, & Williams, 2000). We suggest that to adequately understand the mechanisms through which boys and girls end up at the top or the bottom of a class and thus to remedy the educational gender gap, it is important to consider not only intersexual differences, but those intrasexual differences as well.

We put gender identity forward as a concept that could help to account for these intrasexual differences. Gender identity, a central concept in multifactorial theory (see below), refers to the extent to which someone feels to be masculine or feminine, regardless of biological sex, given what it means to be masculine or feminine in a given society (Perry & Pauletti, 2011; Stets & Burke, 2000a; Tobin et al., 2010; Wood & Eagly, 2009). This concept thus builds on masculinities theory by equally considering the impact of masculinity/femininity on behavior, choices, dispositions, and traits. However, unlike masculinities theory which assesses group-based gender cultures, gender identity gauges masculinity and femininity on an individual level and thus goes beyond masculinities theory’s scope.

In this paper we will investigate the way sex and gender identity link to academic self-efficacy. In should be noted that we will consider self-efficacy for self-regulated learning. This form of self-efficacy is less gender biased than domain-specific self-efficacy (Huang, 2013), such as self-efficacy for mathematics or language, and is more pertinent to general academic achievement than self-efficacy in a specific school subject (Pajares, 2002; Pajares & Valiante, 2001). More specifically, we will consider the interaction between sex and gender identity to investigate inter- and intrasexual differences regarding the association of self-perceived masculinity/femininity on self-efficacy. This way, we hope to better understand the underlying processes that lead to girls’ and boys’ educational success.

**SELF-EFFICACY AND ACHIEVEMENT**

Self-efficacy refers to the beliefs and confidence that one can perform certain tasks or behaviors and is grounded within the larger theoretical framework known as social cognitive theory. This theory proposes that human achievement is dependent on one’s behaviors, internal personal factors (such as cognitive, affective and biological events) and environmental conditions (Bandura, 1997; Schunk & Pajares, 2001). Hence, social cognitive theory encompasses a large array of motivators and
regulators of social, behavioral and cognitive capabilities. Self-efficacy occupies a central role in this theory, because it acts upon several of these determinants. For instance, self-efficacy influences choice of activities, motivational level, execution of capabilities and so on (Bandura, 1997). Considering this broad-ranging impact of self-efficacy, it will come as no surprise that self-efficacy is considered to be one of the most important contributors to academic achievement (Hattie, 2008; Schunk & Pajares, 2001; Yusuf, 2011). A meta-analysis has shown, for instance, that self-efficacy beliefs account for 14% of variation in student’s academic performance with an effect-size of .38 (Multon et al., 1991). Moreover, studies have shown that it impacts on achievement both directly and indirectly, as will be discussed below.

Firstly, self-efficacy contributes directly to achievement (Bandura, 1997; Bandura et al., 1996; Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011; Greene, Miller, Crowson, Duke, & Akey, 2004; Multon et al., 1991; Yusuf, 2011), because self-efficacy enables people to effectively use their perceived skills (Bandura, 1997), according to the observed demands of the situation (Salomon, 1984). Self-efficacy beliefs thus function as a sort of self-fulfilling prophecy, by affecting how consistently and effectively people apply what they know (Bandura, 1997). Consequently, it turns out to be a better predictor of performance than skills or intellectual aptitude alone. For instance, the association between self-efficacy and achievement remains, even when controlling for prior achievement (Yusuf, 2011); And, in children of the same ability level, more efficacious children obtained better scores than their less efficacious counterparts (Bandura, 1997).

Secondly, self-efficacy also contributes indirectly to academic achievement through important learning behaviors, such as strategic thinking and motivational processes. For instance, it has been shown that high-efficacious students manage their work-time better (Bandura, 1997; Pajares, 2002), are more persistent (Bandura, 1997; Multon et al., 1991; Zimmerman, 2000), derive more pleasure and enjoyment from their tasks, (Bandura, 1997) and employ more deep cognitive processing (Pintrich & De Groot, 1990; Walker, Greene, & Mansell, 2006). Efficacious students also set higher aspirations for themselves (Bandura, 1997; Bandura et al., 1996; Pajares, 2002; Zimmerman, 2000) and are less likely to suffer from test anxiety (Bandura, 1997; Pintrich & De Groot, 1990) or other negative emotional reactions (Bandura, 1997; Bandura et al., 1996; Zimmerman, 2000).

In conclusion, through both its direct and indirect influence, self-efficacy has an encompassing impact on academic achievement and is an important variable to consider in educational research.
SELF-EFFICACY AND GENDER

Since self-efficacy is a central factor in educational achievement, it is interesting to see how it can contribute to the explanation of the gender gap in academic attainment. Most research finds some gender differences, however, these results were often inconsistent (Meece, Glienke, & Burg, 2006). Huang (2013) performed a meta-analysis, which included over 200 studies, to better answer the question regarding gender differences in self-efficacy. Her results show that age is an important moderator. Apparently, gender differences in self-efficacy only start to occur in early adolescence and increase with age, while there are usually no gender differences in elementary school (Huang, 2013).

Not only age is an important moderator for gender differences in efficacy beliefs, the school subject turns out to be a central moderator as well (Huang, 2013). Boys have higher self-efficacy in mathematics (effect size ($\beta$) = .18) and computer sciences ($\beta$ = .18). Girls, on the other hand, have a higher self-efficacy in language arts ($\beta$ = .16) and a small advantage in general academic self-efficacy ($\beta$ = .03). Instead of investigating self-efficacy for specific school subjects, Pajares (2002) investigated self-efficacy for self-regulated learning. His results show that girls usually outperform boys in their capability to organize their work, employ meta-cognitive strategies and show effort management.

However, as Pajares and Valiante (2001) justly note, gender differences cannot be accounted for within social cognitive theory. They state: ‘Social cognitive theory does not endow gender self-beliefs with agentic and motivating properties. And neither does it endow gender itself with such properties” (Pajares & Valiante, 2001, p. 12). So, researchers usually end up accounting for these differences on an ad-hoc basis, without linking these to an overarching theory. They usually refer to the perceived, stereotypic nature of the school domains in which gender differences are observed. Mathematics and computer sciences are seen as male domains, so boys wish to excel in these subjects because their self-perceptions are infused with the notion that success is a requirement (Pajares, 2002; Pajares & Valiante, 2001; Schunk & Pajares, 2001). Conversely, girls’ mathematics’ self-efficacy would be lower because they feel that this is a male domain. This would imply that it is harder for girls to excel in mathematics and this notion would then undermine both girls’ confidence and interest in this subject. Pajares and Valiante (2001) confirmed that these gender differences are greater when someone holds more gender stereotypic beliefs by demonstrating how a feminine gender orientation accounts for the gender differences in writing self-efficacy.
While few studies make this connection, these explanations and gender differences could be perfectly framed within masculinities theory, as will be discussed in the next section.

**CONNECTING MASCULINITIES THEORY TO SELF-EFFICACY**

Masculinities theory puts gender differences in educational outcomes in a broader context by situating it within cultures of masculinity and femininity. In schools, this takes the shape of gendered peer cultures. In this theory, the concept of hegemonic masculinity takes central stage. Hegemonic masculinity is the dominant form of masculinity, which is superior in the gender order (Connell, 1996). In order for hegemonic masculinity to exist, it must define itself against what it is not, what it is superior to. Therefore, masculinity is constructed within this gender order against subordinated “others” (Brutsaert, 2006; Connell, 1996; Epstein, 1997; Schippers, 2007). These others include not only femininity, but also marginalized masculinities, such as homosexuality. Often, these marginalized masculinities would be conflated with femininity, thereby firmly ensuring their subordinate position (Connell, 1996; Schippers, 2007).

Masculinities theory shows how gendered peer cultures have a profound impact on the way children enact masculinity and femininity, and “do gender” in their everyday lives at school. For instance, studying, neatness and an inquisitive attitude are usually regarded as feminine in a school setting (Epstein, 1998; Martino, 1999). Boys who do work hard at school or cannot compensate good marks with “appropriate masculine behavior”, such as excelling in sports or being the class clown, get ridiculed and called names such as “poofter” or “fag” (Epstein, 1998; Martino, 1999; Stoudt, 2006; Swain, 2005; Warrington et al., 2000). Masculinities theory has proven valuable in understanding gender differences in school attitudes and achievement. This approach, for instance, succeeds in explaining why age and school subject are important moderators for academic self-efficacy. Firstly, certain subjects are considered to be “feminine”, such as language, arts or home economics (Connell, 1996; Martino, 1996). Consequently, boys do not want to be caught liking these subjects or even paying attention for fear of being associated with femininity. This is reflected in their lower self-efficacy for these subjects. Students tend to feel safer and more at ease in subjects which fit within their sex-category (Pajares, 2002; Pajares & Valiante, 2001; Schunk & Pajares, 2001), which conversely raises boys’ self-efficacy for “masculine” subjects, such as mathematics and computer sciences (Connell, 1996; Martino, 1996). Secondly, this theory succeeds in explaining why these differences only come to play when adolescence
kicks in. Adolescence is the life phase where parents’ importance starts to become less paramount and peers start to take center stage (Coleman & Johnstone, 1981; Larson & Richards, 1991; Van Houtte, 2004b). Consequently, this is the time where the impact of —gendered— peer cultures becomes most notable.

Masculinities theory not only succeeds in theoretically framing gender differences in self-efficacy, it also links to the mechanisms which contribute to self-efficacy, such as vicarious learning, mastery experiences and persuasion. Firstly, vicarious learning refers to obtaining a heightened sense of self-efficacy for an activity by watching others successfully perform it (Bandura, 1977). Of course, the higher the perceived similarity with the model, the more their success contributes to an individual’s self-efficacy (Schunk & Pajares, 2001). Masculinities theory considers this process from a gendered perspective and frames how boys or girls might feel less interest, self-efficacy and motivation for courses which they deem are “not for them”, because they do not see other boys/girls signing on for these courses or succeeding in them. This is confirmed through research that shows that perceived similarity with someone in a certain field of study was the strongest mediator between sex and interest, and mediated men’s interest in English (Cheryan & Plaut, 2010).

A second contributor to self-efficacy is persuasion. Persuasion refers to people being led, through suggestion, to believe that they can successfully perform a certain activity (Bandura, 1977). This is often interpreted as persuasion by another individual, such as a friend, parent or therapist. However, since masculinities theory usually considers the group level, this could be interpreted more broadly as suggestions from society, such as stereotypes. Several studies have shown how both parents and teachers —unintentionally— transmit the reigning stereotypes that mathematics is harder for girls, undermining girls’ confidence and performance (Gunderson, Ramirez, Levine, & Beilock, 2012; Meece et al., 2006; Tiedemann, 2000).

A third contributor is mastery experiences. In contrast to persuasion and vicarious learning which are second-hand indications of efficacy, mastery experiences refer to an individual experiencing success for him/herself. Success raises mastery experiences and consequently self-efficacy, while repeated failures lower them (Bandura, 1977). While this has rarely been tested within masculinities theory because of the dominance of ethnographic studies within the field, a central tenet of masculinities theory does tie into the impact of mastery experiences. Masculinities theory recognizes several variations of ideal or hegemonic masculinity within different contexts (Connell, 1989; Mac an Ghaill, 1994). These different forms of masculinity are often crosscut by socio-economic position and/or race.
and tend to emphasize different aspects of the masculinity concept. For instance, boys with a higher socio-economic position tend to have masculinity concepts revolving around responsibility and competence, which better allow the acceptance of good marks (Connell, 1989). Lower SES boys, on the other hand, tend to emphasize the pride and aggression aspects of masculinity and have a stronger anti-school attitude. It is assumed within masculinities theory that this may stem from the differential experience of success for children from different social backgrounds (Epstein, 1998; Jackson, 2003; Mac an Ghaill, 1994). Since children from lower economic classes tend to have less mastery experiences at school, they are assumed to emphasize other aspects of masculinity as a self-protection strategy.

While, as discussed above, masculinities theory is able to contribute some valuable insights and extensions to self-efficacy research, it has several drawbacks as well. For instance, there is a regrettable lack of attention for femininities in both research and theory (Connell & Messerschmidt, 2005; Schippers, 2007), thereby effectively handicapping one half of the research into educational gender differences (that is, the research into the achievement of girls). Secondly, masculinities theory has a group-based approach, which leads to overlooking nuances and intra-sexual differences (Francis, 2000, 2010).

**GENDER IDENTITY**

We propose that research on gender identity could remedy the earlier-discussed shortcomings of masculinities theory. Unlike masculinities theory, gender identity is an individual-based characteristic that accounts for both masculinity and femininity and can be assessed in a qualitative and quantitative fashion. The concept of gender identity, which is central in multifactorial theory, refers to the extent to which someone feels masculine or feminine, regardless of biological sex (Perry & Pauletti, 2011; Stets & Burke, 2000a; Tobin et al., 2010; Wood & Eagly, 2009). According to multifactorial theory, gender identity is bipolar. This means that people can pinpoint their place on a continuum with masculine at the one hand of the line and feminine on the other (Spence, 1984, 1999). The theory further states that people would base their sense of gender identity on a self-assessment of their attitudes, behavior, traits, looks, interests, and so on. This assessment is idiosyncratic, indicating that what is most central to someone’s gender identity tends to differ from person to person (Perry & Pauletti, 2011; Spence, 1993; Tobin et al., 2010) (for a more extensive review of multifactorial theory, see Vantieghem, Vermeersch and Van Houtte [2014b]).
Interestingly, the concept of gender identity, and the multifactorial theory from which it is derived, are easily compatible with the identity-based motivation model (IBM) (Oyserman & Destin, 2010), a model which is used to explain academic (under)achievement of certain groups. While this model has mostly been employed to explain the impact of class and ethnic identities (e.g., Oyserman, Bybee, & Terry, 2006; Oyserman & Destin, 2010), it can also be used to ascertain the influence of gender identity (Elmore & Oyserman, 2012). According to the model, identities that are salient in certain contexts determine interpretation of contextual cues and which courses of action come to mind. This way, identities guide behavior (Elmore & Oyserman, 2012; Oyserman & Destin, 2010). For instance, when boys feel that studying is not “masculine” or typical for boys, they will interpret difficulty with homework as a signal that studying does not come natural to boys. Conversely, if studying is considered to be part of the gender identity, difficulty can be interpreted as a sign that the task is worthwhile, not impossible. The authors have shown how children develop more academically oriented goals and how boys persist longer at school tasks by connecting their gender category to success (Elmore & Oyserman, 2012).

While Oyserman and colleagues’ identity-based motivation model has proven to be an interesting framework, it tends to be mostly used in experimental research where certain identities are primed or in interventions which adapt identity contents (e.g., Elmore & Oyserman, 2012; Oyserman et al., 2006; Oyserman & Destin, 2010). This way, the research employing IBM has remained mostly group-based and, just as masculinities theory, has ignored intrasexual differences. Interestingly, the model does provide the possibility to study intrasexual variations via the postulate of dynamic construction of identity content (Oyserman & Destin, 2010). This postulate states that what identities are taken to depends on which traits a person associates with the social group, in casu gender groups. As such, intrasexual variation in identity content is explicitly part of the model, yet it has remained underresearched where gender identity is concerned (for an application of intragroup differences in ethnic identity see Oyserman, Kemmelmeier, Fryberg, Brosh, & Hart-Johnson, 2003). By considering individual variations in gender identity through survey research, this gap could be remedied. However, in survey research, gender identity has only rarely been connected to educational outcomes. It has mostly been used in research regarding children’s psychological adjustment, demonstrating that atypical children tend to feel worse (e.g., Carver, Yunger, & Perry, 2003; Egan & Perry, 2001; Yunger, Carver, & Perry, 2004). When it has been used in educational research, results show that gender typicality was not associated with schoolgirls’ general motivation for STeM or English (Leaper, Farkas, &
Brown, 2012) nor with their academic achievement (Patterson & Pahlke, 2011). On the other hand, it did relate to college boys’ academic motivation and choice for gender (a)typical courses (Leaper & Van, 2008). These empirical findings suggest that there would be a relation for boys, but not for girls. However, girls and boys were not simultaneously included in the samples of these studies, which is necessary in order to make claims about gender-specific effects. Furthermore, there was a selection bias and sizeable age difference between the samples of these studies (10- to 18-year-old school girls versus college boys), which further complicates comparing the results. Hence, further research is necessary to assess these possible sex-specific connections between gender identity and educational parameters.

RESEARCH QUESTIONS AND HYPOTHESES

As discussed above, the survey-research that connects gender identity with educational outcomes has so far mainly focused on general motivation or achievement itself. To our knowledge, research regarding the link between gender identity on self-efficacy is non-existent. Research regarding sex differences in self-efficacy is more prevalent, however, and has shown that such differences are usually non-existent in elementary school, but increase throughout adolescence (Huang, 2013). Hence, it would be interesting to investigate the self-efficacy of children who start secondary school and are at the beginning of adolescence. We could expect that any gender differences found at this age are wont to increase throughout the years. As such, considering gender differences among 12- to 13-year-olds poses a conservative test of the association between sex, gender identity and self-efficacy. Consequently, the first research question of this paper is: (1) how do sex and gender identity link to academic self-efficacy of children in the first year of secondary education? It should be noted that self-efficacy can be assessed on different academic domains, such as computer sciences, language, and so on, with quite different results. In this study we will consider self-efficacy for self-regulated learning. This specific form of self-efficacy has several advantages. First, because self-regulated learning comprises skills necessary for academic achievement in any school course, it is more pertinent to general achievement than self-efficacy in a specific school subject (Pajares, 2002; Pajares & Valiante, 2001). Furthermore, it is less gender biased than self-efficacy for a specific school subject (Huang, 2013). Nevertheless, in line with previous research findings (Huang, 2013; Pajares, 2002), girls are still expected to exhibit a somewhat higher self-efficacy for self-regulated learning than boys. For the hypothesis on the connection between gender identity
and self-efficacy, we rely on empirical results from Leaper and Van (2008) and an IBM-study (Elmore & Oyserman, 2012), which both show positive connections between gender identity and educational parameters. Hence, we expect higher levels of self-efficacy among gender typical children.

Additionally, the earlier-cited studies suggest that there would be no association between gender identity and girls’ motivation or achievement (Leaper et al., 2012; Patterson & Pahlke, 2011), whereas there would be among boys (Leaper & Van, 2008). Therefore, the second research question is: (2) Does the association between gender identity and self-efficacy differ for boys and girls? It should be noted, however, that it is difficult to infer clear sex-specific associations from these studies because of sample issues (as discussed above). Moreover, masculinities theory and the identity-based motivation model would predict quite different outcomes. Both frameworks state that, unlike the masculine identity, the feminine identity does allow academic achievement and striving (Connell, 1996; Elmore & Oyserman, 2012; Epstein, 1997; Schippers, 2007). Consequently, it could be hypothesized that this leads “feminine” girls to estimate their academic capabilities as higher than “masculine” boys would. Furthermore, when it comes to interpretations of difficulty, girls are more likely to judge difficulties with academic tasks as an indication that the behavior is important, not impossible (Elmore & Oyserman, 2012). Boys on the other hand are more prone to judge these difficulties as a confirmation that academic success is “not for them”. Such a mindset hampers persistence, reducing mastery experiences and undermining self-efficacy beliefs (Bandura, 1997). Hence, a double hypothesis exists here. On the one hand, based on empirical survey studies (Leaper et al., 2012; Leaper & Van, 2008; Patterson & Pahlke, 2011), we could expect a stronger association between gender identity and self-efficacy for boys than for girls. On the other hand, based on masculinity theory and the IBM model, we could expect a stronger relation for girls.

Lastly, we also wish to break through the gender dichotomy that pervades gender research by equally considering intrasexual differences. Hence, the last research question is: (3) Are there important intrasexual differences for the relation between gender identity and self-efficacy? We could expect children who feel they deviate from the gender norms to score quite differently on self-efficacy than children who perceive themselves to be quite masculine or feminine. To our knowledge, no research has ever considered this question before. However, given the research that demonstrates the lower well-being of atypical children (Carver et al., 2003; Egan & Perry, 2001; Yunerger et al., 2004) and that well-being has been shown to impact on self-efficacy (Bandura, 1997), we expect atypical children to score lower on self-efficacy.
METHODS

Data
The data are part of the “Teaching in the bed of Procrustes” - investigation and were gathered in the first half of the school year 2012-2013. Schools were selected so each geographical region within Flanders (Belgium) was equally represented. Selection was further dependent on school denomination and a proportional representation of rural versus city schools. Within these parameters, three random samples were drawn. For each school which refused, a matched school from the next random sample was contacted. In the end, of the 124 contacted schools 59 schools participated in the study, which translates to a response rate of 47.6%. Within these schools, all 7th graders were asked to complete the questionnaire. Researchers were always present during the completion of the survey to explain the procedure and answer questions of students. The response rate among the students was 96.9%, translating to a total of 6380 students. The response-rate is so high since only students who were not present during the visit, for example due to sickness, did not complete the questionnaire. This study design guaranteed that we obtained data from 12 to 13-year-old pupils from various backgrounds, and from several regions in Flanders. Of the students in the sample 46.2% are girls.

Design
In this study, we took self-efficacy as the dependent variable and investigated how this differed between boys and girls in our sample (see hypothesis 1). Furthermore, the earlier-cited investigations suggested there are important differences between boys and girls for the association between gender identity and education-related parameters (Leaper et al., 2012; Leaper & Van, 2008; Patterson & Pahlke, 2011). However, due to sample issues (as discussed above), it is hard to assess such sex-specific associations. Our sample, which included both boys and girls of the same age, allowed us to make intersex comparisons and thus helped to account for this important hiatus in research (see hypothesis 2). Since intra-sexual differences were a main focus of this study, we also considered sex and gender identity interactions. This way, we could see how self-perceived femininity and masculinity helped to account for intra-sexual differences (see hypothesis 3).

To account for the nested structure of the data, we employed multilevel analysis, using the statistical program MLwiN. First, we ran a zero-model, where no variables were included. This way, we could ascertain how much variance was located at the individual versus the school level. In the subsequent models, variables were included. All continuous variables were grand-mean centered to
improve interpretability of the intercept (Hox, 2010). As a second step, we ran a model with sex as main determinant, to assess the extent of the gender gap in academic self-efficacy. In the next model, we added gender identity to ascertain the influence of self-perceived masculinity and femininity, regardless of biological sex. In the following model, we included the interaction effect between sex and gender identity, to determine whether gender identity is associated differently with self-efficacy for boys and girls. In the last model, we included the control variables.

We controlled for socio-economic position, initial scholastic ability, subjective well-being, school track and age. We controlled for well-being since gender atypical children tend to feel less happy about themselves (Egan & Perry, 2001) and well-being has been shown to impact on self-efficacy through the way events are interpreted and remembered (Bandura, 1997). For instance, people who are feeling down tend to remember past failings rather than success, which lowers their self-efficacy. Since self-efficacy functions as a self-fulfilling prophecy, people who underestimate their capabilities, also tend to perform worse and engage in less challenging tasks, reducing further developmental possibilities (Bandura, 1997; Bandura et al., 1996). We also controlled for age at the time of the questionnaire, since self-efficacy beliefs tend to better reflect actual capabilities when students get older, while younger students tend to be overconfident (Schunk & Pajares, 2001). Lastly, we controlled for which track the children attended. In the Flemish educational context, there are two tracks in the first grade: the general and the vocational-oriented track (Van Houtte, Demanet, & Stevens, 2012). The tracks are commonly regarded hierarchically, with the vocational track at the lower end, and lower self-efficacy is commonly found in lower ability groups (Schunk & Pajares, 2001).

Variance components were random for the intercept and individual error term. Variance components for other variables were kept fixed, since there was no significant random slope variance in preliminary models, indicating that the associations did not fluctuate in different schools. All assumptions for linear regression models have been tested and were upheld in the analysis.

**Variables**

**Dependent variables.** To assess academic self-efficacy, we used an eight-item five-point LIKERT scale based on Caprara and colleagues (Caprara et al., 2011). These items concern self-efficacy for self-regulated learning, such as the capacity to plan and organize academic activities, to structure environments conducive to learning and to motivate themselves for schoolwork.
The scale was construed using mean sum of scores and the Cronbach’s alpha was .84. On average, pupils scored 2.62 on this scale (standard deviation (SD) = .66), with a range from 0 to 4. (Table 7)

**Independent variables.** To measure gender identity, we employed the gender typicality subscale from the self-concept questionnaire by Egan and Perry (Egan & Perry, 2001). This measure consists of six items with a five-point LIKERT scale tapping the extent to which one feels to be typical for their gender (sample items: “I feel that my skills and interests are similar to those of other [girls/boys]”, “I feel that I am a good example of a typical [boy/girl]”). Thus, when interpreting this measure in light of masculinity/femininity: typical boys would feel to be masculine and typical girls feel to be feminine. On the other hand, atypical children feel that they deviate from what is expected for their gender. The gender typicality scale employed here is an identification scale, which is in accordance with multifactorial theory and means that researchers do not force a definition of masculinity or femininity upon the respondent. Instead, respondents are free to consider the characteristics they find most pertinent for their felt gender identity. We have decided to employ an identification scale, since multifactorial theory states and research has confirmed that what is most central for felt gender identity tends to differ from person to person (Perry & Pauletti, 2011; Spence, 1993; Tobin et al., 2010) and rarely coincides with the traits that are considered to define masculinity or femininity in identity-measures, such as the Bem Sex Role Inventory (Egan & Perry, 2001).

This scale was constructed using mean sum of scores and Cronbach’s alpha was .77. On average, pupils rated themselves 2.47 on this scale (SD = .68), with a range from 0 to 4. (Table 7)

**Control variables.** The socio-economic position of the students was derived from the work situation of the parents. We asked students to fill out the last or current employment of both their father and mother. Where the information provided by the children was missing or insufficient, we relied on data provided by the parents themselves in another questionnaire. Employment was attributed a score based on the EGP-classification (Erikson & Goldthorpe, 2002) ranging from 1 (unskilled manual labor) to 8 (high-grade professionals and managers). The child received the highest SES-score of both parents, reflecting the dominant socio-economic position of the family as a whole (Erikson & Goldthorpe, 1992; Forehand, Middletin, & Long, 1987). In this analysis, this measure is used as a continuous variable with a mean score of 5.04 (SD = 1.87) (Table 7).
### Table 7.
Descriptive statistics for variables: Frequencies (%), means, standard deviation (SD), Cronbach’s alpha and N.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total sample</th>
<th>Girls</th>
<th>Boys</th>
<th>Mean difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>Mean</td>
<td>SD</td>
<td>alpha</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>2.62</td>
<td>.66</td>
<td>.84</td>
<td>6214</td>
</tr>
<tr>
<td>Sex (0 = boys)</td>
<td>53.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender typicality</td>
<td>2.47</td>
<td>.68</td>
<td>.77</td>
<td>6284</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>5.04</td>
<td>1.87</td>
<td></td>
<td>6194</td>
</tr>
<tr>
<td>Well-being</td>
<td>2.83</td>
<td>.52</td>
<td>.81</td>
<td>6314</td>
</tr>
<tr>
<td>Initial ability: math</td>
<td>26.79</td>
<td>10.77</td>
<td></td>
<td>5982</td>
</tr>
<tr>
<td>Initial ability:: Dutch</td>
<td>14.41</td>
<td>5.37</td>
<td></td>
<td>5749</td>
</tr>
<tr>
<td>Age</td>
<td>12.23</td>
<td>.51</td>
<td></td>
<td>6367</td>
</tr>
<tr>
<td>Track (0 = general)</td>
<td>88.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *p < .05. **p < .01. ***p < .001
Subjective well-being was measured by asking students how they had felt in the last 30 days. They indicated on a five-point scale ranging from “never” to “very often” the frequency of positive (satisfied, happy, peaceful, …) and negative feelings (nervous, worthless, …) (Keyes, Shmotkin, & Ryff, 2002). The scale consisted of 12 items, equally divided between positive and negative items. Based on principal component analysis, we constructed a single scale from these items. The items pertaining to negative affect were reverse coded so higher scores on the scale indicate a higher well-being. This scale was constructed using a mean sum of scores and has a Cronbach’s alpha of .81. Pupils scored 2.83 on this scale (ranging from 0 to 4), with a standard deviation of .52 (Table 7).

Initial ability was measured by a test in mathematics and in Dutch (the students’ first language), which has given valid scores for students of this age. The mathematics test consisted of 50 questions regarding problematical cases and calculations (Dudal, 2003). The Dutch test consisted of four texts for which the children had to answer questions concerning content and vocabulary (CITO, 2010). Pupils scored on average 26.79 (SD = 10.77) on the mathematics test and 14.41 (SD = 5.37) on the language test (Table 7).

We also controlled for age at the time of the questionnaire. The mean age in this sample was 12 years old, with a range from 10 to 15 years old, indicating that some pupils skipped a year while others had to repeat—several—years. However, 76.6% of the sample was twelve years old, demonstrating that the majority of the pupils were on track (Table 7).

Lastly, we controlled for whether the children attended the general versus vocational-oriented track. In our sample, 11.7% of pupils were in the vocational-oriented track.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Self-efficacy</th>
<th>Gender typicality</th>
<th>SES</th>
<th>well-being</th>
<th>Math</th>
<th>Dutch</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy</td>
<td>—</td>
<td>.136***</td>
<td>-.008</td>
<td>.289***</td>
<td>.062***</td>
<td>.022</td>
<td>.001</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.255***</td>
<td>—</td>
<td>.03</td>
<td>.222***</td>
<td>.085***</td>
<td>.051**</td>
<td>-0.048**</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td>.017</td>
<td>.05**</td>
<td>—</td>
<td>.52**</td>
<td>.318***</td>
<td>.318***</td>
<td>-0.26***</td>
</tr>
<tr>
<td>Well-being</td>
<td>.325***</td>
<td>.293***</td>
<td>.097***</td>
<td>—</td>
<td>.114***</td>
<td>.095***</td>
<td>-0.088***</td>
</tr>
<tr>
<td>Initial ability: math</td>
<td>.082***</td>
<td>.084***</td>
<td>.33***</td>
<td>.108***</td>
<td>—</td>
<td>.552***</td>
<td>-0.341***</td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td>-.01</td>
<td>.034</td>
<td>.334***</td>
<td>.113***</td>
<td>.582***</td>
<td>—</td>
<td>-0.315***</td>
</tr>
<tr>
<td>Age</td>
<td>-.05**</td>
<td>-.045</td>
<td>-.336***</td>
<td>-.087***</td>
<td>-.363***</td>
<td>-.334***</td>
<td>—</td>
</tr>
</tbody>
</table>

*Note. Correlations for boys are above the diagonal; Correlations for girls are below the diagonal. *p < .05. **p < .01. ***p < .001
RESULTS

The zero-order correlations showed that the associations between the variables did not differ much between boys and girls, with the exception of the association between gender typicality and self-efficacy (see Table 8). Table 8 shows that the positive association between gender typicality and self-efficacy was almost twice as large for girls as it was for boys. This positive association is in accordance with the first hypothesis, and the stronger association for girls supports the second hypothesis derived from masculinities theory and the IBM model. These connections and the third hypothesis have been more thoroughly examined in the subsequent multilevel regression analysis.

The zero-model from the multilevel analysis showed that academic self-efficacy varied somewhat between schools, however only 1.85% ($e_{0ij} = .425$, $u_{0j} = .008$, $p < .001$) of total variance was situated at the school-level. We continued using multilevel analysis to correctly account for the nested structure of the data, though we did not introduce school-level variables because of the individually-determined nature of academic self-efficacy.

Model one showed a modest, but significant association between sex and academic self-efficacy ($\gamma = .181$, $p < .001$), indicating that girls scored somewhat higher on academic self-efficacy than boys. In the second model, we included gender typicality. This model showed that both sex ($\gamma = .207; p < .001$) and gender typicality ($\gamma = .181; p < .001$) were positively related to academic self-efficacy (Table 9), confirming hypothesis 1. This indicates that both girls and gender typical students scored higher on academic self-efficacy. Introducing the interaction-term between sex and gender typicality ($\gamma = .100; p < .001$) in model three showed that the influence of typicality significantly differed according to students’ sex. Gender atypical students, both boys and girls, scored similarly low (see Figure 6). However, on the other side of the spectrum, a sizeable difference in academic self-efficacy between typical boys and girls appeared, with typical girls outperforming boys. This indicates that the effect of gender typicality is larger for girls, supporting the second hypothesis derived from masculinities theory and the IBM model. Introducing the control variables in the last model did not change the overall significance of the earlier model. Nevertheless, even though gender typicality remained significant, its effect size had diminished ($\gamma = .096; p < .001$). Analysis showed this to be mostly due to the introduction of subjective well-being ($\gamma = .348; p < .001$), confirming the third hypothesis.

All in all, the final model explained 15.18% of the total variance in academic self-efficacy.
Table 9.
Multilevel analysis: Unstandardized coefficients and standard errors in parentheses

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.518***</td>
<td>2.512**</td>
<td>2.525***</td>
<td>2.516***</td>
<td>2.500***</td>
</tr>
<tr>
<td></td>
<td>(.016)</td>
<td>(.016)</td>
<td>(.016)</td>
<td>(.016)</td>
<td>(.017)</td>
</tr>
<tr>
<td>Sex (0 = boy)</td>
<td>.181***</td>
<td>.207***</td>
<td>.207***</td>
<td>.207***</td>
<td>.231***</td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td>(.017)</td>
<td>(.017)</td>
<td>(.017)</td>
<td>(.018)</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.181***</td>
<td>.134***</td>
<td>.076***</td>
<td>.096***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.012)</td>
<td>(.017)</td>
<td>(.016)</td>
<td>(.018)</td>
<td></td>
</tr>
<tr>
<td>Sex X typicality</td>
<td>.100***</td>
<td>.082***</td>
<td>.063*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.024)</td>
<td>(.023)</td>
<td>(.025)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being</td>
<td>.345***</td>
<td>.348***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.016)</td>
<td>(.017)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Track (0 = general)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.006**</td>
<td>(.002)</td>
<td></td>
</tr>
<tr>
<td>Math</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>.005***</td>
<td>(.001)</td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-.009</td>
<td>(.005)</td>
<td></td>
</tr>
<tr>
<td>Variance components</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School u_{0j}</td>
<td>.007</td>
<td>.007</td>
<td>.006</td>
<td>.008</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.002)</td>
<td>(.002)</td>
<td>(.002)</td>
<td>(.002)</td>
</tr>
<tr>
<td>Student e_{0ij}</td>
<td>.417</td>
<td>.403</td>
<td>.402</td>
<td>.371</td>
<td>.357</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>(.007)</td>
<td>(.007)</td>
<td>(.007)</td>
<td>(.007)</td>
</tr>
</tbody>
</table>

Note. tests represented here were two-sided. N_{School}: 58, N_{Student}: 5269
*: p ≤ .05. **: p ≤ .01 ***: p ≤ .001
DISCUSSION

Throughout the years, the gender gap in education has been widely investigated in several countries and from several viewpoints. As one of these theoretical viewpoints, masculinities theory has helped to clarify the influence of gendered peer cultures and “doing gender” on students’ everyday school behavior and educational outcomes. However, these investigations have necessarily remained qualitative and group-based, because of the scope and characteristics of masculinities theory. Another viewpoint is the identity-based motivation model, which considers the impact of social identities, such as gender or ethnicity, on academic achievement. However, similar to masculinities theory, intrasexual differences have not been the focus of research within the IBM model. This study’s goal therefore was to account not only for intersexual differences in educational achievement, but intrasexual differences as well. To adequately consider masculinity and femininity on a micro level, we introduced the concept of gender typicality (Egan & Perry, 2001).

The analyses show not only that girls and typical students score higher on self-efficacy, confirming the first hypothesis; they also show important intrasexual differences for the association between gender identity and educational outcomes. This way, the current study provides valuable insights and additions to both
masculinities theory and the identity-based motivation model. The results show that girls scoring high on typicality, who feel they live up to the expectations for their gender-category, scored the highest on academic self-efficacy. Typical boys scored notably lower than these typical girls. However, they still outperformed students scoring low on typicality, both boys and girls, who lagged behind. Note that by considering these intrasexual differences in gender identity, these low-achieving girls and rather well-achieving boys have been rendered visible, something most educational gender gap research tends to overlook.

Quite notable in the results is the smaller association between boys’ gender identity and academic self-efficacy than girls’. This was against the hypothesis based on the findings of Leaper, Farkas and Brown (2012), Leaper and Van (2008) and Patterson and Pahlke (2011) who reported an effect of gender identity on boys’ study motivation, but not girls’. However, these investigations have been conducted with samples which consisted of either girls or boys, but never both sexes at the same time. This finding is one of the important contributions of this investigation, since our sample is among the first capable of comparing boys and girls simultaneously on this variable.

The findings thus suggest that the association between gender identity and academic self-efficacy is larger for girls than for boys. This matches the hypothesis based on the IBM model and masculinities theory (see hypothesis 2). This could be explained by a congruous identity content of the typical feminine role and the student role, which might amplify the effect on academic self-efficacy. After all, qualities that characterize the good student tend to match the characteristics of the —stereotypically— feminine girl, such as being calm, cooperative, disciplined, communicative, compliant, tidy, punctual, docile and having an eye for detail (Beaman, Wheldall, & Kemp, 2006; Jones & Myhill, 2004). This would imply that girls who feel to be typically feminine, tend to be better students and thus also score higher on academic self-efficacy.

The other side of the coin could be that boys’ behavior is simply less affected by their gender identity because of the laddish anti-school culture in most schools. It has been shown that boys experience more pressure for gender conformity (Egan & Perry, 2001) and that anti-school cultures tend to have a more severe impact on boys’ school behavior (Demanet, Vanderwegen, Vermeersch, & Van Houtte, 2013; Derks & Vermeersch, 2001; Francis, 2000; Van Houtte, 2004a). Some boys even actively negotiate their masculine image by adapting their pre-school behavior throughout the years, for instance by cooperating less in class (Jones & Myhill, 2004). This implies that boys’ behavior would become less extreme and that they would become more similar to each other, regardless of their
gender typicality. Because of the anti-school focus of most boys’ peer cultures, this convergence of behavior would be especially true for school-related conduct and attitudes, such as academic self-efficacy.

Even though we found a different relation between gender identity and self-efficacy for typical boys than for typical girls, we encountered quite a different picture at the other side of the spectrum. Students scoring low on typicality, both boys and girls, occupied a similarly vulnerable position. It is quite interesting that by employing the concept of gender identity, we have uncovered not only trends which differ between the sexes, such as the stronger association between gender identity and academic self-efficacy for girls than for boys; We have also encountered a unifying trend across the sexes, which juxtaposes students who score high versus low on gender typicality (in that atypical boys and girls are similarly poor off when compared to their peers who do feel more gender typical). Moreover, the present analyses showed that a large part of the association between gender identity and academic efficacy was explained by well-being, confirming the third hypothesis. This suggests that students scoring low on typicality feel worse than their typical counterparts, and this lower well-being is negatively related to their academic self-efficacy. The finding that atypical children score lower on well-being has been found in other studies as well (Carver et al., 2003; Egan & Perry, 2001; Saxvik & Joireman, 2005; Yunger et al., 2004). Several ethnographic accounts of pupil’s lives have demonstrated how showing cross-gender behavior is accompanied by ridiculing and bullying by peers and even teachers (Blakemore, 2003; Epstein, 1998; Renold, 2004; Stoudt, 2006; Swain, 2005). Indeed, schools are often shown to demonstrate a strong preference for traditional gender expressions and an intolerance for a blurring of what is considered “normal” or appropriate (Epstein, 1997; Ueno & McWilliams, 2010). As such, being confronted with an environment that is not supportive or accepting of who they are, it is not surprising that these self-perceived atypical pupils demonstrate a lower sense of well-being. However, it should be noted that well-being did not explain the complete association between gender identity and academic self-efficacy, and so part of this relation remained unaccounted for. Possible avenues for future research are an identity struggle and a heightened pressure for gender conformity. Adolescence is usually regarded as a critical developmental period for identity formation (Erikson, 1968; Faircloth, 2009) and identity formation connects to several important educational parameters, such as achievement (Cadely, Pittman, Kerpelman, & Adler-Baeder, 2011), school belonging (Faircloth, 2009) and academic self-efficacy (Hejazi, Shahraray, Farsinejad, & Asgary, 2009). It is quite possible that this identity formation process poses more of a struggle for gender atypical students, who feel
they deviate from the norm and from what is expected for their gender. Coming to
terms with who they are might not be as self-evident for these self-perceived
atypical students as it is for adolescents who fit within what society and significant
others expect of them, and thus might negatively connect to the self-efficacy of
these atypical students. Furthermore, when these atypical children experience a
strong pressure for gender conformity from parents and/or peers, this could
exacerbate the negative relation with well-being (Egan & Perry, 2001; Yunger et al.,
2004) and ultimately with their academic self-efficacy. However, these connections
should be investigated in future research.

Even though we feel this research has added important insights to the study of
the educational gender gap, this investigation also had some limitations. For
instance, the children in this study are about twelve year olds and still at the
beginning of adolescence. They still have a long way to go in exploring and
constructing their identity (Klimstra, Hale Iii, Raaijmakers, Branje, & Meeus, 2009),
including their gender identity. Therefore, certain links we have found in this study
might still be subject to change. Hence, we would like to call for longitudinal
research, in order to test our results in different life phases. Furthermore, this study
is based on cross-sectional survey data. Consequently, even though we assume a
pathway from gender identity to self-efficacy, no causality claims can be made.
While it seems unlikely that self-efficacy impacts on gender identity, whose
foundations are theorized to be established in early childhood (Spence, 1984), it
could be that self-efficacy and gender identity are co-constructed. This possibility
and even reversed pathways should be subject to further longitudinal or
experimental research.

CONCLUSION

This research into the educational gender gap is exceptional because of the
consideration of intrasexual differences by employing the concept of gender
identity. The results show that self-perceived typical girls score the highest on
academic self-efficacy, followed by typical boys. This indicates that individually-
assessed masculinity and femininity help to understand the workings of the gender
gap in education. Apparently, self-perceived femininity in girls is more positively
connected to academic self-efficacy than masculinity in boys. This might be
because of the traits, attitudes and behavior which are typically connected to
femininity and which might be more conducive for educational achievement.
However, students scoring low on gender typicality, both girls and boys, score
similarly low on academic self-efficacy. A large part of this association is explained
by the lower well-being of these students. This shows that feeling gender atypical in our societies is a challenge for youngsters, which negatively connects to their well-being and ultimately to educational parameters. This research also demonstrates that by considering intrasexual differences, low-achieving girls and high-achieving boys can be rendered visible.
Chapter 8.
Differences in study motivation within and between genders:
An examination by gender typicality among early adolescents.

Wendelien Vantieghem, & Mieke Van Houtte
Published online in Youth & Society.
Doi: 10.1177/0044118X15602268

Despite boys’ educational underachievement, gender differences in study motivation have received little research attention. Guided by self-determination theory and the identity-based motivation model, this study investigates differences in study motivation between boys and girls, as well as within each gender. To adequately consider these within-gender differences, we investigate gender and gender typicality interactions in a sample of 6380 Flemish 7th graders collected in 2012-2013. Results from multilevel analyses show that, in line with the educational gender gap, girls display higher levels of autonomous motivation. Furthermore, gender-typical girls score highest on autonomous motivation. Gender-typical boys score considerably lower, though they outperform self-perceived atypical boys and girls. In controlled motivation, no differences are observed between boys and girls of equal ability. Nevertheless, higher scores on gender typicality contribute to a higher sense of controlled motivation. The results are discussed in light of well-being, the need for autonomy, and gendered expectations of teachers.
INTRODUCTION

In recent decades, the gender gap in education has captured the minds of policymakers and researchers alike. Since the 1990s, it has become clear that girls are outperforming boys on several educational parameters: Boys repeat grades more often (Fergusson & Horwood, 1997; Van Landeghem et al., 2010), have lower marks in class (Duckworth & Seligman, 2006; Epstein, 1998; Fergusson & Horwood, 1997), drop out more often (Buchmann et al., 2008; Fergusson & Horwood, 1997; Van Landeghem et al., 2010), represent lower enrollment in higher education (Buchmann et al., 2008) and are overrepresented in special education services and remedial classes (Benjamin, 2003) These findings are encountered in most western industrialized countries (PISA, 2009) and constitute a pervasive challenge.

In line with the gender-intensification hypothesis, early adolescence seems to be an important phase for the manifestation and consolidation of these gender-differentiated patterns (Galambos, Almeida, & Petersen, 1990). That is, gender plays a central role in several developmental tasks of early adolescence, such as the consolidation of gender identity (Carver et al., 2003; Tobin et al., 2010) physical maturation, management of sexual interests and the prospect of adult roles, such as planning an academic and occupational future (Perry & Pauletti, 2011; Tobin et al., 2010). Notably, these tasks concerning gender-role development coincide with a substantial widening of the educational gender gap. Differences between boys and girls tend to be rather minimal in elementary education (Buchmann et al., 2008; Derks & Vermeersch, 2001; Fergusson & Horwood, 1997), while the early years of secondary education mark the expansion of gendered educational patterns. This suggests that early adolescence is an important phase for the manifestation and consolidation of gender-differentiated achievement patterns. Hence, attention to the gender-role developmental challenges of adolescence could increase our understanding of the concurrent enlargement of boy-girl differences in education.

Furthermore, in order to remedy the underachievement of boys, it is important to understand how achievement is determined. Study motivation is one of the central determinants of educational achievement (Hattie, 2008) and within the research of study motivation, self-determination theory has become the dominant theoretical framework. Self-determination theory (SDT), which states that self-determined behavior is characterized by motivation of better quality, considers the antecedents, variations and consequences of several types of study motivation (Deci, Vallerand, Pelletier, & Ryan, 1991; Ryan & Connell, 1989). For instance, autonomous study motivation refers to behavior characterized by
enjoyment and/or personal relevance (Deci et al. 1991; Vansteenkiste et al. 2009) and has shown to be one of the central determinants of academic achievement (Hattie, 2008). Conversely, controlled study motivation refers to behavior prompted by external or internal pressure (Deci et al. 1991) and connects negatively to academic achievement (Vansteenkiste, Sierens, Soenens, Luyckx, & Lens, 2009).

A model which might theoretically link gendered issues with study motivation is the identity-based motivation model (IBM) (Elmore & Oyserman, 2012). The model proposes that the traits associated with a salient social identity guide people’s behavioral options. For example, to the extent that a feminine gender identity is stereotypically linked to being well-behaved, dutiful and cooperative (Beaman et al., 2006; Jones & Myhill, 2004), girls’ chances of demonstrating these behaviors at school are raised (Elmore & Oyserman, 2012). Conversely, because masculine gender identities are often linked to being an academic underachiever, class clown, or effortless achiever (Francis, 2000; Mac an Ghaill, 1994), boys’ chances of academic engagement are reduced (Elmore & Oyserman, 2012).

The theoretical frameworks of self-determination theory (SDT) and the identity-based motivation model (IBM) complement each other nicely. For one, where IBM lacks a clear focus by considering a wide range of educational parameters, SDT concentrates on the key educational parameter of study motivation. Furthermore, where SDT tends to be concerned mostly with individual differences (which has contributed to its lack of gender-driven research – as will be discussed below), IBM draws attention to the way social groups shape individuals’ experiences. Thus, linking IBM and SDT provides a coherent framework for examining gender differences in study motivation.

So far, research has revealed that gender differences in study motivation are in line with IBM and the gender gap in education, with secondary-school girls displaying higher average levels of autonomous motivation than boys (Marsh & Myers, 1986; Vallerand, Fortier, & Guay, 1997; Vansteenkiste et al., 2009). However, no study that we are aware of has taken these gender differences as its focus. Thus, gender differences in autonomous versus controlled motivation remain little understood. Moreover, within-gender variations have received little attention in SDT- and IBM-driven research. Despite the general tendency for girls to outperform boys, there are quite some high-performing boys and low-achieving girls (Warrington et al., 2000). These students tend to remain invisible due to most research’s focus on general girl-boy differences.

We would like to suggest self-perceived gender typicality as a social-identity concept that ties into the developmental tasks of early adolescence and could help
to account for these within-gender variations. Gender typicality refers to the extent to which someone feels typical for his or her gender category and is a central aspect of gender identity (Egan & Perry, 2001). In this paper, we investigate the way gender and self-perceived gender typicality link to study motivation in a sample of early adolescents. Additionally, we consider the interaction between gender and gender typicality to investigate differences in study motivation within genders. In doing so, we hope to better understand the underlying gendered processes which determine whether girls and boys achieve educational success.

**SELF-DETERMINATION THEORY AND STUDY MOTIVATION**

Self-determination theory has been developed by Deci and Ryan (Deci & Ryan, 1985). Central to the theory is the distinction between autonomous and controlled motivation. Controlled motivation reflects an external perceived locus of causality, whereas autonomous motivation reflects an internal locus of control (Vansteenkiste, Lens, De Witte, De Witte, & Deci, 2004). Hence, autonomously motivated people engage freely and willingly in behavior because they feel these actions are a reflection of who they are. Autonomous study motivation has two bases: identified and intrinsic motivation (Deci et al., 1991; Vansteenkiste et al., 2009). Identified regulation is characterized by students’ understanding of the personal relevance and importance of the learning activity. Intrinsic motivation represents students engaging in an activity because of interest in the task itself or because of the enjoyment they derive from it. Hence, autonomously motivated students study because they enjoy it or because they feel studying is important to them.

Conversely, controlled motivated people engage in an activity because they feel they have no choice, either by external or internal force. Sources of external force may be rewards or punishments offered by others, whereas internal force refers to people pressuring themselves into action by feelings of guilt, shame, or pride. Motivation based on external forces is commonly referred to as external regulation, whereas motivation spurred by internal force is called introjected regulation. These types make up the two bases of controlled study motivation (Deci et al., 1991; Vansteenkiste et al., 2009).

In general, self-determination theory proposes that motivation is of better quality when behavior is self-determined and volitional. It also postulates that high-quality study motivation is essential for educational success (Deci et al., 1991; Ryan & Connell, 1989; Ryan & Deci, 2013). As such, autonomous motivation, which is characterized by volitional study behavior, is expected to connect to better
educational outcomes than controlled motivation. Indeed, controlled study motivation has been shown to result in lower achievement because students who experience controlled motivation tend to employ less adaptive learning strategies. For instance, they use more surface-level learning (Meece et al., 2006; Walker et al., 2006), such as rote memorization and basic rehearsal (Walker et al., 2006). They also display more cheating, maladaptive coping strategies (Ryan & Connell, 1989), less persistence (Vansteenkiste et al., 2009), and more test anxiety (Ryan & Connell, 1989; Vansteenkiste et al., 2009).

Conversely, autonomous study motivation is characterized by the use of more meta-cognitive strategies (Pintrich & De Groot, 1990; Vansteenkiste et al., 2009), such as planning and time-management. Autonomously motivated students persist longer at tasks (Meece et al., 2006), show more effort (Pintrich & De Groot, 1990; Ryan & Connell, 1989; Vansteenkiste et al., 2009), more deep-level learning, better retention and less procrastination (Deci et al., 1991). This translates into higher achievement (Deci et al., 1991; Pintrich & De Groot, 1990), which is demonstrated by higher marks in class and lower drop-out among autonomously motivated students (Vallerand et al., 1997).

GENDER DIFFERENCES IN STUDY MOTIVATION

On average, boys tend to score lower on autonomous motivation than girls (Marsh, Martin, & Cheng, 2008; Vallerand et al., 1997; Vansteenkiste et al., 2009). These motivational differences might help to explain why boys consistently display less conducive study behavior, such as being more inattentive and disruptive in class (Fergusson & Horwood, 1997; Francis, 2000; Warrington et al., 2000) and having less self-discipline where homework is concerned (Duckworth & Seligman, 2006). These study behaviors ultimately translate into boys having less academic success than girls (Duckworth & Seligman, 2006; Epstein, 1998; Fergusson & Horwood, 1997).

Despite this issue of boys’ reduced educational engagement, none of the studies that reported gender differences in autonomous motivation had gender as its research focus (e.g., Vansteenkiste et al. [2009] on motivational profiles; Marsh, Martin, & Cheng [2008] on the potential benefits of male teachers; and Vallerand et al. [1997] on high-school drop-out). The reported gender differences are usually secondary findings in a larger research project. Conversely, some studies using the identity-based motivation model (IBM) have focused on the influence of gender. For instance, Elmore and Oyserman (2012) observed how children develop more academically-oriented goals and how boys persist longer at school tasks by
connecting their gender category to success. Nevertheless, study motivation has not been considered in IBM-driven research regarding gender identity. Consequently, both in IBM and SDT-driven research, gender-differentiated study motivation has received little research attention. Therefore, it remains unclear whether significant gender differences in controlled study motivation exist. It could be hypothesized that boys score higher on controlled study motivation than girls. That is, based on ethnographic accounts of how traditional notions of masculinity infuse boys’ peer groups with anti-school attitudes (e.g., Francis, 2000; Swain, 2005), one could expect boys to study more because they have to than because they want to, resulting in higher levels of controlled motivation.

**WITHIN-GENDER DIFFERENCES AND GENDER TYPICALITY**

While differences between girls and boys received little attention in SDT-driven research, differences within the genders have not been researched at all. Even in IBM-research, little attention has been devoted to variations within the genders. That is, IBM tends to be mostly used in experimental research where certain identities are primed, or in interventions that adapt identity contents (e.g., Elmore & Oyserman, 2012; Oyserman et al., 2006; Oyserman & Destin, 2010). Because of this, the research employing IBM has remained mostly group-based, whereas individual differences regarding the content or salience of identities have rarely been investigated.

Interestingly, the identity-based motivation model (IBM) does provide the possibility of studying within-gender variations via the postulate of dynamic construction of identity content (Oyserman & Destin, 2010). This postulate states that the influence of identities depends on which traits a person associates with the social group. As such, individual variation in identity content is explicitly part of the model, yet it has been rarely investigated where gender identity is concerned (for an application of within-group differences in ethnic identity, see Oyserman et al., [2003]). We propose that this gap could be remedied by considering individual variations in gender identity through survey research concerning gender typicality.

Gender typicality refers to the extent to which people feel to be good examples of their gender category (Egan & Perry, 2001). In other words, it reflects individual variations in the extent to which people feel they fulfill typical gender notions. In line with the expectations of the identity-based motivation model (IBM), gender typicality is an important motivator for gendered behavior (Tobin et al., 2010; West & Zimmerman, 1987). That is, it has been shown that individuals who score higher on gender typicality also display more gender-congruent behavior.
Moreover, one of the important developmental tasks of middle childhood and early adolescence is consolidating one’s gender typicality through processes of social comparison (Carver et al., 2003; Tobin et al., 2010). Feeling gender atypical might be especially challenging for early adolescents, because gender norms tend to become more limiting in early adolescence (Galambos et al., 1990; Horn, 2007) and peer-acceptance is often conditional on displaying gender-conforming behavior (Blakemore, 2003; Stoudt, 2006). Indeed, gender atypical adolescents often score lower on well-being (Carver et al., 2003; Egan & Perry, 2001; Yung et al., 2004), and this has been shown to negatively influence their educational functioning (Vantieghem, Vermeersch, & Van Houtte, 2014a). Hence, because of its connection to development in early adolescence, IBM, and educational functioning, gender typicality constitutes an interesting variable for the study the educational gender gap and study motivation in early adolescence.

So far, gender typicality has only rarely been employed within educational research. Leaper and colleagues showed that gender typicality was associated with college men’s ability beliefs, interests and choice for gender-typical courses (Leaper & Van, 2008). Contrary to expectations, they found that it was not associated with secondary school girls’ value and ability beliefs for science, mathematics or English (Leaper et al., 2012). Similarly, Patterson and Pahlke (2011) found no effect of gender typicality on the academic achievement of ten- to thirteen-year-old girls in a single-gender school. Based on this work, we could expect that there would be a different relation between gender typicality and educational parameters for boys than girls. However, both genders were not simultaneously included in these samples, which is necessary to make claims about gender-specific relations. Furthermore, there was a sizeable age difference between the boys and girls in these studies (school girls versus college boys), which could further complicate comparing the results. Thus, while this earlier research suggests a stronger relationship between gender typicality and educational parameters for boys than girls, it remains a tentative hypothesis that should be confirmed in other research.

THE CURRENT STUDY

The current study investigates differences in study motivation between and within genders by employing the concept of gender typicality in a sample of early adolescents. Given the higher academic performance of girls and previous research concerning study motivation (Marsh et al., 2008; Vallerand et al., 1997; Vansteenkiste et al., 2009), we expect girls to display a more advantageous
motivational profile than boys, with higher scores on autonomous motivation and lower scores on controlled motivation (hypothesis 1).

Despite the link IBM posits between gender identity and educational functioning, research that connects the concept of gender typicality with educational outcomes has been limited. The research that has considered gender typicality in an educational setting has focused on ability beliefs (e.g., Leaper et al., 2012; Leaper & Van, 2008) or achievement (e.g., Patterson & Pahlke, 2011), whereas study motivation has not been taken into account. Furthermore, since this above-cited research has used single-gender samples, it remains unclear if an overarching effect of gender typicality on educational outcomes for both genders exists. However, given the research that demonstrates the reduced well-being of atypical adolescents (Carver et al., 2003; Egan & Perry, 2001; Yunger et al., 2004) and knowing that negative emotional states connect to less adaptive motivational profiles (Assor, Kaplan, Kanat-Maymon, & Roth, 2005), we could expect self-perceived atypical adolescents to have a less advantageous motivational profile than gender-typical adolescents (hypothesis 2).

Lastly, we consider within-gender differences in motivation by investigating the interaction between gender and self-perceived gender typicality. The identity-based motivation model (IBM) suggests that the traits associated with a social identity determine the courses of action that come to mind (Elmore & Oyserman, 2012). Given stereotypically feminine behaviors tend to be more compatible with students’ requirements than traditionally masculine behavior, IBM predicts that girls should display higher educational engagement than boys. However, IBM-driven research has not yet considered individual variations in the identification with such norms. Studies employing the concept of gender typicality, which assess respondents’ perceived fit with their gender category, suggest that there is an association between young men’s felt typicality and ability beliefs in gender-congruent courses (Leaper & Van, 2008). Contrary to expectations, studies found no association between gender typicality and adolescent girls’ ability beliefs or achievement (Leaper et al., 2012; Patterson & Pahlke, 2011). However, because of sample-issues with these studies (as discussed above), it remains to be seen whether the current study with early adolescents can confirm a stronger association between gender typicality and study motivation for boys than girls (hypothesis 3).
METHOD

Participants and Procedure

The data are part of the “Teaching in the Bed of Procrustes”-study and were gathered in the first half of the 2012-13 school year. Schools were selected so each geographical region within Flanders (Belgium) was equally represented. Selection was further dependent on school denomination and a proportional representation of rural versus city schools. Within these parameters, three random samples were drawn. For each school that refused participation, a matched school from the next random sample was contacted. This way, we aimed to obtain schools that represent the Flemish educational context on important parameters (such as region, school denomination and location), while maintaining randomness within these subpopulations. In the end, of the 124 contacted schools 59 schools participated, which translates to a response rate of 47.6%. This response rate is less than 50% due to schools in Flanders being swamped with research requests. Consequently, schools accept research requests on a first-come, first-serve basis. Analyses in which we compared our sample to the Flemish school population, based on information from the Flemish Educational Department, showed no important differences in school sector, curriculum or student composition. This suggests that no systematic biases occurred and that the schools were representative of the Flemish context.

In each school, all 7th graders were asked to complete the questionnaire. A waiver of parental consent and the use of child assent were approved by the school and the Belgian Commission for the Protection of Privacy, based on the minimal risk of the study. Researchers were always present during the completion of the survey to explain the procedure and answer questions. The students had 50 minutes to complete the paper and pencil survey. Pupils were told that the survey was not a test and were assured that it was completely confidential. The response rate among the students was 96.9%, translating to a total of 6380 students.

Of the students in the sample, 53.8% were boys and the majority were of Belgian or western-European descent (79.9%). In line with the migration history of Belgium, children of non-western European descent (20.1%) had predominantly Turkish, Moroccan or Eastern-European roots. The socio-economic status (SES) of the students was based on the work situation of the parents. In this sample, 22% of the students had a blue collar background (e.g., parents in a manual labor position), while 78% had a white collar background (e.g., parents in service or management positions). In the Flemish educational context, there are officially two tracks in the first grade: the general and the vocational-oriented track (Boone & Van Houtte, 2012; Van Houtte et al., 2012). In our sample, 11.7% of pupils were in
the vocational track. Because the sample consisted of 7th graders, the mean age was 12 years old. However, there was a range from 10 to 15 years old, indicating that some pupils skipped a year while others had repeated years. Nevertheless, 76.6% of the sample was twelve years old, demonstrating that the majority were on track.

**Measures**

**Study motivation.** To assess study motivation, we used the Academic Self-regulation Scale (Ryan & Connell, 1989). This scale consists of four subscales, each containing four five-point likert items (0 = completely disagree, 4 = completely agree), measuring external, introjected, identified and intrinsic study motivation (Deci et al., 1991). The scale presents a stem followed by rationales, for instance “I am motivated to study because…”: “others force me to do it” (external), “I would feel ashamed if I didn’t” (introjected), “this is an important goal in my life” (identified), “studying is fun” (intrinsic).

Similar to other researchers (e.g., Black & Deci, 2000; Haerens, Kirk, Cardon, De Bourdeaudhuij, & Vansteenkiste, 2010; Vansteenkiste et al., 2004), we created two composite scores. First, autonomous motivation was created by averaging the subscales of intrinsic and identified regulation. Second, controlled motivation was created by averaging the subscales of introjected and external regulation. This approach was justified, as a principal-components analysis indicated a clear drop in eigenvalues (i.e., 5.67, 2.40, 1.28, 1.03) between the second and third retained factor. Together, the first two components explained 50.42% of the variance in the motivation items. After orthogonal rotation (VARIMAX), all autonomous motivation items had loadings of at least .50 on the first component, whereas all controlled motivation items had loadings of at least .40 on the second factor. No cross-loadings were found. Cronbach’s alpha of autonomous motivation was .89, and .79 for controlled motivation.

**Gender typicality.** To measure gender typicality, we employed the gender typicality subscale from the Self-Concept Questionnaire by Egan and Perry (2001). This measure consists of six items with a five-point likert scale (0 = completely disagree, 4 = completely agree) tapping the extent to which one feels typical of one’s gender (e.g., “I feel that my skills and interests are similar to those of other [girls/boys]”, “I feel that I am a good example of a typical [boy/girl]”). The gender typicality score was computed by using the mean sum of scores. It had a Cronbach’s alpha of .77.
Subjective well-being. Subjective well-being was measured by asking students how often they had felt certain emotions in the past 30 days. They indicated the frequency of positive and negative feelings on a 5-point scale ranging from 0 = never to 4 = very often (Keyes et al., 2002). The scale consisted of 12 items, equally divided between positive (e.g., happy, peaceful) and negative items (e.g., nervous, worthless). Based on principal component analysis, we constructed a single scale from these items. The items pertaining to negative affect were reverse coded so higher scores on the scale indicate a higher well-being. This scale was computed using a mean sum of scores, and it had a Cronbach’s alpha of .81.

Initial ability. Initial ability was measured by two validated tests for students of this age in mathematics and in Dutch (the students’ first language). The mathematics test consisted of 50 questions regarding problematic cases and calculations (Dudal, 2003). The Dutch test consisted of four texts and questions concerning content and vocabulary (CITO, 2010). The scales were constructed separately for Dutch and mathematics by a sum score of the correct answers on each test.

Data Analytic Plan
We began with descriptive analyses that gave us an insight into overall gender comparisons in the data. We then used regression analyses to analyze the research questions while taking the confounding effects of the control variables into account. Because the sample consisted of 7th grade students clustered in schools, we employed multilevel analysis to account for the nested structure of the data, using the statistical program MLwiN.

We ran separate multilevel regression analyses for the outcome variables autonomous versus controlled motivation. For each outcome, we first ran a zero-model to calculate the variance at the school versus the individual level (see Results). In the next model, we included the main variables, gender, self-perceived gender typicality and its interaction, to assess the hypotheses. All continuous variables were grand mean centered to improve interpretability of the intercept (Hox, 2010). In the final model, we included the control variables to check whether the associations remained, diminished or disappeared after controlling for socioeconomic position, initial scholastic ability, subjective well-being, school track, age and ethnicity.

Variance components were random for the intercept and individual error term. Variance components for other variables were kept fixed to improve statistical power and because there was no significant random slope variance in
preliminary models, indicating that the effects did not fluctuate in different schools (Hox, 2010).

Table 10.
Descriptive statistics.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total sample</th>
<th>Girls</th>
<th>Boys</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Autonomous motivation</td>
<td>2.03</td>
<td>.82</td>
<td>5865</td>
</tr>
<tr>
<td>Controlled motivation</td>
<td>2.07</td>
<td>.71</td>
<td>5871</td>
</tr>
<tr>
<td>Sex (0 = boys)</td>
<td>53.8</td>
<td></td>
<td>6380</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>2.47</td>
<td>.68</td>
<td>6284</td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being</td>
<td>2.83</td>
<td>.52</td>
<td>6314</td>
</tr>
<tr>
<td>Math</td>
<td>26.79</td>
<td>10.77</td>
<td>5982</td>
</tr>
<tr>
<td>Dutch</td>
<td>14.41</td>
<td>5.37</td>
<td>5749</td>
</tr>
<tr>
<td>Age</td>
<td>12.23</td>
<td>.51</td>
<td>6367</td>
</tr>
<tr>
<td>Track (0 = general)</td>
<td>88.3</td>
<td>6380</td>
<td>90.7</td>
</tr>
<tr>
<td>Ethnicity (0 = ethnic</td>
<td>79.9</td>
<td>6170</td>
<td>79.8</td>
</tr>
<tr>
<td>major)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES (0 = white-collar)</td>
<td>78</td>
<td>6194</td>
<td>79.8</td>
</tr>
</tbody>
</table>

**RESULTS**

**Descriptive Statistics**
Means and standard deviations are presented by gender for all variables in Table 10. Compared to boys and in line with hypothesis 1, girls had a somewhat higher score on autonomous motivation, $t(5863) = 5.28, p < .001, d = .14$, and a negligibly lower sense of controlled motivation, $t(5869) = 2.55, p = .011, d = .06$. Girls also scored somewhat lower on gender typicality than boys, $t(6282) = 7.45, p < .001, d = .19$. In line with the educational gender gap and previous research, girls were better represented in the academic-oriented track than boys, $\chi^2(1, N = 6380) = 31.46, p < .001, V = .07$, were slightly younger, $t(6365) = 5.43, p < .001, d = .14$ (indicating
that they had repeated fewer grades than boys), and scored somewhat higher on the language test, $t(5747) = 12.23, p < .001, d = .33$. Conversely, boys did a little better than girls on the mathematics test, $t(5980) = 5.06, p < .001, d = .13$. There were no differences between boys and girls on well-being, $t(6312) = 1.77, p = .076, d = .04$, and ethnicity status, $\chi^2(1, N = 6170) = .05, p = .831, V = .003, ns$. However, girls had a white-collar background negligibly more often than boys, $\chi^2(1, N = 6194) = 10.05, p = .002, V = .04$.

The zero-order correlations in Table 11 show that the associations between the variables did not differ much between boys and girls, with two exceptions. First, in contrast to hypothesis 3, Table 11 shows that the positive association between gender typicality and autonomous motivation was about three times as large for girls as it was for boys, $Z = 2.82, p < .01$. Second, the association between autonomous and controlled motivation was substantially stronger among boys than among girls, $Z = 7.84, p < .001$, indicating that both types of motivation tend to be more concurrent for boys than for girls.

### Table 11.
Zero-order correlations divided by gender.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Autonomous motivation</th>
<th>Controlled motivation</th>
<th>Gender typicality</th>
<th>Well-being</th>
<th>Math</th>
<th>Dutch</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autonomous motivation</td>
<td>—</td>
<td>.472***</td>
<td>.036*</td>
<td>.174***</td>
<td>−.054**</td>
<td>−.092***</td>
<td>.114***</td>
</tr>
<tr>
<td>Controlled motivation</td>
<td>.298***</td>
<td>—</td>
<td>.059**</td>
<td>−.031</td>
<td>−.016</td>
<td>−.034</td>
<td>.021</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.110***</td>
<td>.068***</td>
<td>—</td>
<td>.222***</td>
<td>.085***</td>
<td>.051**</td>
<td>−.048**</td>
</tr>
</tbody>
</table>

**Control Variables**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Autonomous motivation</th>
<th>Controlled motivation</th>
<th>Gender typicality</th>
<th>Well-being</th>
<th>Math</th>
<th>Dutch</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-being</td>
<td>.148***</td>
<td>−.098***</td>
<td>.293***</td>
<td>—</td>
<td>.114***</td>
<td>.095***</td>
<td>−.088***</td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td>−.023</td>
<td>.010</td>
<td>.084***</td>
<td>.108***</td>
<td>—</td>
<td>.552***</td>
<td>−.341***</td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td>−.092***</td>
<td>−.032</td>
<td>.034</td>
<td>.113***</td>
<td>.582***</td>
<td>—</td>
<td>−.315***</td>
</tr>
<tr>
<td>Age</td>
<td>.084***</td>
<td>.028</td>
<td>−.045*</td>
<td>−.087***</td>
<td>−.363***</td>
<td>−.334***</td>
<td>—</td>
</tr>
</tbody>
</table>

*Note. Correlations for boys are above the diagonal; correlations for girls are below the diagonal.  
*p < .05. **p < .01. ***p < .001*
Autonomous Motivation
We conducted a multilevel regression analysis on autonomous motivation. The zero-model showed that autonomous motivation varied somewhat between schools, however only 5.19% ($e_{0ij} = .639, u_{0j} = .035, p < .001$) of total variance was situated at the school-level.

As can be seen in Table 12, model 1 showed a significant association between gender and autonomous motivation ($\gamma = .125, p < .001$), indicating that girls scored somewhat higher than boys on autonomous motivation, confirming hypothesis 1. Gender typicality was also positively related to autonomous motivation ($\gamma = .053, p < .05$). This means that gender typical students scored higher on autonomous motivation, which is in line with hypothesis 2. The significant Gender x Gender Typicality interaction ($\gamma = .080, p < .01$) showed that the impact of typicality significantly differed according to students’ gender (see Figure 7). Gender atypical students (both boys and girls) scored similarly low.

However, a sizeable difference in autonomous motivation between typical boys and girls appeared, with typical girls outperforming boys. This indicates that the effect of gender typicality is larger for girls, which runs against the expectations based on previous empirical research concerning gender typicality and educational parameters (see hypothesis 3).

Introducing the control variables in the last model dispelled the significant effect of gender typicality and the interaction effect between gender and self-perceived gender typicality. Analysis showed this to be mostly due to the introduction of subjective well-being ($\gamma = .270, p < .001$). Overall, the final model explained 12.01% of the total variance in autonomous motivation.
Figure 7.
Association between sex, gender typicality, and autonomous study motivation.

![Graph showing the association between sex, gender typicality, and autonomous study motivation.]

Table 12.
Results of multilevel analysis on autonomous motivation.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.987*** (.030)</td>
<td>1.862*** (.025)</td>
</tr>
<tr>
<td>Sex (0 = boy)</td>
<td>.125*** (.022)</td>
<td>.157*** (.024)</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.053* (.022)</td>
<td>.024 (.023)</td>
</tr>
<tr>
<td>Sex X Gender typicality</td>
<td>.08** (.031)</td>
<td>.044 (.033)</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being</td>
<td></td>
<td>.270*** (.022)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>.053* (.026)</td>
</tr>
<tr>
<td>Track (0 = general)</td>
<td></td>
<td>.027 (.052)</td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td></td>
<td>-.012*** (.003)</td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td></td>
<td>.003** (.001)</td>
</tr>
<tr>
<td>SES (0 = white-collar)</td>
<td></td>
<td>.035 (.031)</td>
</tr>
<tr>
<td>Ethnicity (0 = ethnic majority)</td>
<td></td>
<td>.340*** (.034)</td>
</tr>
<tr>
<td><strong>Variance components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School $\mu_{ij}$</td>
<td>.036 (.008)</td>
<td>.014 (.004)</td>
</tr>
<tr>
<td>Student $\theta_{ij}$</td>
<td>.632 (.012)</td>
<td>.579 (.012)</td>
</tr>
</tbody>
</table>

*Note. Unstandardized coefficients and standard errors are provided in parentheses. Tests represented here were two-sided. $N_{\text{School}}$: 57, $N_{\text{Student}}$: 4,900.

*p < .05, **p < .01, ***p < .001.
Controlled Motivation

We conducted a multilevel regression analysis on controlled motivation. The zero-model showed that controlled motivation varied minimally between schools, with only 0.99% ($e_{0ij} = .499, u_{0j} = .005, p < .001$) of total variance being situated at the school-level.

Model 3 showed that gender ($\gamma = -.038, p < .05$) was negatively related to controlled motivation, whereas gender typicality ($\gamma = .061, p < .01$) was positively related to controlled motivation (Table 13). This indicates that girls scored slightly lower on controlled motivation, whereas gender typical students scored somewhat higher. Against expectations, the Gender x Gender typicality interaction was not significant ($\gamma = .011, p > .10$), indicating that the impact of typicality did not differ according to students’ gender.

Introducing the control variables in the last model dispelled the significant effect of gender ($\gamma = -.026, p > .10$), whereas the significant effect of gender typicality remained relatively unchanged ($\gamma = .089; p < .001$). Analysis showed that the disappearance of the gender-effect was mostly due to the introduction of the initial ability measures. Overall, the final model explains 5.54% of the total variance in controlled motivation.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.089*** (.016)</td>
<td>2.069*** (.019)</td>
</tr>
<tr>
<td>Sex (0 = boy)</td>
<td>$-.038^* (.019)$</td>
<td>$-.026 (.021)$</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>$.061** (.019)$</td>
<td>$.089*** (.021)$</td>
</tr>
<tr>
<td>Sex x Gender typicality</td>
<td>$.011 (.027)$</td>
<td>$.011 (.030)$</td>
</tr>
</tbody>
</table>

Control Variables

<table>
<thead>
<tr>
<th>Control Variables</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial ability: Dutch</td>
<td>$-.004^* (.002)$</td>
<td></td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td>$.002* (.001)$</td>
<td></td>
</tr>
<tr>
<td>Well-being</td>
<td>$-.119*** (.02)$</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>$.006 (.024)$</td>
<td></td>
</tr>
<tr>
<td>Track (0 = general)</td>
<td>$-.096^* (.046)$</td>
<td></td>
</tr>
<tr>
<td>SES (0 = white-collar)</td>
<td>$-.019* (.028)$</td>
<td></td>
</tr>
<tr>
<td>Ethnicity (0 = ethnic majority)</td>
<td>$.132*** (.03)$</td>
<td></td>
</tr>
</tbody>
</table>

Variance components

<table>
<thead>
<tr>
<th>Variance components</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>School $\mu_{0j}$</td>
<td>$.005 (.002)$</td>
<td>$.004 (.002)$</td>
</tr>
<tr>
<td>Student $\epsilon_{0ij}$</td>
<td>$.494 (.009)$</td>
<td>$.472 (.01)$</td>
</tr>
</tbody>
</table>

Note. Unstandardized coefficients and standard errors are provided in parentheses. Tests represented here were two-sided. $N_{School} = 57$, $N_{Student} = 4,907$.

*p < .05. **p < .01. ***p < .001
DISCUSSION

In recent decades, research has documented that girls outperform boys on several educational parameters in most western industrialized countries. The identity-based motivation model (IBM) offers an explanation for the observed gender-differential achievement patterns by pointing towards the traits typically associated with femininity versus masculinity (Elmore & Oyserman, 2012). However, research employing this model has not considered individual variations in self-perceived adherence to these typical gender traits. This study’s goal was to add to this research by accounting for individual variations, as well as overall differences between the genders. Furthermore, we argue that gender-related issues connected to early adolescence might be of special importance to the development and consolidation of the educational gender gap. To adequately consider within-gender differences in early adolescence from a social-identity perspective, we introduced the concept of gender typicality (Egan & Perry, 2001). We link this concept to the self-determination theory (SDT) of Deci and Ryan (1985), because their conceptualization of study motivation has been of central importance to the study of academic achievement throughout the school career. This way, this research also contributes to self-determination theory, given gender differences in study motivation have rarely been considered from this perspective.

The analyses show that there are in fact important differences in study motivation within and between genders among Flemish 7th graders. In accordance with the first hypothesis and as suggested by previous research (Marsh et al., 2008; Vallerand et al., 1997; Vansteenkiste et al., 2009), girls possessed a more advantageous motivation profile than boys. On average, girls had higher scores than boys on autonomous motivation and slightly lower scores on controlled motivation. In the end, the difference in controlled motivation disappeared when controlling for students’ initial ability scores. Thus, although boys and girls of similar ability demonstrated an equal level of controlled motivation, girls continued to have a more favorable level of autonomous motivation. These findings are in line with the educational gender gap, with girls displaying higher levels of academic engagement and achievement than boys (Epstein et al., 1998; Francis, 2000; Warrington et al., 2000).

In addition to considering overall gender differences in study motivation, this study also examined within-gender differences. Although important differences within each gender category were observed for autonomous motivation (as will be discussed below), the interaction effect between self-perceived gender typicality and gender never reached significance for controlled motivation. It seems plausible that
the limited differences between boys and girls in controlled motivation left little room for the detection of within-gender differences.

Nonetheless, the main association between gender typicality and controlled motivation was significant throughout the analyses. This indicates that gender typical students scored somewhat higher on controlled motivation than atypical students. Hence, typical students tended to feel more pressure to study – both from others and from themselves through feelings of guilt and shame. It could be that these students are just more prone to internalize external pressures, both regarding gender norms (which leads to their higher sense of gender typicality) as well as pressures concerning studying (leading to a higher sense of controlled study motivation). Nevertheless, future research should investigate how expectations from teachers link with these processes. On the one hand, the identity-based motivation model draws attention to the compatibility between the qualities of good students and girls who are stereotypically feminine, such as being calm, cooperative, compliant, and tidy (Beaman et al., 2006; Jones & Myhill, 2004). Although this might raise typical girls’ educational functioning (indeed, typical girls scored highest on autonomous motivation in this study, as will be discussed below), it might also raise teachers’ expectations of these girls. In fact, studies have shown that girls are usually perceived to be ideal students (Beaman et al., 2006; Jones & Myhill, 2004), and this might be even more true of girls who adhere to stereotypically feminine behavioral norms. Consequently, gender-typical girls might experience more pressure from teachers to live up to their expectations, resulting in a higher sense of controlled motivation.

On the other hand, as suggested by IBM, gender-typical boys might display more behavior associated with traditionally masculine anti-school cultures, such as challenging the authority of teachers and playing the class-clown (Francis, 2000; Jackson, 2003; Swain, 2005). Such behavior tends to evoke more controlling teaching methods from teachers (Reeve, 2009), which has been shown to result in more controlled motivation among students (Assor et al., 2005). This vicious circle of disruptive behavior and controlling teaching methods might lead to the higher sense of controlled motivation observed among typical boys in this study. Although we see a similar association between self-perceived gender typicality and controlled motivation for both girls and boys, we do presume that different processes might be at the basis of these outcomes (e.g., high expectations for typical girls versus controlling teaching methods for typical boys). This, however, should be subject of future research that investigates the impact of teacher behaviors and expectations on study motivation of boys and girls.
Although we did not observe significant Gender x Gender typicality interactions for controlled motivation, we did observe within-gender differences for autonomous motivation. More specifically, typical girls scored the highest on autonomous motivation (see Figure 7). Typical boys scored notably lower than typical girls. However, gender-typical students still substantially outperformed atypical boys and girls, who scored equally low on autonomous motivation due to lower well-being.

Note that by considering the impact of gender typicality within each gender, these low-motivated girls and rather well-motivated boys have been rendered visible, something most educational gender gap research tends to overlook. In part because of this focus in research and media, some teachers have developed the generalized expectation that girls are better students than boys (Jones & Myhill, 2004; Leaper & Brown, 2014; Van Houtte, 2007). This influences the educational experiences of both boys and girls, with boys being subject to disproportionate rates of disciplinary action (Consuegra, Halimi, & Engels, 2015), while girls in need of special education services more often escape official detection (Beaman et al., 2006; Benjamin, 2003). More studies that nuance the reigning dichotomy in gender educational research are necessary to break through these stereotypes affecting teachers’ expectations. That is, teachers should be aware that the general trend of girls outperforming boys masks a wide range of variation within each gender.

When we consider this within-gender variation, the results show a notable difference between girls and boys with respect to the association between gender typicality and autonomous motivation. More specifically, the current study uncovered weaker associations among boys than girls. This is quite different from the findings of previous researchers who found an association between gender typicality and ability beliefs among young men (Leaper & Van, 2008), but not among adolescent girls (Leaper et al., 2012; Patterson & Pahlke, 2011). However, these investigations employed samples that consisted of either girls or young men, but not both genders at the same time. In addition, the age-discrepancies in these samples (school girls versus college men) further complicated inferring clear conclusions. The rejection of this hypothesis is one of the important contributions of this investigation, given this is one of the first studies to compare both boys and girls on the association between gender typicality and educational variables. The results suggest that the connection between gender typicality and motivation is stronger for girls than it is for boys. As suggested before, the explanation for this might be that girls who score high on self-perceived femininity could be more motivated and at home in a school-setting than typical boys, since this fits better
within their self-concept. However, this gender-specific link between gender typicality and educational parameters should be confirmed with other samples.

Even though we found a differing impact of gender identity on autonomous motivation for typical boys and girls, we encountered quite a different picture among atypical students. Both self-perceived atypical boys and girls occupied a similarly vulnerable position. The present analyses showed that the association between gender typicality and autonomous motivation was explained away by well-being. This suggests that feeling atypical is linked to lower well-being, and this in turn connects to a lower sense of autonomous motivation. The finding that atypical children score lower on well-being has often been found in survey-research (e.g., Egan & Perry, 2001; Yuner et al., 2004), as well as by ethnographic studies which demonstrated that showing cross-gender behavior during adolescence is met with ridiculing and bullying (Epstein, 1998; Swain, 2005). The finding that atypical students demonstrate lower well-being is also consistent with self-determination theory, which suggests that the need for autonomy is a necessary prerequisite for autonomous motivation (Deci et al., 1991) and that this need is only fulfilled when people feel accepted for who they are (Legate, Ryan, & Weinstein, 2012; Uysal, Lin, & Knee, 2010). Although this process has predominantly been investigated within a lesbian, gay, and bisexual (LGB) population (Legate et al., 2012; Weinstein et al., 2012), some research suggests that experiencing gender-conformity pressures results in negative consequences for heterosexuals as well (Good & Sanchez, 2010). Hence, these findings imply that fitting into gender groups seems to be of central importance for the well-being and academic functioning of early adolescents. Future research should explore these connections between gender atypicality, the need for autonomy, well-being and academic functioning further, as well as changes in these associations in different developmental phases.

If replicated in other research, we believe a strong case could be made for urging secondary schools to pay more attention to gender issues. That is, schools are often perceived to be heteronormative environments, where gender non-conforming behavior is met with bullying and ridicule by peers and teachers (Epstein, 1998; Stoudt, 2006; Swain, 2005). Negative consequences of school victimization based on gender non-conformity or sexual orientation include poor psychological well-being, lower grade point average and impeded school belonging (Toomey, McGuire, & Russell, 2012). Nevertheless, schools can be crucial agents of change by promoting a more tolerant and gender equal environment. Research has shown that, for instance, LGBT-inclusive curricula, anti-bullying policies and gay-straight alliance clubs can significantly improve the perceived school safety for LGBT and gender non-conforming students (Toomey et al., 2012). By promoting a
more tolerant and gender equal environment, such policies could go a long way in
improving the well-being, school experiences and study motivation of both LGBT
and “atypical” students.

Next to positively impacting the school experiences of atypical students by
augmenting their well-being, extra attention should be paid to the suboptimal
motivational profile of boys. That is, previous research (Marsh et al., 2008;
Vallerand et al., 1997; Vansteenkiste et al., 2009) and the current study confirm that
boys tend to score lower on autonomous motivation, which seems especially
pertinent in light of the educational gender gap. Since autonomous motivation is
central in ensuring positive learning behaviors for both boys and girls (Pintrich &
De Groot, 1990), interventions that train teachers in autonomy-supportive
methods seem warranted (for an elaborate discussion of such an intervention, we
refer to Reeve [2009]).

Limitations
This investigation was subject to some limitations. First, the effect sizes observed in
this study are rather small, indicating a limited influence of gender typicality on
motivation. However, if gender typicality has similar influences on other
educational parameters, as has for instance been demonstrated for academic self-
efficacy (Vantieghem et al., 2014a), then these processes should not be discounted.
That is, since self-efficacy and study motivation contribute independently to
educational achievement (Pintrich & De Groot, 1990; Yusuf, 2011), these
influences of gender typicality on educational parameters could add up to a
substantial alteration in achievement across the years.

Second, the amount of variance accounted for in the models is rather
modest. Although this is an often-encountered issue in social research, it does
imply that other processes not addressed in this study are important in explaining
study motivation among Flemish 7th graders. For instance, self-determination
theory highlights the importance of the basic psychological needs of autonomy,
relatedness, and competence for developing autonomous motivation (Deci et al.,
1991). Some research has investigated how parents and teachers can heighten
students’ motivation by providing supportive environments and using specific
teaching methods (Greene et al., 2004; Wentzel, 1998). Consequently, for a full
understanding of study motivation, such aspects should be taken into account as
well.

Third, the children in this study were early adolescents and still have a long
way to go in constructing their identity (Klimstra et al., 2009), including their
gender identity. Therefore, certain links found in this study might still be subject to
change. Hence, we would like to call for longitudinal research, in order to test these results in different life phases.

Fourth, this study is based on cross-sectional survey data. Consequently, even though we assume a pathway from gender typicality to study motivation, no causality claims can be made. Although it seems unlikely that study motivation impacts on gender identity, because its foundations are theorized to be established in early childhood (Spence, 1984), it could be that study motivation and gender typicality are partly co-constructed. That is, the current cross-sectional sample cannot rule out possible confounders. Thus, the possibility remains that the association between gender typicality and motivation is an artifact of a common underlying process. The existence of confounders and the direction of the causal path should be assessed in experimental or longitudinal research, which is better suited for ruling out these issues.

Fifth, this study used self-report methods. Future research might consider peer or teacher report data, even though this might be problematic where gender typicality is concerned. What is most central for felt gender identity tends to differ from person to person (Perry & Pauletti, 2011; Tobin et al., 2010) and might result in considerable discrepancy between self-assessed and third-party-observed gender typicality.

Finally, given this is one of the first studies to delve deeper into differences between and within genders from a self-determination perspective, explanatory mediational pathways could not yet be researched. Therefore, the suggested pathways, such as beliefs regarding appropriate school behavior, frustration of the need for autonomy and teacher expectations, should become the focus of future research in order to better understand the gender gaps in study motivation.

CONCLUSION

This study investigated differences in study motivation within and between genders in a sample of Flemish 7th graders. The results suggest that femininity is linked to more beneficial educational outcomes than masculinity. That is, girls scored higher than boys on autonomous motivation and somewhat lower on controlled motivation. Additionally, girls who perceived themselves to fulfill typical gender notions scored higher on autonomous motivation than their peers.

Conversely, feeling that one does not live up to typical gender notions was equally detrimental for girls and boys. That is, self-perceived atypical students had a reduced well-being, which in turn translated to a lower sense of autonomous
motivation. These findings highlight the importance of fitting into gender groups for the well-being and educational functioning of early adolescents.

Nevertheless, feeling to fulfill traditional gender expectations is not necessarily without negative consequences, since self-perceived gender typical students scored higher on controlled motivation. A possible explanation for why these students experienced more pressure from themselves and others for studying, lie in the gendered expectations of teachers.

In conclusion, this study demonstrates that, next to boy-girl differences, differences in study motivation within each gender category exist as well. The results suggest that gender norms and the well-being that is derived from living up to these gendered expectations could be powerful in accounting for within and between gender differences in the educational gender gap.
Chapter 9.
Are girls more resilient to gender-conformity pressure?
The association between gender-conformity pressure and academic self-efficacy.

Wendelien Vantieghem, & Mieke Van Houtte
Published in Sex Roles, 73(1), 1-15.
Doi: 10.1007/s11199-015-0509-6

Since the 1990s, it has become clear that Belgian girls tend to outperform boys on educational parameters. Similar educational gender gaps are encountered in other western industrialized countries and are often attributed to cultural conventions concerning typical masculinity and femininity, which would inform gendered study cultures. This paper investigates the influence of gendered peer cultures in Flanders on an individual level by employing the concept of pressure for gender conformity. More specifically, the gender-differentiated impact of gender-conformity pressure is investigated in relation to academic self-efficacy. We examine these inter- and intragender differences in a sample of 6380 seventh-grade students in Flanders (the northern, Dutch-speaking part of Belgium), clustered in 59 schools. The data were collected at the start and the end of the 2012-13 school year. Results from multilevel regression analysis at two waves show that boys’ academic self-efficacy is lower when experiencing more pressure for gender conformity. Girls’ academic self-efficacy, however, does not decline when experiencing similar levels of pressure. In addition, when taking into account the negative toll that pressure for gender
conformity has on girls’ well-being, their academic self-efficacy is higher when experiencing more pressure. This gender-differentiated impact of pressure for gender conformity is not apparent at the start of the school year, but emerges in the course of seventh grade. The results are discussed in light of gendered expectations for boys and girls in the Belgian context.
INTRODUCTION

Since the 1990s, analyses show that Belgian girls outperform boys on several educational parameters (e.g., Derks & Vermeersch, 2002; Van Landeghem, Goos, & Van Damme, 2010). Most western industrialized countries tend to experience similar gender-differentiated achievement patterns (for an analysis in OECD countries, see PISA, 2009). Nevertheless, substantial variation in this gender gap exists depending on countries’ educational context and gender-equality levels. Given Belgium is a relatively gender-equal country, ranked 12th on the gender equality index (UNDP, 2011), it constitutes an interesting test environment for gendered processes such as the educational gender gap.

Usually, the educational gender gap is attributed to gender-specific study cultures, with girls’ peer groups displaying more conducive study attitudes and behaviors than boys'. That is, traditional notions of masculinity would infuse boys’ peer groups with an aversion to school and studying (e.g., UK: Mac an Ghaill, 1994; Australia: Martino, 1999; U.S.: Pascoe, 2007). Conversely, traditional femininity is thought to be more compatible with schools’ requirements of students’ behavior (UK, U.S. & Australia: Beaman et al., 2006; UK: Jones & Myhill, 2004). Most of the research concerning gendered study cultures has been qualitative in nature (e.g., UK: Epstein, 1998; Mac an Ghaill, 1994; Australia: Martino, 1999; U.S.: Morris, 2012; Pascoe, 2007). Relatively little research has been quantitative (e.g., Van de Gaer et al., 2006a; Van Houtte, 2004b). Most of this research, however, has focused on documenting the gender gap by considering simple boy-girl differences. In this article, we consider how gender-conformity pressures may partly account for the educational gender gap.

Pressure for gender conformity refers to the degree to which one experiences pressure from significant others and self to exhibit gender-congruent behavior (e.g., Egan & Perry, 2001). This variable can be assessed quantitatively on an individual level for both boys and girls, while maintaining a strong gender focus by considering the differential impact of culture-specific gender norms. By doing so, this study elaborates on previous studies published in Sex Roles on the educational gender gap in Flanders (Van de Gaer et al., 2006a; 2008), masculine study cultures in German schools (Heyder & Kessels, 2013) and gender-conformity pressures among pre-teens and early adolescents in the United States (Carver et al., 2003). Moreover, by focusing on gendered processes in an understudied population such as Belgian adolescents, this article helps to highlight possible cross-national variations in gendered processes. More specifically, this study will investigate
gendered study cultures among Belgian adolescents by linking pressure for gender conformity to a central educational parameter, that is, academic self-efficacy.

Academic self-efficacy refers to people’s beliefs in their capability to perform certain academic tasks (Bandura, 1997; Schunk & Pajares, 2001; Zimmerman, 2000), such as the capability to motivate oneself for studying and to finish school work on time (Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011). Given self-efficacy has proven to be one of the central predictors of academic achievement according to a meta-analysis in several countries (Hattie, 2008), it is an interesting mediator in the study of the educational gender gap. Furthermore, because early adolescence is a crucial time-period, both for the development of gender roles (U.S.: Carver et al., 2003; Horn, 2007) and for the emergence of gender-differential achievement patterns (Belgium: Derks & Vermeersch, 2002; OECD-countries: PISA, 2009), we will research developmental changes by considering associations at the start and the end of the first year of secondary education. Hence, the present study examines the gender-differentiated impact of gender-conformity pressure on academic self-efficacy in a sample of Belgian seventh graders, using descriptive and multilevel regression analysis at two time-points.

In what follows, we consider international and Belgian research concerning gender-specific study cultures. Subsequently, we delve deeper into pressure for gender conformity and how it relates to study cultures. Next, we frame the processes under study in its cultural context by considering gender (in)equality and education in Belgium. Then, we link the educational gender gap findings to academic self-efficacy, before we derive our hypotheses.

**GENDER CULTURES AT SCHOOL**

In Belgium, the educational gender gap has manifested itself on several parameters in the last few decades. For instance, boys repeat grades more often than girls (Van Landeghem et al., 2010), have lower marks in class (Van de Gaer et al., 2006a), are less enrolled in academic-oriented tracks (Van Landeghem & Van Damme, 2007; Van Woensel, 2007), drop out more often without qualification (Van Landeghem et al., 2010), have lower enrollment in higher education (Van Woensel, 2007) and are overrepresented in special education services (Van Landeghem & Van Damme, 2007).

Similar gender-differentiated achievement gaps are observed among most western industrialized countries. Although there is substantial variation in the gender gap among these countries, boys are overrepresented in special education
services in all OECD-countries (Benjamin, 2003). PISA research in OECD countries (2009) also indicates that, on average, girls outperform boys in reading throughout the school career, are more present in academic-oriented tracks, drop out less often without qualification and graduate more often from tertiary education.

Ethnographic studies suggest that these gender-differentiated achievement patterns may emanate from cultures of masculinity that hinder boys in attaining educational success (e.g., UK: Mac an Ghaill, 1994; Australia: Martino, 1999; U.S.: Pascoe, 2007). In several countries, for instance, certain school courses and study behaviors are considered feminine, impeding boys’ identification with and motivation for school (U.S.: Bhanot & Jovanovic, 2005; Pajares & Valiante, 2001; Australia: Connell, 1996; Martino, 1996, 1999; Belgium: Derks & Vermeersch, 2002; Van de Gaer, Pustjens, Van Damme, & De Munter, 2006b; UK: Epstein, 1998; Germany: Heyder & Kessels, 2013). Conversely, characteristics associated with traditional femininity (such as being tidy, cooperative, and passive) are more in line with a studious attitude, which may give girls an advantage at school (UK, U.S. & Australia: Beaman et al., 2006; UK: Jones & Myhill, 2004). In Belgium, quantitative evidence of gendered peer cultures has been provided by Van Houtte (2004a, 2004b). She showed that boys’ peer groups in secondary schools display less conducive study attitudes than girls’ peer groups and that this difference was largely responsible for boys’ lower achievement (Van Houtte, 2004b).

**PRESSURE FOR GENDER CONFORMITY**

We suggest that the concept of pressure for gender conformity could be a promising addition to the study of gendered study cultures. Pressure for gender conformity refers to the degree to which one experiences pressure from significant others and self to exhibit gender-congruent behavior (e.g., Egan & Perry, 2001). According to U.S. research, conformity pressures are most prominent in childhood and early adolescence (Carver et al., 2003; Horn, 2007) and results in displaying less cross-gender behavior (Carver et al., 2003). However, when young people limit their activities, the exploration of their talents is impeded (Yunger et al., 2004). Therefore, it will come as no surprise that pressure for gender conformity has been consistently shown to negatively impact adolescents’ psychological adjustment in both the U.S. and the UK (U.S.: Carver et al., 2003; Corby, Hodges, & Perry, 2007; Egan & Perry, 2001; Yunger, Carver, & Perry, 2004; UK: Menon, 2011); this demonstrates the importance of well-being as a control variable in research considering gender-conformity pressures. Although these processes are expected to
be similar in Belgium, no research into pressure for gender conformity has been conducted with Belgian students.

So far, few studies have linked pressure for gender conformity with educational parameters. As a notable exception, Leaper, Farkas, and Brown (2012) showed that U.S. adolescent girls’ ability and value beliefs for English were higher when experiencing pressure for gender conformity, whereas their beliefs for STeM (science, technology, math) courses were lower when experiencing such pressure. Note that these course-specific influences can be linked to the perceived gender-stereotypic nature of these school subjects in the U.S. That is, STeM courses are often perceived as being masculine, whereas language is seen as being rather feminine (Meece, Glienke, & Burg, 2006; Pajares & Valiante, 2001).

The above findings demonstrate the importance of considering context and identity contents when considering gendered processes. This is a notion that has been proposed by several theories. For instance, the Social Constructionist Theory states that people’s behavior is enacted in accordance to their gender categorization, which is based on cultural conventions for appropriate conduct (West & Zimmerman, 1987). As such, expectations for girls and boys would be different, based on the culture-specific conventions regarding appropriate boyish or girlish behavior at school.

Similarly, identity content is an important aspect of the identity-based motivation model. This model proposes that the traits, that are associated with a social identity, determine which behavior will come to mind (Oyserman, Bybee, & Terry, 2006; Oyserman & Destin, 2010). When a masculine gender identity is predominantly linked to being an academic underachiever, class clown, or effortless achiever—as it is for instance in the UK, U.S. and Belgium (Belgium: Derks & Vermeersch, 2002; UK: Francis, 2000; Mac an Ghaill, 1994; U.S.: Morris, 2012)—, this will reduce boys’ chances of academic engagement (for a confirmation with a U.S. sample, see Elmore & Oyserman, 2012). Conversely, when a feminine identity is linked to being a hard-working and well-behaved student (UK, U.S. & Australia: Beaman, Wheldall, & Kemp, 2006; UK: Jones & Myhill, 2004), this will raise girls’ chances of demonstrating such behavior (U.S.: Elmore & Oyserman, 2012). These processes have been confirmed for a Belgian context with 12-year-olds. Among those who perceived themselves to be typical for their gender, girls were more likely to demonstrate higher autonomous study motivation (Vantieghem et al., 2013) and self-efficacy than boys (Vantieghem, Vermeersch, & Van Houtte, 2014a).
Self-efficacy contributes in both direct and indirect ways to achievement in various countries, including Belgium (Belgium: Ferla, Valcke, & Schuyten, 2010; Italy: Bandura, Barbaranelli, Caprara, & Pastorelli, 1996; Caprara et al., 2011; U.S.: Greene, Miller, Crowson, Duke, & Akey, 2004). A meta-analysis covering various countries has shown that self-efficacy beliefs account for 14% of variation in student’s academic performance with a standardized effect size of .38 (i.e., students’ performance increased with about one third of a standard deviation; Multon, Brown, & Lent, 1991). Because self-efficacy beliefs affect how effectively people apply their knowledge, it is a better predictor of performance than skills or intellectual aptitude alone. Hence, in U.S. and Canadian research with children of the same ability level, more efficacious children still obtained better scores (Bandura, 1997).

Self-efficacy also contributes indirectly to academic achievement through important learning behaviors. It has, for instance, been shown that self-efficacious university students in Belgium tend to persist longer, have more mastery goals and adopt a deep-learning approach more often than non-efficacious peers (Ferla, Valcke, & Schuyten, 2010). Similar results have been found for U.S. students (e.g., Multon et al., 1991; Pintrich & De Groot, 1990; Walker, Greene, & Mansell, 2006; Zimmerman, 2000).

With regard to gender differences, self-efficacy poses a nuanced and complex picture. For instance, according to a meta-analysis conducted in 15 countries (Huang, 2013), gender differences in self-efficacy start to occur in early adolescence and increase with age. Specifically, school subject was found to be a central moderator for gender differences. That is, boys have on average a higher self-efficacy in mathematics (effect size: \( g = .18 \)) and computer sciences (\( g = .18 \)). Girls have a higher self-efficacy in language arts (\( g = .16 \)) and a negligible advantage in general academic self-efficacy (\( g = .03 \)). Findings on course-specific self-efficacy in Belgium are similar, with lower mathematic self-efficacy among 15-year old girls than boys (Ferla, Valcke, & Cai, 2009).

As highlighted in the previously reviewed findings, investigating course-specific self-efficacy is necessarily gender-biased in either boys’ or girls’ favor. Hence, it would be more interesting to investigate a less gender-biased form of self-efficacy, such as efficacy for self-regulated learning—which has only a mild female advantage (meta-analysis: Huang, 2013)— and refers to the capacity to plan and organize academic activities. Because self-regulated learning comprises skills necessary for any school course, it is also more pertinent to general achievement.
than self-efficacy in a specific subject (Pajares, 2002; Pajares & Valiante, 2001) and is thus an interesting mediator for educational gender gap research. Self-efficacy for self-regulated learning functions similarly in Belgian samples, with a small female advantage (Vantieghem et al., 2014a) and relating to academic skills above and beyond general academic self-efficacy (Ferla et al., 2010).

THE BELGIAN CONTEXT

Because culture-specific conditions could alter the influence of gender-conformity pressures on educational parameters, we will consider both gender-equality levels and the educational context of Belgium.

In Belgium, there are officially two tracks in the first grade of secondary education: the general and the vocational-oriented track (Van Houtte, Demanet, & Stevens, 2012). When students have not achieved well enough, they can be required to change track or repeat a year (Van Houtte et al., 2012). Such grade-retention is common practice, resulting in over 30% of Belgian students being older than their classmates at the end of secondary education (Van Landeghem & Van Damme, 2007). Also, students of non-western European descent and with lower socioeconomic background disproportionally end up in the lower esteemed tracks (Boone & Van Houtte, 2012; Van Houtte & Stevens, 2009). Track, age, initial academic ability, socioeconomic and ethnic background are influential aspects of the Belgian educational system. They have also been shown to influence a range of important educational parameters, including academic self-efficacy (for research in several countries regarding the influence of age, SES and ability: Bandura, 1997; for an overview regarding the influence of track, ethnicity, age, and prior achievement: Schunk & Pajares, 2001), and hence constitute important control variables.

Another set of factors that might influence the processes under study are gendered beliefs and practices. Belgium is a relatively gender-equal country, ranked 12th on the gender equality index among the very high human development countries (UNDP, 2011). This composite measure reflects inequality between men and women on reproductive health, empowerment and the labor market, with more gender egalitarian countries ranking higher (in comparison, the U.S. occupies place 47, the UK place 34). As such, we can posit that Belgium is a progressive country that scores rather high on gender equality measures. Nevertheless, full gender equality is not achieved in Belgium, as evidenced by gender-differentiated achievement patterns and rather traditional work-care distributions in most households (RoSa, 2014). That is, male roles have changed little in the last decades, with few men performing care-related and domestic tasks. For example, just 1% of
unemployed Belgian men is a housemaker and less than 10% of wage-earning men works part time (FOD-Economie, 2013). Conversely, most Belgian women are combining professional and care-related tasks. Because the scope of the male role has remained more traditional and limiting, this might reduce Belgian boys’ acceptable educational options. And, because of the lack of flexibility in the male role, it seems plausible that pressure for gender conformity might be higher among Belgian boys than girls.

THE PRESENT STUDY

In this study, we investigate the impact of pressure for gender conformity on self-efficacy for self-regulated learning in a sample of Belgian seventh graders. In line with international research (Huang, 2013; Pajares, 2002) and research from a Belgian context (Vantieghem et al., 2014a), we expect girls to score somewhat higher than boys on academic self-efficacy in the regression analyses (hypothesis 1a). Also, based on the evolution of gendered beliefs in Belgium (see section: The Belgian context), we expect to see boys experiencing more pressure for gender conformity than girls in the analyses of variance (hypothesis 1b).

Furthermore, using regression analysis, we expect a positive association between gender-conformity pressure and self-efficacy for girls (hypothesis 2a) and a negative association for boys (hypothesis 2b). This expectation is based on an empirical study with U.S. girls by Leaper and colleagues (2012), the identity-based motivation model, and Belgian research that demonstrated that traditional masculinity is problematic for academic achievement (Derks & Vermeersch, 2002), whereas feminine identities are linked to higher academic self-efficacy (Vantieghem et al., 2014a).

To better assess changes in the critical developmental period of early adolescence, longitudinal data from two time-points will be employed. We expect stronger effects in the regression analyses at the end of the school year than at the beginning (hypothesis 3). This is based on a meta-analysis in different countries demonstrating that self-efficacy differences tend to increase with age (Huang, 2013) and the social dosage-effect, which states that gender-conformity pressure is likely to be stronger the longer students are together (Martin & Fabes, 2001).

In summary, this study considers three hypotheses:
1) Belgian girls experience less pressure for gender conformity and have a higher sense of academic self-efficacy in the first year of secondary education than boys, controlling for students’ subjective well-being, track, age, initial academic ability, ethnicity, and socioeconomic status.
2) The net effect of pressure for gender conformity has a positive influence on girls’ academic self-efficacy, and a negative influence on boys’ self-efficacy.

3) The net associations between gender-specific pressure for gender conformity and academic self-efficacy are stronger at the end of the school year than at the beginning.

To account for the nested structure of the data with students clustered in schools, we employ multilevel analysis using the statistical program MLwiN. The benefit of this method is, in contrast to simple regression analysis, that it correctly accounts for the successive sampling method (as will be discussed below), and thus avoids the problem of autocorrelation. Another benefit is that multilevel analysis can tease out the amount of variance on each level. That is, we can ascertain how much variance is located at the individual versus the school level. To do this, we first run a so-called zero-model, where no variables are included and calculate the variance at each level (see results). In the next model, we include the main variables. All continuous variables are grand mean centered to improve interpretability of the intercept (for a more detailed discussion of multilevel analysis: Hox, 2010).

In the final model, we include the control variables. This way we can assess whether the associations increase, diminish, or disappear after taking socioeconomic position, initial scholastic ability, subjective well-being, school track, ethnicity and age into account, since these variables have shown to be of importance for academic self-efficacy, the Belgian educational context (see section: the Belgian context) or with regards to pressure for gender conformity (see section: Pressure for gender conformity). To investigate the evolution of the impact of pressure for gender conformity on academic self-efficacy (see hypothesis 3), we analyze these associations on two time-points: the start and the end of seventh grade.

METHOD

Participants
The data are part of the “Teaching in the Bed of Procrustes”-study and were gathered in two waves (at the beginning and the end of the 2012-13 school year). Schools were selected so each geographical region within Flanders was equally represented. Flanders is the northern, Dutch-speaking part of Belgium and has its own Department of Education. Therefore, Flanders has a more homogenous educational policy than Belgium as a whole, which is why the sample was selected from this region rather than the whole country. Selection was further dependent on school denomination (public versus parochial) and a proportional representation of
rural versus city schools. Within these parameters, three random samples were drawn. For each school that refused, a matched school from the next random sample was contacted. This way, we aimed to obtain schools that represent the Flemish educational context on important parameters (such as region, school denomination, and location), while maintaining randomness within these subpopulations. Of the 124 contacted schools 59 schools participated in the study, which translates to a response rate of 47.6%. Analyses in which we compared our sample to the Flemish school population, based on information from the Flemish Educational Department, showed no important differences in school sector, curriculum, or student composition. This suggests that no systematic biases occurred and that the schools were representative of the population.

Within these schools, all seventh graders were asked to complete the questionnaire. The use of child assent was approved by the school and the Belgian Commission for the Protection of Privacy, based on the minimal risk of the study. Researchers were always present during the completion of the survey to explain the procedure and answer questions. Pupils were told that the survey was not a test and were assured that it was completely confidential. The response rate among the students was 96.9%, translating to a total of 6380 students. For the second wave, cooperation with one school was terminated because of problems with upholding appointments. So, in the second wave, 58 schools cooperated, which translated to 6234 students filling out the questionnaire. Of the students in the sample, the majority were boys (53.8%) and of Belgian or western-European descent (79.9%). The socioeconomic status (SES) of the students was based on parents’ occupation. In this sample, 22% of the students had a blue-collar background. Furthermore, 11.7% of pupils were in the vocational-oriented track. Because the sample consisted of seventh graders, the mean age was 12 years old. There was a range from 10 to 15 years old, indicating that some pupils skipped a year, whereas others had repeated years. Nevertheless, 76.6% of the sample was 12 years old, demonstrating that the majority were on track. The demographic characteristics of the study participants in the final sample divided by gender can be seen in Table 14.
Table 14.
Demographic characteristics of the participants in the final sample by gender.

| Time-point 1: | Girls | | | Boys | | |
|-------------|-------|-------------------|---|-------------------|---|
| Variables   | Mean  | SD    | Range | %   | N    | Mean  | SD    | Range | %   | N    |
| % West-European ethnicity | 84.1 | 2312 |     | 84.9 | 2555 |     |
| % blue collar background | 16.9 | 2312 |     | 19.8 | 2555 |     |

| Time-point 2: | Girls | | | Boys | | |
|-------------|-------|-------------------|---|-------------------|---|
| Variables   | Mean  | SD    | Range | %   | N    | Mean  | SD    | Range | %   | N    |
| % West-European ethnicity | 84  | 2302 |     | 85.0 | 2504 |     |
| % blue collar background | 17.1 | 2302 |     | 19.7 | 2504 |     |
Measures

**Academic self-efficacy.** To assess academic self-efficacy, we used an eight-item scale based on Caprara and colleagues (Caprara et al., 2011). These items concern the self-efficacy for self-regulated learning activities, such as the capacity to plan and organize academic activities, to structure environments conducive to learning and to motivate oneself for schoolwork. The scale presents a stem followed by a question (e.g., “How well can you...”: “finish homework assignments by deadlines?”; “Arrange a place to study without distractions?”). Students answered on a 5-point likert scale, ranging from 0 = cannot do at all, to 4 = can do very well. The scale was construed using mean sum of scores, so higher scores constitute a higher sense of academic self-efficacy. The scale displayed good internal reliability (see Table 15).

**Pressure for gender conformity.** To measure pressure for gender conformity, we used the pressure for gender conformity subscale from the gender-identity questionnaire by Egan and Perry (2001). This subscale measures the extent to which one experiences pressure for gender-conforming behavior, with four items related to pressure from peers and four items related to pressure from oneself (e.g., “I think it is important to act just like other [girls/boys]”, “My friends would be upset if I wanted to play with [boys’/girls’] toys”). The scale consists of separate versions for boys and girls. The Dutch translation and answering format with a 4-point likert scale (completely disagree, disagree, agree, completely agree) are based on the work of Bos and Sandfort (2010). This scale was constructed using mean sum of scores, with high scores indicating more pressure for gender conformity. The scale displayed good internal reliability (see Table 15).

**Subjective well-being.** Subjective well-being was measured by asking students how often they had felt certain emotions in the past 30 days. They indicated the frequency of these feelings on a 5-point scale ranging from 0 = never to 4 = very often (Keyes, Shmotkin, & Ryff, 2002) (Keyes et al., 2002) (Keyes et al., 2002). The scale consisted of 12 items, equally divided between positive (e.g., happy, peaceful) and negative items (e.g., nervous, worthless). Based on principal component analysis, we constructed a single scale from these items. The items pertaining to negative affect were reverse coded so higher scores on the scale indicate a higher well-being. This scale was constructed using a mean sum of scores and displayed good internal reliability (see Table 15).
### Table 15.
Descriptive statistics

<table>
<thead>
<tr>
<th>Variables</th>
<th>Range</th>
<th>Cronbach's alpha</th>
<th>% refcat</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>% refcat</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy (Time 1)</td>
<td>0-4</td>
<td>.84</td>
<td></td>
<td>2.73</td>
<td>.62</td>
<td>2515</td>
<td></td>
<td>2.53</td>
<td>.66</td>
<td>2864</td>
</tr>
<tr>
<td>self-efficacy (Time 2)</td>
<td>0-4</td>
<td>.87</td>
<td></td>
<td>2.55</td>
<td>.69</td>
<td>2392</td>
<td></td>
<td>2.37</td>
<td>.73</td>
<td>2624</td>
</tr>
<tr>
<td>Gender pressure (Time 1)</td>
<td>0-3</td>
<td>.77</td>
<td></td>
<td>1.19</td>
<td>.50</td>
<td>2515</td>
<td></td>
<td>1.69</td>
<td>.53</td>
<td>2864</td>
</tr>
<tr>
<td>Gender pressure (Time 2)</td>
<td>0-3</td>
<td>.81</td>
<td></td>
<td>1.23</td>
<td>.49</td>
<td>2392</td>
<td></td>
<td>1.72</td>
<td>.51</td>
<td>2624</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being (Time 1)</td>
<td>.36-4</td>
<td>.81</td>
<td></td>
<td>2.84</td>
<td>.52</td>
<td>2515</td>
<td></td>
<td>2.86</td>
<td>.50</td>
<td>2864</td>
</tr>
<tr>
<td>Well-being (Time 2)</td>
<td>0-4</td>
<td>.86</td>
<td></td>
<td>2.83</td>
<td>.61</td>
<td>2392</td>
<td></td>
<td>2.87</td>
<td>.57</td>
<td>2624</td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td>0-50</td>
<td></td>
<td></td>
<td>26.54</td>
<td>9.97</td>
<td>2515</td>
<td></td>
<td>28.06</td>
<td>10.78</td>
<td>2864</td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td>0-25</td>
<td></td>
<td></td>
<td>15.5</td>
<td>5.20</td>
<td>2515</td>
<td></td>
<td>13.83</td>
<td>5.27</td>
<td>2624</td>
</tr>
<tr>
<td>Age</td>
<td>10-15</td>
<td></td>
<td></td>
<td>12.16</td>
<td>.45</td>
<td>2515</td>
<td></td>
<td>12.22</td>
<td>.52</td>
<td>2864</td>
</tr>
<tr>
<td>Track (Time 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 = academic</td>
<td></td>
<td></td>
<td></td>
<td>90.6</td>
<td></td>
<td>3009</td>
<td></td>
<td>85.7</td>
<td></td>
<td>3514</td>
</tr>
<tr>
<td>1 = vocational</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Track (Time 2)</td>
<td></td>
<td></td>
<td></td>
<td>90.5</td>
<td></td>
<td>2826</td>
<td></td>
<td>85.6</td>
<td></td>
<td>3227</td>
</tr>
<tr>
<td>0 = academic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 = vocational</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 = white-collar</td>
<td></td>
<td></td>
<td></td>
<td>79.8</td>
<td></td>
<td>2870</td>
<td></td>
<td>76.4</td>
<td></td>
<td>3324</td>
</tr>
<tr>
<td>1 = blue-collar</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 = West-European</td>
<td></td>
<td></td>
<td></td>
<td>79.8</td>
<td></td>
<td>2867</td>
<td></td>
<td>80.0</td>
<td></td>
<td>3303</td>
</tr>
<tr>
<td>1 = Non-West-European</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. Gender pressure = pressure for gender conformity. % refcat = percent of students in the reference category or value = 0 of a categorical variable.*
Initial academic ability. Initial academic ability was measured by two validated tests for students of this age in mathematics and Dutch, which was administered at the start of seventh grade. The mathematics test consisted of 50 questions regarding problematical cases and calculations (Dudal, 2003). The Dutch test consisted of four texts the children had to read and for which they had to answer questions concerning content and vocabulary (CITO, 2010). The scales were constructed separately for Dutch and mathematics by a sum score of the correct answers on each test.

RESULTS

Descriptive Statistics
We conducted a preliminary repeated measures ANCOVA to test for a possible Time x Variable x Gender interaction. Time (Time 1 or Time 2) and Variable (self-efficacy, well-being, gender-conformity pressure) were Within-Subjects factors. Gender was a Between-Groups factor. Age, Ethnicity (0 = western-European descent, 1 = Non-western-European descent), track (0 = academic track, 1 = vocational track), SES (0 = white-collar background, 1 = blue-collar background) and initial academic ability were covariates. The Time x Variable x Gender interaction was not significant, $F(2, 4681) = 1.53, p < .216, \eta^2 = .001$, indicating that the responses of boys and girls on the variables did not differ across time. Given the absence of this interaction, subsequent analyses were done separately for each time-point.

To test for average gender differences in continuous measures on Time 1 and 2, a MANOVA was run with gender as Between-Groups factor. Means and standard deviations for all variables can be seen in Table 15. The MANOVA-tests revealed a multivariate effect for gender on both time-points, $F_1(6, 5372) = 297.39, p < .001, \eta^2 = .249; F_2(6, 5009) = 270.78, p < .001, \eta^2 = .245$, indicating that girls and boys scored differently on average on the variables. To assess whether girls outscored boys on self-efficacy and scored lower on pressure for gender conformity (see hypothesis 1), we consider the univariate MANOVA results. In support of hypothesis 1, girls had a somewhat higher score on academic self-efficacy than boys at both time-points, $F_1(1, 5377) = 133.89, p < .001, \eta^2 = .024; F_2(1, 5014) = 84.09, p < .001, \eta^2 = .016$, and a lower score on pressure for gender conformity, $F_1(1, 5377) = 1279.00, p < .001, \eta^2 = .192; F_2(1, 5014) = 1197.04, p < .001, \eta^2 = .193$. In line with previous research, girls scored somewhat higher on the language test, $F(1, 5377) = 136.60, p < .001, \eta^2 = .025$, whereas boys did negligibly better on the mathematics test, $F(1, 5377) = 28.41, p < .001, \eta^2 = .005$. Girls were
negligibly younger (indicating that they had repeated less grades than boys) \( F(1, 5377) = 21.10, p < .001, \eta^2 = .004 \), and scored a little lower on well-being at Time 2, \( F(1, 5014) = 6.87, p = .009, \eta^2 = .001 \). At Time 1, there was no significant gender difference on well-being, \( F(1, 5377) = .77, p = .381, \eta^2 = 0 \).

Table 16.
Zero-order correlations at Time 1 divided by gender.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Self-efficacy</th>
<th>Gender-conformity pressure</th>
<th>Well-being</th>
<th>Math</th>
<th>Dutch</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy</td>
<td>—</td>
<td>.005</td>
<td>.289**</td>
<td>.062**</td>
<td>.022</td>
<td>.001</td>
</tr>
<tr>
<td>Pressure for gender conformity</td>
<td>-.004</td>
<td>—</td>
<td>-.081**</td>
<td>-.076**</td>
<td>-.125**</td>
<td>.053**</td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being</td>
<td>.325**</td>
<td>-.151***</td>
<td>—</td>
<td>.114**</td>
<td>.095**</td>
<td>-.088**</td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td>.082**</td>
<td>-.107***</td>
<td>.108**</td>
<td>—</td>
<td>.550**</td>
<td>-.341**</td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td>-.01</td>
<td>-.168***</td>
<td>.113**</td>
<td>.584**</td>
<td>—</td>
<td>-.315**</td>
</tr>
<tr>
<td>Age</td>
<td>-.05**</td>
<td>.043*</td>
<td>-.087***</td>
<td>-.363**</td>
<td>-.334***</td>
<td>—</td>
</tr>
</tbody>
</table>

*Note.* Correlations for boys are above the diagonal; Correlations for girls are below the diagonal.  
* p < .05. ** p < .01. *** p < .001.

Table 17.
Zero-order correlations at Time 2 divided by gender.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Self-efficacy</th>
<th>Gender-conformity pressure</th>
<th>Well-being</th>
<th>Math</th>
<th>Dutch</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-efficacy</td>
<td>—</td>
<td>-.056**</td>
<td>.280**</td>
<td>.061**</td>
<td>.017</td>
<td>-.037*</td>
</tr>
<tr>
<td>Pressure for gender conformity</td>
<td>-.001</td>
<td>—</td>
<td>-.074**</td>
<td>-.087**</td>
<td>-.138**</td>
<td>.065**</td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being</td>
<td>.353**</td>
<td>-.161***</td>
<td>—</td>
<td>.172**</td>
<td>.147**</td>
<td>-.123**</td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td>.099**</td>
<td>-.084***</td>
<td>.151**</td>
<td>—</td>
<td>.550**</td>
<td>-.341**</td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td>.025</td>
<td>-.174***</td>
<td>.128**</td>
<td>.584**</td>
<td>—</td>
<td>-.315**</td>
</tr>
<tr>
<td>Age</td>
<td>-.041*</td>
<td>.024</td>
<td>-.135***</td>
<td>-.363**</td>
<td>-.334***</td>
<td>—</td>
</tr>
</tbody>
</table>

*Note.* Correlations for boys are above the diagonal; Correlations for girls are below the diagonal.  
* p < .05. ** p < .01. *** p < .001.
To test for average gender differences on categorical variables, chi-square tests were run. In line with the educational gender gap, girls were more present in the academic-oriented track throughout the school year than boys, $X^2_1(1, N = 6523) = 35.56, p < .001, \nu = .074; X^2_2(1, N = 6053) = 33.43, p < .001, \nu = .074$. There were no differences between boys and girls on ethnicity status, $X^2(1, N = 6170) = .05, p = .831, \nu = .003, ns$, while girls had negligibly more often a white-collar background than boys, $X^2(1, N = 6194) = 10.05, p < .01, \nu = .04$.

The zero-order correlations in Table 16 (for Time 1) and 17 (for Time 2) show that the associations between the variables did not differ much between boys and girls, with the exception of the association between well-being and pressure for gender conformity. This association is on both times almost twice as large among girls, $r_1(2872) = -.15, p < .001; r_2(2742) = -.16, p < .001$, as it is among boys, $r_1(3345) = -.08, p < .001; r_2(3059) = -.07, p < .001$. This difference is significant on both time-points, $Z_1 = 2.79, p < .01, Z_2 = 3.356, p < .001$. The largest correlation was between the initial ability tests for Dutch and mathematics, $r(2642) = .58$. Nevertheless, there were no problems of multicollinearity since VIFs for all variables were smaller than two. Quite notable is that at Time 1, there is no significant association between pressure for gender conformity and academic self-efficacy, $r_{girls}(2849) = -.004, p = .81; r_{boys}(3294) = .005, p = .78$. At Time 2, this changes for boys, $r_{boys}(3052) = -.056, p < .01$, but not for girls, $r_{girls}(2727) = -.001, p = .98$. We explore these associations more deeply in the multilevel regressions, discussed below.

**Analysis of Time 1**

The zero-model from the multilevel analysis showed that academic self-efficacy varied somewhat between schools. However, only 2.11% ($e_{0ij} = .417, u_{0j} = .009$) of total variance was situated at the school-level. Variance components are random for the intercept and individual error term, which is necessary in multilevel analysis and allows variance in scores between individuals and schools. Variance components for other variables were kept fixed to increase statistical power, given there was no significant random slope variance. This indicates that the associations between the independent variables and academic self-efficacy did not fluctuate in different schools (for a more detailed discussion of random slopes: Hox, 2010).

In line with hypothesis 1a, we expect a female advantage in academic self-efficacy. To assess the extent of this gender difference, we consider the effect of gender on academic self-efficacy. In line with hypothesis 2, we expect a negative association between pressure for gender conformity and self-efficacy for boys and a positive association for girls. To determine whether this is the case, we consider the
interaction effect between gender and pressure for gender conformity. Model 1, which included the main variables gender, pressure for gender conformity and its interaction, gives us a first insight into these associations on Time 1.

Table 18.
Results of multilevel analysis at Time 1: Unstandardized coefficients and standard errors in parentheses

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.529***</td>
<td>2.47***</td>
</tr>
<tr>
<td></td>
<td>(.018)</td>
<td>(.018)</td>
</tr>
<tr>
<td>Sex (0 = boy)</td>
<td>.171***</td>
<td>.233***</td>
</tr>
<tr>
<td></td>
<td>(.02)</td>
<td>(.02)</td>
</tr>
<tr>
<td>Gender-conformity pressure</td>
<td>.11</td>
<td>.046</td>
</tr>
<tr>
<td></td>
<td>(.022)</td>
<td>(.023)</td>
</tr>
<tr>
<td>Gender X pressure for conformity</td>
<td>-.023</td>
<td>-.007</td>
</tr>
<tr>
<td></td>
<td>(.034)</td>
<td>(.035)</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being</td>
<td>.403***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.021)</td>
<td></td>
</tr>
<tr>
<td>Track (0 = general)</td>
<td>.042</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.040)</td>
<td></td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td>-.004</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td></td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td>.005***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td></td>
</tr>
<tr>
<td>SES (0 = white collar)</td>
<td>-.019</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.025)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity (0 = West-European)</td>
<td>.188***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.027)</td>
<td></td>
</tr>
<tr>
<td><strong>Variance components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School $u_{ij}$</td>
<td>.007</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.002)</td>
</tr>
<tr>
<td>Student $e_{0ij}$</td>
<td>.411</td>
<td>.358</td>
</tr>
<tr>
<td></td>
<td>(.008)</td>
<td>(.007)</td>
</tr>
</tbody>
</table>

*Note.** Tests represented here were two-sided. $N_{School}$: 58, $N_{Student}$: 4867

*p < .05. **p < .01. ***p < .001
As seen in Table 18, model 1 showed a modest, but significant association between gender and academic self-efficacy ($\gamma = .171, p < .001$). This indicates that girls scored somewhat higher on academic self-efficacy than boys, supporting hypothesis 1a. Pressure for gender conformity, however, did not have a significant effect ($\gamma = .11; p = .617$). The interaction-term between gender and pressure for gender conformity ($\gamma = -.023; p = .499$) neither yielded significant results, which runs against hypothesis 2. Introducing the control variables in the second model did not change the overall pattern of the model.

Overall, the final model explained 14.79% of the total variance in academic self-efficacy.

**Analysis of Time 2**

The zero-model from the multilevel analysis on Time 2 showed that academic self-efficacy continued to vary somewhat between schools. Similar to the value at Time 1, 2.29% ($e_{0ij} = .513, u_{0i} = .012$) of total variance was situated at the school-level (see Table 19). Furthermore, variance components are random for the intercept and individual error term. The significant random slope for pressure for gender conformity at Time 2 indicated that the impact of pressure for gender conformity on academic self-efficacy varied somewhat between schools at the end of the school year. Variance components for other variables were kept fixed to increase statistical power (Hox, 2010).

Similar to the analyses on Time 1, we expect a female advantage in academic self-efficacy (hypothesis 1a), and a negative association between pressure for gender conformity and self-efficacy for boys and a positive association for girls (hypothesis 2). To assess the extent of this gender difference, we consider the effect of gender on academic self-efficacy, and the interaction effect between gender and pressure for gender conformity. Model one, which included the main variables gender, pressure for gender conformity and its interaction, gives us a first insight into these associations on Time 2.

As seen in Table 19, model 1 showed a significant association between gender and academic self-efficacy ($\gamma = .148, p < .001$). This indicates that girls continued to score somewhat higher on academic self-efficacy than boys, confirming hypothesis 1a. Pressure for gender conformity now yielded a modest negative effect ($\gamma = -.078; p < .01$), while the interaction-term between gender and gender-conformity pressure ($\gamma = .076; p < .046$) showed that the influence of pressure significantly differed according to students’ gender. As can be seen in Figure 8, boys’ academic self-efficacy was lower when experiencing higher levels of
pressure, supporting hypothesis 2b. Girls’ academic self-efficacy, however, tended to remain at the same level regardless of the amount of pressure they experienced.

### Table 19.
Results of multilevel analysis at Time 2:
Unstandardized coefficients and standard errors in parentheses

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.369***</td>
<td>2.311***</td>
</tr>
<tr>
<td></td>
<td>(.02)</td>
<td>(.018)</td>
</tr>
<tr>
<td>Sex (0 = boy)</td>
<td>.148***</td>
<td>.229***</td>
</tr>
<tr>
<td></td>
<td>(.022)</td>
<td>(.022)</td>
</tr>
<tr>
<td>Gender-conformity pressure</td>
<td>-.078**</td>
<td>-.038</td>
</tr>
<tr>
<td></td>
<td>(.025)</td>
<td>(.032)</td>
</tr>
<tr>
<td>Gender X pressure for conformity</td>
<td>.076*</td>
<td>.108**</td>
</tr>
<tr>
<td></td>
<td>(.038)</td>
<td>(.04)</td>
</tr>
</tbody>
</table>

### Control variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Well-being</td>
<td>.38***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.017)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.031</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.023)</td>
<td></td>
</tr>
<tr>
<td>Track (0 = general)</td>
<td>.151***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.043)</td>
<td></td>
</tr>
<tr>
<td>Initial ability: Dutch</td>
<td>-.003</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td></td>
</tr>
<tr>
<td>Initial ability: Math</td>
<td>.005***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td></td>
</tr>
<tr>
<td>SES (0 = white collar)</td>
<td>-.063*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.028)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity (0 = West-European)</td>
<td>.178***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.03)</td>
<td></td>
</tr>
</tbody>
</table>

### Variance components

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>School u_{0j}</td>
<td>.001</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>(.001)</td>
</tr>
<tr>
<td>Gender-conformity pressure u_{02j}</td>
<td>.012</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.006)</td>
<td></td>
</tr>
<tr>
<td>Student e_{0ij}</td>
<td>.508</td>
<td>.443</td>
</tr>
<tr>
<td></td>
<td>(.009)</td>
<td>(.009)</td>
</tr>
</tbody>
</table>

*Note. Tests represented here were two-sided. N_{School:} 58, N_{Student:} 4806
" p < .05. ** p < .01. *** p < .001.
Figure 8.
Association between sex, gender-conformity pressure and academic self-efficacy

Figure 9.
The association between sex, gender-conformity pressure and academic self-efficacy, controlled for well-being
Introducing the control variables in the last model somewhat changed the patterns of the model. The main effect of pressure for gender conformity lost its significance ($\gamma = -0.038; p = .118$). However, the interaction effect with gender continued to be significant and even increased its effect size ($\gamma = .108; p < .01$). This finding indicates that by considering the opposing ways gender and pressure interact, the more general effects of pressure for gender conformity disappear. Analysis showed this to be mostly due to the introduction of subjective well-being ($\gamma = .38; p < .001$). This effect of subjective well-being is further clarified when considering Figure 9. Similar to Figure 8 and in line with hypothesis 2b, boys demonstrated lower levels of academic self-efficacy when experiencing more pressure. In contrast to Figure 8 and confirming hypothesis 2a, once controlled for well-being, girls demonstrated higher levels of self-efficacy when experiencing more pressure. This implies a suppression effect, given the positive effect of pressure on girls’ self-efficacy is only visible when keeping in mind the negative effect of pressure on well-being.

Overall, the final model explained 16.95% of the total variance in academic self-efficacy.

DISCUSSION

As in many western industrialized countries, the educational gender gap in Belgium manifests with girls outperforming boys on many educational parameters (Van Landeghem et al., 2010; Van Landeghem & Van Damme, 2007; Van Woensel, 2007). In the UK, Australia and U.S., this gender gap is often studied ethnographically by considering gendered study cultures (e.g., UK: Mac an Ghaill, 1994; Australia: Martino, 1999; U.S.: Pascoe, 2007). Few quantitative studies concerning gender-specific study cultures have been conducted (e.g., Belgium: Derks & Vermeersch, 2001; Van de Gaer et al., 2009; Van Houtte, 2004b; Germany: Heyder & Kessels, 2013) and most of these tend to focus on simple boy-girl differences without delving deeper into the subtle workings of gender. We propose studying the impact of gender cultures through the concept of pressure for gender conformity. So far, pressure for gender conformity has mostly been researched with regard to psychological adjustment in U.S. adolescent samples (e.g., Carver et al., 2003; Corby et al., 2007; Egan & Perry, 2001) and has only rarely been linked to educational parameters. We sought to address this gap by studying the impact of pressure for gender conformity on academic self-efficacy in a Flemish sample of seventh graders. This way, we demonstrate the viability of quantitatively assessing gendered processes in educational gender gap research. Also, by
providing findings from a Flemish sample, possible culture-specific variations in gender and its influence on achievement can be highlighted.

In line with the first hypothesis, our results show that girls scored somewhat higher in both descriptive and regression analyses on self-efficacy for self-regulated learning throughout the school year. This is in line with previous research which found that girls continue to score somewhat higher than boys on self-efficacy for self-regulated learning (meta-analysis: Huang, 2013; U.S.: Pajares, 2002; Belgium: Vantieghem et al., 2014a), even though this is a less gender-biased form of self-efficacy than course-specific efficacy (meta-analysis: Huang, 2013). Furthermore, as could be expected based on the gender situation in Belgium, boys scored consistently higher on perceived gender-conformity pressure throughout the first year of secondary education.

As we expected based on the second hypothesis, our results show that pressure for gender conformity impacts academic self-efficacy in gender-specific ways. Note that it is only by considering the interaction effect between gender and pressure for gender conformity that these effects can be detected. Indeed, the effects for boys and girls are each other’s opposites with boys’ academic self-efficacy being lower when experiencing pressure for gender conformity, whereas girls’ self-efficacy becomes higher. Without considering this interaction, the influence of pressure for gender conformity would probably pass unnoticed, given the effects tend to cancel each other out. This seems to confirm the important and divergent ways gender cultures influence educational parameters in Flanders.

As mentioned above, our results show that when boys experience more pressure, their self-efficacy tends to suffer. For girls, we find quite a different picture. At first glance, pressure for gender conformity does not seem to significantly impact their academic self-efficacy (see Figure 8). However, when taking subjective well-being into account, this association changes and girls’ self-efficacy is higher when experiencing more pressure (see Figure 9). This indicates a suppression effect between pressure for gender conformity and well-being on academic self-efficacy. More specifically, as has often been found in U.S. based samples (e.g., Carver et al., 2003; Egan & Perry, 2001; Younger et al., 2004), pressure for gender conformity takes a negative toll on girls’ well-being. Nevertheless, this gender-conformity pressure influences girls’ academic self-efficacy in a positive way. Because the effect of pressure on well-being tends to nullify the effect of pressure on self-efficacy, this results in girls’ self-efficacy being seemingly unaffected regardless of the amount of pressure they experience (see Figure 8).

These findings illustrate the gender-specific study cultures that many boys and girls experience in Flemish schools. For Flemish boys, this study culture tends
to revolve around opposing school authority, acting tough, and a rejection of femininity (Derks & Vermeersch, 2001). So, when boys experience a lot of pressure from peers and themselves to uphold masculine norms, their educational functioning tends to suffer. These findings are in line with the second hypothesis based on the identity-based motivation model and previous international research about the impact of masculinity cultures on boys’ school behaviors (e.g., UK: Mac an Ghaill, 1994; Australia: Martino, 1999; U.S.: Pascoe, 2007). The gender norms for girls, on the other hand, seem to evoke quite different responses. When we tease out the effect of pressure for gender conformity on well-being versus self-efficacy, it can be seen that girls’ self-efficacy is actually higher when experiencing more pressure. This is in line with international research into gendered study cultures, which has demonstrated that stereotypically feminine traits tend to match those of a good student, as evidenced by characteristics such as being calm, cooperative, compliant, and tidy (UK, U.S. & Australia: Beaman et al., 2006; UK: Jones & Myhill, 2004). This has even led teachers to often perceive girls as ideal students. Similarly for the Belgian context, feminine identities link to higher scores on academic self-efficacy (Vantieghem et al., 2014a) and Belgian teachers’ preference for girls is partly explained by the higher perceived teachability of girls (Van Houtte, 2007). Because people who experience pressure tend to exaggerate their behavior (Weinstein et al., 2012; Willer et al., 2013), girls under pressure could try to emulate the ideal type of the typically feminine girl and inadvertently raise their compliance to educational standards, heightening their academic self-efficacy in the process.

Our finding that gender-conformity pressure links positively with girls’ self-efficacy also adds to the research by Leaper and colleagues (2012). They discovered positive effects of pressure for gender conformity on U.S. girls’ value and ability beliefs for English, whereas it had negative effects on their value and ability beliefs for mathematics. Hence, this research demonstrated the divergent effects of gender-conformity pressures on courses that are either considered to be gender-congruent for girls (such as English in the U.S.-context [Pajares & Valiante, 2001]) or gender-incongruent (such as math [U.S.: Bhanot & Jovanovic, 2005]). Our results add to these findings by demonstrating that such processes extend to more general academic measures, such as self-efficacy for self-regulated learning. Apparently, the study behaviors that are gauged by the measure for self-regulated learning are considered to be rather feminine among Flemish students, resulting in girls’ heightened sense of self-efficacy when experiencing more gender-conformity pressure and vice versa for boys. This ties into previous research that demonstrated that study behaviors, and even the school institute itself, are considered to be
feminine in several countries. (Australia: Martino, 1999; Belgium: Derks & Vermeersch, 2002; Van de Gaer et al., 2006b; UK: Epstein, 1998; Germany: Heyder & Kessels, 2013).

Quite interestingly, the gender-specific relations between gender-conformity pressure and self-efficacy are observed only at the end of the school year. When considering the start of the school year, pressure for gender conformity did not have much impact, let alone provide evidence for gendered study cultures. That these relations can change so much over the course of one school year indicates the importance of this first year in secondary education for the socialization of pupils into the gender norms and study cultures of their respective schools. Hence, 7th grade could be a decisive moment for addressing boys’ educational disadvantage. That is, schools and teachers could try to counterbalance these effects by reducing gender-biases among students. For instance, more attention could go to gender-mixing in classrooms, since gender segregation contributes to gender-stereotyped behavior among students (review: Martin, Fabes, & Hanish, 2014). Similarly, schools could stimulate both teachers (Boysen, 2013) and students (Lamb, Bigler, Liben, & Green, 2009) to confront sexism, since this reduces students’ sexist attitudes.

It might not be surprising, however, that the effect of pressure for gender conformity is only apparent at the end of the school year. That is, gendered study cultures differ between schools in Flanders (Van Houtte, 2004a) and students are more likely to comply to these specific norms the longer they have been exposed to them (Derks & Vermeersch, 2001). Hence, it might take some time for students to familiarize themselves with the new rules of being an adolescent boy or girl at their school. Interestingly, these findings echo the social dosage effect described by Martin and Fabes (2001). They found a stronger impact of gender norms on U.S. preschool children’s behavior the longer they had been exposed to same-sex peers. This social dosage effect with regards to gender-conformity pressures should be examined in other research. For instance, it would be interesting to see whether the impact of pressure for gender conformity continues to grow throughout the years. We could expect, for instance, that the influence of pressure for gender conformity on educational parameters might plateau around mid-adolescence, given that some evidence suggests that gender-conformity pressures tend to be most prevalent in childhood and early adolescence (e.g., Carver et al., 2003; Horn, 2007). Hence, we call for further longitudinal research into the impact of pressure for gender conformity on students’ education.

Overall, it is remarkable that such gender-specific effects can be discovered among Flemish students, despite Belgium’s relative gender equality. This implies
that aspects of gender-typing remain prevalent in gender-egalitarian nations, despite high scores on formal gender equality indexes. Additionally, this study suggests that peer relations in early adolescence, and the conformity-pressures that go with it, seem to be of central importance in transmitting and enforcing gender-normative behavior. By researching similar patterns in samples from different contexts, the cultural particularities or transcendence of these processes could be assessed. It could, for instance, be hypothesized that effects might be more extreme in less gender equal environments. We also propose replicating the study with other Flemish samples. This follows from the successive sampling method used in this research, where schools were selected before students were approached. Because many schools declined to participate in the research, a participation rate of 47.6% was achieved. Even though analyses showed that the schools in the sample did not differ substantially from other Flemish schools in school sector, curriculum or student composition, it would strengthen the research findings if they could be replicated in other research with Flemish samples.

**CONCLUSION**

Our research investigated gender cultures at Belgian schools from an individual, quantitative and gendered perspective by employing the concept of pressure for gender conformity. The results show that seventh-grade boys who experienced more pressure for gender conformity tended to have a lower sense of academic self-efficacy. On the other hand, girls’ academic self-efficacy did not suffer despite enduring similar levels of pressure. More specifically, when controlling for well-being, girls’ self-efficacy even increased when they experienced more gender-conformity pressure. Hence, pressure for gender conformity had opposite effects on boys and girls. This is illustrative of the gender-specific study cultures at Flemish schools, with boys’ peer groups displaying more anti-school attitudes than girls’ peer groups. In our study, the relation between pressure for gender conformity and academic self-efficacy appears in the course of the first year of secondary education. This is indicative of the social dosage effect and suggests that the first year of secondary education is an important time phase in socializing students within the gender cultures of their schools.
Chapter 10.
The impact of gender divergence on adolescents’ well-being:  
Does the school context matter?

Wendelien Vantieghem, & Mieke Van Houtte

Schools are often heteronormative environments that push toward traditional gender binaries and heterosexuality. This negatively influences the well-being and academic functioning of LGB-students. Less is known about the impact on gender-divergent adolescents. Hence, this study investigates the well-being of gender-divergent students — as evidenced by their gender (a)typicality and pressure for gender conformity — in different school cultures — as evidenced by the level of traditional gender ideology and homonegativity —. We examine these context-dependent associations in a sample of 4987 Flemish students, who were questioned three times in the course of 2012-2014. Results from multilevel regression analyses show that both gender typical and atypical adolescents feel worse in heteronormative than liberal schools. Furthermore, the social acceptance of boys seems to be more conditional on displaying gender-conforming behavior in heteronormative than liberal schools. Overall, the results suggest that liberal schools ensure higher well-being for all students (regardless of gender-divergence or sex) and that student-teacher relationships seem to be central in explaining these school differences. Implications for school policy, as well as research considering heteronormativity and gender divergence are discussed.
GENDER DIVERGENCE AND ADOLESCENTS’ WELL-BEING: DOES THE SCHOOL CONTEXT MATTER?

Schools are often heteronormative environments where heterosexuality and traditional gender differences are assumed and reinforced through both formal (e.g., curriculum and policies) and informal school features (e.g., rituals, daily interactions and vocabulary) (Chesir-Teran, 2003; Toomey et al., 2012; Wilkinson & Pearson, 2009). Because of this, they can be unwelcoming surroundings to students who defy traditional gender norms, with recent studies uncovering the impact of such school contexts on same-sex attracted youth (e.g., Aerts, Dewaele, Cox, & Van Houtte, 2014; Goodenow, Szalacha, & Westheimer, 2006; Murdock & Bolch, 2005; Toomey & Russell, 2013; Wilkinson & Pearson, 2009). For instance, lesbian, gay and bisexual (LGB) students experience more interpersonal problems with peers and teachers than straight students (Ueno, 2005), and report high prevalence of school-based victimization (Goodenow et al., 2006; Poteat, Scheer, & Merieish, 2014a; Toomey et al., 2012). Consequently, LGB youth tend to feel worse (Poteat et al., 2014a; Sandfort, Bos, Collier, & Metselaar, 2010; Ueno, 2005; Wilkinson & Pearson, 2009), have a reduced sense of school belonging (Murdock & Bolch, 2005; Poteat et al., 2014a), and worse academic achievement than their straight counterparts (Aerts, Van Houtte, Dewaele, Cox, & Vincze, 2011; Murdock & Bolch, 2005; Poteat et al., 2014a; Wilkinson & Pearson, 2009).

Nevertheless, same-sex attracted youth are not the only ones affected by heteronormativity at school. Studies have shown that heterosexual adolescents, especially boys, are also subject to homophobic name-calling, which influences their well-being in adverse ways (Collier, Bos, & Sandfort, 2013; Poteat, Scheer, DiGiovanni, & Merieish, 2014b). It is theorized that such homophobic bullying is used as a way to enforce gender conformity and sanction gender role transgressions among adolescents, regardless of sexual orientation. However, beyond homonegative bullying, few studies have considered the impact of heteronormative school environments on heterosexual students or those who defy gender norms (for an exception, see Toomey et al., 2012). Usually, it is assumed that patterns found among LGB-students could be extrapolated to gender-theradent students (e.g., Poteat et al., 2014a). Nevertheless, there is a lack of empirical evidence to support this notion.

Another issue with previous research is that most studies have used individual-level experiences (such as victimization) as indicators of schools’ heteronormativity (e.g., Birkett, Espelage, & Koenig, 2009; Murdock & Bolch, 2005; Sandfort et al., 2010). However, to properly assess and understand contextual
influences, and hence, for a thorough understanding of heteronormativity in schools, multi-level analyses are necessary (Hox, 2010). Consequently, this study considers in a multi-level fashion the associations between Flemish schools’ heteronormative cultures and early adolescents’ well-being. We pay special attention to the context-dependent influence of gender divergence, as measured by gender typicality and pressure for gender conformity, on students’ well-being.

HETERONORMATIVITY IN SCHOOLS

Heteronormativity refers to the institutional organization of heterosexual privilege (Chesir-Teran, 2003; Toomey et al., 2012; Wilkinson & Pearson, 2009). It builds upon and enforces an underlying binary of gender and sexuality that is presumed to be natural. Hence, heteronormativity functions by defining and enforcing ideas about what is “normal” in everyday life, and by pathologizing that which is not. This involves highlighting the differences between men and women, sanctioning gender transgressions that disrupt this binary, and celebrating heterosexual relationships. As such, heterosexuality and the social construction of gender differences are interwoven, and people who defy one of these norms are often assumed to equally violate the other norm (Collier et al., 2013; Rieger, Linsenmeier, Gygax, Garcia, & Bailey, 2010). That is, those who act in gender non-conforming ways are often assumed to be attracted to the same sex, whereas lesbians and gays are often assumed to behave in “butch” or effeminate ways. Consequently, “gay” or “fag” are used as derogatory terms for boys who behave in non-masculine ways, rather than to indicate boys with same-sex attractions (Pascoe, 2007). Similarly, the word “dyke” is assigned to girls who do not conform to traditional notions of femininity (Nielsen, Walden, & Kunkel, 2000). These processes stigmatize and marginalize both same-sex attraction and gender non-conformity, resulting in people “doing gender” in ways that avoid such social sanctions.

Heteronormativity is evident in several institutions, including schools. Most schools evince celebrations of heterosexuality and the gender binary in both formal and informal ways (Chesir-Teran, 2003; Hong & Garbarino, 2012). For instance, heteronormativity is often evident in formal aspects of schooling by omitting gender and LGB-issues from anti-bullying policies or the curriculum. Informal aspects encompass daily interactions or the general social sphere, and include, for instance, the amount of people who are “out” at school, the use of LGB-slurs by students and teachers, the requirement for girls to wear dresses at prom, and overall gender-related attitudes of peers and teachers.
Since most research concerning heteronormativity in schools has considered the impact on LGB students, the focus has predominantly been on LGB-specific aspects of heteronormativity. For instance, studies focused on the impact of LGB-inclusive curricula (e.g., Chesir-Teran & Hughes, 2009), the presence of gay-straight alliance clubs at school (e.g., Craig & Smith, 2014; Goodenow et al., 2006; Szalacha, 2003; Toomey & Russell, 2013), whether LGB issues where explicitly mentioned in anti-bullying policies (e.g., Chesir-Teran & Hughes, 2009; Goodenow et al., 2006; Szalacha, 2003) or whether teachers were supportive and/or received training regarding LGB issues (e.g., Goodenow et al., 2006; Murdock & Bolch, 2005; Szalacha, 2003). Nevertheless, research has shown that heteronormativity is present and influential in more than just LGB-focused school policies. Wilkinson and Pearson (2009) showed, for instance, the effect of football-participation and religiousness on LGB’s wellbeing. Similarly, Sandfort and colleagues (2010) investigated whether LGB’s mental health was affected by respect for diversity. However, the focus on school policy as an indicator of heteronormativity has sometimes led to contra-intuitive results. For instance, Toomey and colleagues (2012) reported a reduced perception of safety for LGB and gender non-conforming students in schools with more inclusive policies. A possible explanation for this finding might be that these inclusive policies were implemented as a reaction to severe expressions of intolerance to LGB and gender non-conforming students; but this could not be confirmed or denied in this study. Hence, the strong focus on school policies as sole indicator for heteronormativity might obscure rather than expose the patterns under study.

This study aims to supplement previous research by investigating heteronormativity by other indicators than school policy. More specifically, we follow the suggestion of Chesir-Teran (2003), who identifies several aspects of heteronormativity at school. Chesir-Teran (2003) hypothesizes that heteronormativity is not only present in a school’s program or policy, but also in suprapersonal features that reflect average characteristics of school members. This includes demographic features such as mean levels of students’ ability or teacher training, as well as typical attitudes and beliefs of school members. In keeping with school-effects research, we will dub these shared attitudes and beliefs of school members “school culture”, while the demographic features are considered separately as characteristics of school composition (Van Houtte, 2005). Previous research in the U.S. has considered school compositions and has shown, for instance, that schools with large (Goodenow et al., 2006; Szalacha, 2003) and ethnically diverse student bodies tend to be more welcoming to LGB students (Goodenow et al., 2006; Poteat, Espelage, & Koenig, 2009). Other indicators of
school composition have given more mixed results, with some studies finding that schools with a high percentage of girls (Drury, Bukowski, Velásquez, & Stella-Lopez, 2013) or students of lower socio-economic background tend to be more heteronormative (Szalacha, 2003), whereas others find that such schools tend to provide safer environments for LGB or non-conforming students (Drury et al., 2013; Goodenow et al., 2006). Conversely, research into school cultures on heteronormativity has been more limited. This is probably caused by most research’s focus on individual perceptions, without considering the possible impact of shared beliefs (e.g., Murdock & Bolch, 2005; Poteat et al., 2009; Szalacha, 2003). Consequently, little is known about the impact of several indicators of school culture on the day-to-day experiences of LGB and gender-divergent adolescents. Nevertheless, shared attitudes among students could better approximate the actual living environment than school policies. Indeed, research has shown that many students are often unaware of inclusive policies at their schools, and that there would be considerable variance in the extent to which such policies are enforced (Chesir-Teran & Hughes, 2009; Szalacha, 2003).

THE SCHOOL EXPERIENCES OF GENDER-DIVERGENT ADOLESCENTS

It has been well established that LGB-students experience higher rates of school-based victimization than heterosexual students (Goodenow et al., 2006; Poteat et al., 2014a; Toomey et al., 2012), as well as reduced levels of well-being (Ueno, 2005; Wilkinson & Pearson, 2009) and academic achievement (Aerts et al., 2011; Murdock & Bolch, 2005). Even though the research into the school experiences of gender-divergent adolescents is more limited, the results do suggest similar patterns. For instance, a wealth of qualitative studies has shown that boys who do not conform to traditional notions of masculinity tend to be teased and bullied (Epstein, 1998; Renold, 2004; Stoudt, 2006; Swain, 2005). Quantitative studies have also considered the well-being of gender-divergent adolescents in schools. In recent years, a lot of these have built upon the multidimensional gender model by Egan & Perry (2001) and have paid particular attention to two factors of the model: gender typicality and pressure for gender conformity. Self-perceived gender typicality refers to how similar one perceives oneself to be to other gender group members, whereas felt pressure for gender conformity reflects the degree to which one experiences strong expectations to uphold gender-role norms (Egan & Perry, 2001). Adolescents who experience a lot of gender-conformity pressure tend to report lower psychological adjustment (Carver et al., 2003; Egan & Perry, 2001;
Menon, 2011). Similarly, children and early adolescents who identify as gender atypical, have been shown to report lower psychological adjustment (Carver et al., 2003; Egan & Perry, 2001; Menon, 2011), as well as reduced rates of academic functioning (Vantieghem & Van Houtte, 2015b; Vantieghem et al., 2014a).

Importantly, research suggests that social support and acceptance are key factors in these processes for both same-sex attracted and gender-divergent adolescents. That is, peer and teacher support have both been shown to connect to higher rates of school belonging among LGB teens (Murdock & Bolch, 2005; Toomey & Russell, 2013), as well as lower rates of victimization and suicidality (Goodenow et al., 2006; Hansen, 2007). Furthermore, Smith & Leaper (2006) demonstrated that gender-divergent adolescents (as evidenced by feeling atypical and reporting high gender-conformity pressures) who felt accepted by their peers, reported similar levels of psychological adjustment as other adolescents. This suggests that acceptance of non-conformity in the social environment could be key in accounting for the well-being of gender divergent adolescents.

Notably, boys tend to experience more gender-conformity pressure, and also report higher levels of gender typicality (Egan & Perry, 2001). These stronger findings for boys are usually framed within the stricter societal gender norms for boys and men (Herek, 2002; Ueno & McWilliams, 2010; Vantieghem & Van Houtte, 2015a; Young & Sweeting, 2004). That is, whereas stereotypically masculine behavior (such as going out to work, being ambitious and successful) has become more acceptable for women in the last decades, men have not experienced a similar broadening of appropriate role behavior. Consequently, boys have to navigate stricter gender expectations than girls, and are socially sanctioned more readily and severely for transgressions (Young & Sweeting, 2004). These processes are also reflected in research findings which demonstrate that gay boys are usually bullied more heavily than lesbian girls (Herek, 2002).

Since a lot of bullying takes place at school (Schneider, O'Donnell, Stueve, & Coulter, 2012), and school is one of the important socializing and living environments for adolescents (Eccles, 2007), it seems pertinent to take matters of the school context into account. Nevertheless, few researchers have investigated the well-being of gender divergent students while taking school characteristics into account. Of those who have, findings indicate that both gender typical and atypical students reported lower levels of attachment and more depression in schools with more traditional gender cultures (Reynolds, 2014). Furthermore, Drury and colleagues (2013) showed that gender atypical girls experienced more victimization in school environments with higher gender-conformity pressures. Toomey and colleagues (2012) showed that school strategies that question heteronormativity
(such as gay-straight alliances and inclusive curricula) helped to increase the perceived school safety for non-conforming students.

Overall, these studies suggest that the context can have an important influence on the acceptance of gender non-conformity, and that this in turn has an important impact on students’ well-being. Hence, we could expect that the well-being of non-normative students would be lower in schools with higher levels of heteronormativity. Furthermore, we presume that these associations might be stronger among boys than girls, since gender norms are stricter for boys and they are sanctioned more severely for gender transgressions.

THE CURRENT STUDY

In the current study, we investigate how a school’s level of heteronormativity influences the well-being of gender-divergent early adolescents. We will expand on each part of this research question below.

First, a school’s level of heteronormativity is assessed by considering two indicators of school culture: traditional gender role ideology and homonegativity. Hence, we will consider the extent to which a school culture can be defined as supporting traditional male-female gender role divisions, as well as the level of homonegativity in schools (while controlling for aspects of school composition). This way, we tap both the press towards a traditional gender binary and towards heterosexuality theorized to underlie heteronormativity (Chesir-Teran, 2003; Wilkinson & Pearson, 2009). We will do this by separately aggregating students’ attitudes concerning homonegativity and traditional gender role ideology per school (see Measures). In contrast to previous research’s focus on school policies and individual-level experiences, we add to the study of heteronormativity by considering the effect of suprapersonal factors (as suggested by Chesir-Teran, 2003) on individuals’ well-being in a multi-level fashion. We expect that the associations between gender divergence and well-being will be stronger when schools display higher levels of heteronormativity.

Second, we consider gender divergence among adolescents by using two dimensions from the gender identity model proposed by Egan and Perry (Egan & Perry, 2001). More specifically, the context-dependent effects of felt pressure for gender conformity and self-perceived gender typicality on well-being will be considered. We expect that the associations will be stronger among boys, since gender norms tend to be more strict for boys than girls. Furthermore, in line with previous research (Goodenow et al., 2006; Murdock & Bolch, 2005; Sikora, 2014),
we expect that positive relationships with peers and teachers could buffer the adverse consequences of heteronormative contexts on adolescents’ well-being.

Third, these associations will be assessed in a sample of early adolescents. Since gender norms tend to become more stringent around this age (Galambos et al., 1990), adolescents attending middle school make a good test sample for the proposed associations. Last, next to the use of multi-level analyses to correctly account for the context and the nested structure of the data, data from three time-points will be employed to better judge the robustness of these associations throughout middle school.

**METHOD**

**Participants and Procedure**
The data are part of the “Teaching in the Bed of Procrustes”-study and were gathered in three waves in the school years 2012-14 (at the beginning and at the end of seventh grade, and once more at the end of the eight grade). Schools were selected so each geographical region within Flanders was equally represented. Selection was further dependent on school denomination (public versus private) and a proportional representation of rural versus city schools. Within these parameters, three random samples were drawn. For each school that refused, a matched school from the next random sample was contacted. This way, we aimed to obtain schools that represent the Flemish educational context on important parameters (such as region, school denomination, and location), while maintaining randomness within these subpopulations. Of the 124 contacted schools, 59 schools participated in the study, which translated to a response rate of 47.6%. This response rate is less than 50% due to schools in Flanders being swamped with research requests. Consequently, schools accept research requests on a first-come-first-serve basis. Analyses in which we compared our sample to the Flemish school population, based on information from the Flemish Educational Department, showed no important differences in school sector, curriculum, or student composition. This suggests that no systematic biases occurred and that the schools were representative of the population. Two schools dropped out in the course of the study. Cooperation with one school was terminated because of problems with upholding appointments, while another school had ceased to exist. Hence, the end sample consisted of 57 schools. The variation between sample schools in composition and school culture can be assessed in Table 20.
Table 20.

Descriptive statistics on the school level.

<table>
<thead>
<tr>
<th>Variables (N = 58)</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional gender ideology</td>
<td>1.737</td>
<td>.215</td>
<td>1.4</td>
<td>2.19</td>
</tr>
<tr>
<td>Homonegativity</td>
<td>1.191</td>
<td>.372</td>
<td>.60</td>
<td>2.43</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School size</td>
<td>126.83</td>
<td>81.83</td>
<td>22</td>
<td>338</td>
</tr>
<tr>
<td>Gender composition (% of girls)</td>
<td>.432</td>
<td>.212</td>
<td>.00</td>
<td>.81</td>
</tr>
<tr>
<td>Ethnic composition (% of ethnic minority)</td>
<td>.251</td>
<td>.257</td>
<td>.01</td>
<td>.95</td>
</tr>
<tr>
<td>SES composition (% of students with blue-collar background)</td>
<td>.29</td>
<td>.192</td>
<td>.01</td>
<td>.87</td>
</tr>
</tbody>
</table>

Within these schools, all seventh graders (eight graders by wave 3) were asked to complete the questionnaire. The use of child assent was approved by the school and the Belgian Commission for the Protection of Privacy, based on the minimal risk of the study. Researchers were always present during the completion of the survey to explain the procedure and answer questions. Pupils were told that the survey was not a test and were assured that it was completely confidential. The response rate among the students in the first wave was 96.9%, translating to a total of 6380 students. By connecting the data, it showed that 4987 students had participated in all three waves. This reduction in the number of cases is due to the disappearance of two schools from the study, students not being present at school during one of the survey-moments (for instance, due to sickness), or because students transferred to a different school in the course of the study.

Of the students who participated in all three waves, the majority were boys (52.5%) and of Belgian or western-European descent (81.9%). In line with the migration history of Belgium, children of non-western European descent (15.4%) had predominantly Turkish, Moroccan or Eastern-European roots. The socioeconomic status (SES) of the students was based on parents’ occupation. In this sample, 18.5% of the students had a blue-collar background (e.g., parents in a manual labor position), while 79.6% had a white collar background (e.g., parents in service or management positions). In the Flemish educational context, there are officially two tracks in the first grade: the general and the vocational-oriented track (Boone & Van Houtte, 2012; Van Houtte et al., 2012). In our sample, 12.4% of pupils were in the vocational track by the end of eight grade. At the start of the study (e.g., the start of seventh grade), the mean age was 12 years old. There was a
range from 10 to 15 years old, indicating that some pupils skipped a year, whereas others had repeated years. Nevertheless, 80.4% of the sample was 12 years old, demonstrating that the majority were on track. The cronbach alpha’s and variable scores of the study participants in all three waves can be seen in Table 21.

**Measures**

*School-level variables.* Homonegativity was assessed using the scale by Roggemans (2013) (for an elaborate discussion of this scale, see individual-level variables). To assess the extent to which these attitudes were shared at the school-level, we used the Spearman-Brown formula based on the intra-class correlation of a one-way analysis of variance: \( ICC(1, k) = \frac{(BMS - WMS)}{BMS} \) (with \( k \) = number of raters in each school) (Shrout & Fleiss, 1979). Higher results indicate that beliefs are more similar within schools than between schools, implying that each school has a specific culture. In order to allow aggregation at the school level by calculating a mean sum of the scores of the individual members, the result must be a minimum of .60. The result for homonegativity was .94, indicating that homonegative attitudes were clearly shared by students within schools. Furthermore, a repeated measurement ANOVA showed that this aggregated score remained invariant throughout the waves (\( p = .183 \)). Because of the constant and shared character of these beliefs at the school level, it is accurate to speak of more or less homonegative school cultures. The most progressive school scored .6 on this scale, whereas the most homonegative school scored 2.43.
Table 21.
Descriptive statistics on the individual level.

<table>
<thead>
<tr>
<th>Individual-level Variables</th>
<th>Wave 1</th>
<th></th>
<th></th>
<th>Wave 2</th>
<th></th>
<th></th>
<th>Wave 3</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cronbach's Alpha</td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
<td>Cronbach's Alpha</td>
<td>Mean</td>
<td>SD</td>
<td>N</td>
<td>Cronbach's Alpha</td>
</tr>
<tr>
<td>Well-being</td>
<td>.81</td>
<td>2.87</td>
<td>.50</td>
<td>4953</td>
<td>.86</td>
<td>2.88</td>
<td>.57</td>
<td>4927</td>
<td>.86</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.77</td>
<td>2.48</td>
<td>.67</td>
<td>4930</td>
<td>.86</td>
<td>2.73</td>
<td>.69</td>
<td>4898</td>
<td>.84</td>
</tr>
<tr>
<td>Gender-conformity pressure</td>
<td>.77</td>
<td>1.46</td>
<td>.56</td>
<td>4909</td>
<td>.81</td>
<td>1.48</td>
<td>.55</td>
<td>4814</td>
<td>.82</td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trust in teachers</td>
<td>.91</td>
<td>2.83</td>
<td>.48</td>
<td>4304</td>
<td>.86</td>
<td>2.64</td>
<td>.50</td>
<td>4955</td>
<td>.92</td>
</tr>
<tr>
<td>Equal treatment by teachers</td>
<td>.87</td>
<td>3.29</td>
<td>.71</td>
<td>4559</td>
<td>.89</td>
<td>3.15</td>
<td>.76</td>
<td>4950</td>
<td>.92</td>
</tr>
<tr>
<td>Homonegativity</td>
<td>.90</td>
<td>1.04</td>
<td>.78</td>
<td>4941</td>
<td>.91</td>
<td>1.04</td>
<td>.81</td>
<td>4914</td>
<td></td>
</tr>
<tr>
<td>Traditional genderrole ideology</td>
<td>.80</td>
<td>1.61</td>
<td>.55</td>
<td>4960</td>
<td>.84</td>
<td>1.63</td>
<td>.59</td>
<td>4956</td>
<td>.85</td>
</tr>
<tr>
<td>Sexual orientation:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterosexual</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>96.0</td>
</tr>
<tr>
<td>Lesbian-gay-bisexual (LGB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.3</td>
</tr>
<tr>
<td>Questioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.7</td>
</tr>
<tr>
<td>Social integration:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16.4</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19.2</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>19.3</td>
</tr>
<tr>
<td>Improved</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24.2</td>
</tr>
</tbody>
</table>
School culture of traditional gender ideology. To measure traditional gender role attitudes, we used the scale developed by Vermeersch and colleagues (2010) (for an elaborate discussion of this scale, see individual-level variables). Similar to the measure of homonegative school culture (see above), we employed the Spearman-Brown formula to assess the extent to which traditional gender attitudes were shared at the school level. The result was .92, indicating that there were more differences in these attitudes between schools than within schools. A repeated measurement ANOVA indicated that the first wave differed somewhat from the last two waves ($p = .026$), while the last two waves did not differ significantly from each other. Since the first wave was gathered in the first few months of 7th grade, and students still had to familiarize themselves with their new school, this did not seem to indicate problematic variance in the school culture. The most progressive school scored 1.4 on this scale, whereas the most traditional school scored 2.19.

Individual-level variables. Subjective well-being. Subjective well-being was measured by asking students how often they had felt certain emotions in the past 30 days. They indicated the frequency of positive and negative feelings on a 5-point scale ranging from 0 = never to 4 = very often (Keyes et al., 2002). The scale consisted of 12 items, equally divided between positive (e.g., happy, peaceful) and negative items (e.g., nervous, worthless). Based on principal component analysis, we constructed a single scale from these items. The items pertaining to negative affect were reverse coded so higher scores on the scale indicate a higher well-being. This scale was constructed using a mean sum of scores and displayed good internal reliability on all time-points (see Table 21).

Gender typicality. To measure gender typicality, we employed the gender typicality subscale from the gender-identity Questionnaire by Egan and Perry (2001). This measure consists of six items with a five-point likert scale (0 = completely disagree, 4 = completely agree) tapping the extent to which one feels typical of one’s gender (e.g., “I feel that my skills and interests are similar to those of other [girls/boys]”, “I feel that I am a good example of a typical [boy/girl]”). The gender typicality score was computed by using the mean sum of scores. The scale displayed good internal reliability on all time-points (see Table 21).

Pressure for gender conformity. To measure pressure for gender conformity, we used the pressure for gender conformity subscale from the gender-identity questionnaire by Egan and Perry (2001). This subscale measures the extent
to which one experiences pressure for gender-conforming behavior, with four items related to pressure from peers and four items related to pressure from oneself (e.g., “I think it is important to act just like other [girls/boys]”, “My friends would be upset if I wanted to play with [boys'/girls’] toys”). The scale consists of separate versions for boys and girls. The Dutch translation and answering format with a 4-point likert scale (1 = completely disagree, 2 = disagree, 3 = agree, 4 = completely agree) are based on the work of Bos and Sandfort (2010). This scale was constructed using mean sum of scores, with high scores indicating more pressure for gender conformity. The scale displayed good internal reliability on all time-points (see Table 21).

**Social integration at school.** To assess social integration, we used a scale developed by Van Damme and colleagues (Van Damme, De Fraine, Van Landeghem, Opdenakker, & Onghena, 2002). The scale consisted of 10 items tapping the extent to which participants were socially integrated at school (e.g., “I have few friends at school”, “Most students of this school treat me well”). The scale employed a 5-point likert scale (0 = completely disagree, 4 = completely agree). Negative items were reverse coded so higher scores indicated higher integration and the scale was constructed using mean sum of scores. The scale displayed good internal reliability on all time-points. Because this scale was not part of the survey on Time 1, we constructed a categorical variable out of the responses on Time 2 and 3. Students who were among the least socially integrated on both time-points (by being part of the 33% lowest scoring), were coded 0 = low. Students who were medium integrated on both time-points (by being in the middle 33%), were coded 1 = medium. Students who scored high on both time-points, were coded 2 = high. Students who reported an improved sense of social integration on Time 3 versus 2 (by moving from the lowest or medium category to a higher group), were coded 3 = improved. There were no students who reported a deteriorated sense of social integration on Time 3 versus 2. We used dummy-coding for this variable, with low as reference category.

**Trust in teachers.** To assess the extent to which students trusted their teachers, we used a scale developed by Forsyth and colleagues (Forsyth, Adams, & Hoy, 2011). The scale had 13 items tapping the extent to which students trusted their teachers to help them out and do a good job (e.g., “It is easy to talk to teachers at this school”, “Teachers at this school are always willing to help”). The scale employed a 5-point likert scale (0 = completely disagree, 4 = completely agree).
scale was constructed using mean sum of scores and displayed good internal reliability on all time-points (see Table 21).

**Equal treatment by teachers.** To assess the extent to which students felt treated fairly by their teachers, we employed a subscale from the “What Is Happening In This Class”-questionnaire (WIHIC) (Dorman, 2003). This subscale consisted of 6 items tapping the extent to which students felt discriminated against by their teachers (e.g., “I get as much help as other students from my teachers”, “I am treated the same as other students in the class”). The scale used a 5-point likert scale (0 = almost never, 4 = almost always) and items were coded so higher scores indicating feeling less discriminated against. The scale was constructed using mean sum of scores and displayed good internal reliability on all time-points (see Table 21).

**Age.** Students indicated their age at the time of the first questionnaire. The mean age in this sample was 12 years old, with a range from 10 to 15 years old.

**Sexual orientation.** To assess the sexual orientation of the participants, they were asked who they usually fell in love with. Answer categories included 1 = boys, 2 = girls, 3 = boys and girls, 4 = I don’t know. Students who indicated they fell in love exclusively with people from the opposite gender were coded as 0 = heterosexual. Students who indicated to fall in love with people from the same gender were coded as 1 = lesbian, gay and bisexual (LGB). Students who indicated they did not know who they fell in love with, received the code 2 = questioning. We used dummy-coding for this variable, with heterosexual as reference category.

**Individual-level homonegativity.** Homonegativity was measured by assessing students’ agreement with 11 items pertaining to acceptance and rights for LGB’s (e.g., “If I discover that a friend of mine is gay, our friendship is over”, “Gay marriage should be illegal”) (Roggemans, 2013). The scale had a five-point likert scale (0 = completely disagree, 4 = completely agree) and was constructed using mean sum of scores, with high scores indicating more homonegative attitudes. Because the scale was only assessed on Time 2 and 3, we employed the scale on the individual rather than the time-level. Answers on Time 2 were considered baseline levels of homonegativity, and we constructed changes in homonegative attitudes by deducting Time 2 responses from Time 3.
**Individual-level traditional gender ideology.** To measure traditional gender role attitudes, we used the scale developed by Vermeersch and colleagues (2010). The scale consisted of 15 items (e.g., “It is best if the husband makes the decisions in a family”, “It is normal that girls pay more attention to their looks than boys”) and employed a 5-point likert scale ($0 = \text{completely disagree}, 4 = \text{completely agree}$). The scale was constructed using mean sum of scores, with high scores indicating more traditional attitudes. The scale displayed good internal reliability on all time-points (see Table 21).

**Data-Analytic Plan**
Descriptive analyses on the school level were used to assess the way compositional and cultural school indicators co-varied. Hence, these analyses give us an insight into the compositional characteristics of more or less heteronormative schools.

To account for the multilevel and longitudinal nature of the data, we used three-level multilevel analyses, using the statistical program MLwiN. The first level reflects the time-variant variables (that is, variables that have been assessed in each wave), the second level reflects individual-level variables (that is, student characteristics that did not change from one wave to another, such as gender or SES), and the third level constituted school variables. In keeping with the longitudinal nature of the data, all models included time-level controls for survey-waves (although coefficients are not shown in the tables).

The main dependent of the study was students’ well-being, and we considered the influence of gender typicality and pressure for gender conformity on this variable. To assess the extent to which school characteristics influence these associations, we estimated cross-level interactions. Hence, the first model includes the main variables and a cross-level interaction between gender typicality and homonegative school culture. In the second model, we included the control variables to check whether the associations remained, diminished or disappeared after controlling for social integration, trust in teachers, equal treatment by teachers age, sexual orientation, SES and ethnic background. To ensure that the school culture effects are no remnants of personal attitudes, we control for individual-level homonegativity when considering homonegative school cultures and for individual-level traditional gender ideology when considering traditional school cultures. Similarly, when considering cross-level interactions with gender typicality, we control for gender-conformity pressure and vice versa. Next to these individual-level controls, we also control for school characteristics such as school size and SES composition (see “descriptive school-level analyses” for the reasons underlying this decision).
This pattern was repeated for all possible combinations between the gender-divergence variables (e.g., gender typicality and pressure for gender conformity) and the school heteronormativity variables (e.g., homonegativity and traditional gender attitudes). Consequently, model 3 assessed the mains and cross-level interaction between gender typicality and traditional school culture. Model 5 considered the cross-level interaction between gender-conformity pressure and homonegativity. Model 7 estimated the cross-level interaction between gender-conformity pressure and traditionality. All appropriate controls (age, sexual orientation, SES, ethnic background, social integration, trust in teachers, equal treatment by teachers, gender typicality/gender-conformity pressure, homonegativity/gender ideology, school size and school’s SES composition) were added in models 4, 6 and 8.

Since we expected different effects for the sexes — and to avoid complex three-way interactions — all analyses were performed separately for boys and girls. Continuous variables were grand mean centered to improve interpretability of the intercept (Hox, 2010), with the exception of the gender-divergence variables. Gender typicality and gender-conformity pressure were school-mean centered, since previous research (Drury et al., 2013; Saxvik & Joireman, 2005) and preliminary analyses showed that students rated themselves as somewhat less gender typical in heteronormative schools and that average levels of gender-conformity pressures tended to be higher in these schools. Hence, absolute scores on these scales would be less comparable across schools, since scoring 3 on gender-conformity pressure in a progressive school would be considered high, whereas this might be considered average in a heteronormative school.

Variance components were random for the intercept, individual and time error term. To allow random slope variance and to correctly assess the cross-level interactions, variance components were also random for gender typicality in models 1 through 4, and for gender-conformity pressure in models 5 through 8. Variance components for other variables were kept fixed to improve statistical power (Hox, 2010).

**RESULTS**

**Descriptive School-level Analyses**

The zero-order correlations in Table 22 show that the associations between the variables were similar for homonegative and traditional school cultures. This implies that, in line with the theoretical notion that heteronormativity refers to traditional notions on sexuality and gender (Chesir-Teran, 2003), the patterns can be considered together as indicators of heteronormative school cultures. The results
show that schools tended to be more heteronormative when they had a lower proportion of female students, a higher proportion of blue-collar students and students with ethnic minority background, and had fewer students attending. These associations were somewhat more pronounced among homonegative school cultures than among traditional school cultures for gender composition, Steiger's $Z = -3.294, p < .001$, and ethnic composition, $Z = -4.129, p < .001$, but not for school size, $Z = -.763, p = .445$, or SES composition, $Z = -.822, p = .411$. Because of the moderate to high correlations between the compositional and culture indicators, as well as the limited number of cases on the school-level, further analyses only take the effects of school size and SES composition into account.

Table 22.

Zero-order correlations on the school-level

<table>
<thead>
<tr>
<th>Variables</th>
<th>Traditionality</th>
<th>Homonegativity</th>
<th>School size</th>
<th>Gender composition</th>
<th>Ethnic composition</th>
<th>SES composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditionality</td>
<td>—</td>
<td>.869***</td>
<td>-.33*</td>
<td>-.590***</td>
<td>.553***</td>
<td>.649***</td>
</tr>
<tr>
<td>Homonegativity</td>
<td>—</td>
<td>-.280*</td>
<td>-.395**</td>
<td>.759***</td>
<td>.690***</td>
<td></td>
</tr>
<tr>
<td><strong>Control variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School size</td>
<td>—</td>
<td></td>
<td>.110</td>
<td>-.201</td>
<td>-.395**</td>
<td></td>
</tr>
<tr>
<td>Gender composition</td>
<td>—</td>
<td></td>
<td>—</td>
<td>-.024</td>
<td>-.182</td>
<td></td>
</tr>
<tr>
<td>Ethnic composition</td>
<td></td>
<td></td>
<td>—</td>
<td></td>
<td>.663***</td>
<td></td>
</tr>
<tr>
<td>SES composition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>—</td>
</tr>
</tbody>
</table>

Note. N: 58. * p < .05. ** p < .01. *** p < .001.

**Multilevel Analyses**

The zero-model from the multilevel analysis showed that well-being varied somewhat between schools. For girls, 3.97% ($e_{ij} = .188, u_{ij} = .175, v_{jk} = .015$) of total variance was situated at the school-level, whereas for boys this was only 1.63% ($e_{ijk} = .160, u_{ijk} = .142, v_{jk} = .005$). For both boys and girls, about half of unexplained variance was located at the time-level (girls = 49.74%, boys = 52.12%), indicating that students well-being changed in the course of the study.
Context-dependent associations between gender typicality and well-being. As can be seen in model 1 of Table 23, gender typicality related positively to well-being, whereas a homonegative school culture related negatively to well-being. This was true for both boys and girls, indicating that students who identify as gender typical scored higher on well-being, whereas schools that had more homonegative cultures took a negative toll on students’ well-being. Similarly, the cross-level interaction between typicality and homonegative school culture was significant for both boys and girls, indicating that the well-being of (a)typical students varied according to school’s level of homonegativity. Notably, this effect was negative and similar in size for both genders ($\gamma_{\text{boys}} = -.093, p < .01; \gamma_{\text{girls}} = -.096, p < .01$). Since the effects were very comparable for both genders, the calculations for boys discussed below can be extrapolated to girls.

To assess the effects of the cross-level interactions, we calculate and compare the effects on the extreme endpoints of the scales. That is, we compare the effects on the well-being of students who scored lowest on the typicality scale (that is, scoring zero) in schools that were either the most progressive (hence, scoring .597 on the homonegativity scale) versus in schools that were the most homonegative (scoring 2.43 on the homonegativity scale) (see Table 20). To do this, we use the formula: $y = b_1*\text{typicality} + b_2*\text{homonegativity} + b_3*\text{typicality}^2*\text{homonegativity}$, where the $b$-coefficients can be found in Table 23 and the variable-scores are the extreme-endpoints discussed above (also, see Table 20). The results of this calculation suggests that the well-being of students who feel strongly atypical is more negatively affected in strongly homonegative schools than in more progressive schools (i.e., liberal schools: $\gamma_{\text{boys}} = .188(0) + -.128(.597) + -.093(0)(.597) = -.076$; homonegative schools: $\gamma_{\text{boys}} = .188(0) + -.128(2.43) + -.093(0)(2.43) = -.311$). We use the same formula to calculate the effect on the well-being of students who feel highly typical (hence, scoring 4 on gender typicality). This calculation shows that the well-being of students who feel highly typical is also lower in homonegative versus liberal schools (i.e., liberal schools: $\gamma_{\text{boys}} = .188(4) + -.128(.597) + -.093(4)(.597) = .454$; homonegative schools: $\gamma_{\text{boys}} = .188(4) + -.128(2.43) + -.093(4)(2.43) = -.464$). By comparing these effects, we see that the well-being difference between schools was more pronounced for typical students than it was for atypical ones (i.e., Atypical: -.076 - (-.311) = .235; Typical: .454 - (-.464) = .918).

We found similar associations for the cross-level interaction between gender typicality and traditional school cultures, as can be seen in model 3 of Table 23. Similar to model 1, gender typicality related positively to well-being, whereas a traditional school climate related negatively to well-being. The cross-level
interaction between typicality and traditional school climates was significant for both boys and girls, indicating that the well-being of (a)typical students varied according to a school’s level of traditionality. Once again, this effect was negative and similar in size for both genders ($\gamma_{\text{boys}} = -0.125, p < .01; \gamma_{\text{girls}} = -0.142, p < .05$). We use the same formula from above, supplementing schools’ scores on homonegativity with their scores on gender ideology —with the most progressive schools scoring 1.4, while the most traditional schools scored 2.19 (see Table 20). This gives the formula: $y = b_1 \times \text{typicality} + b_2 \times \text{ideology} + b_3 \times \text{typicality} \times \text{ideology}$. This calculation suggests that the well-being of students who feel strongly atypical is somewhat more negatively affected in strongly traditional schools than in more liberal schools (i.e., liberal schools: $\gamma_{\text{boys}} = 0.186(0) + -0.17(1.4) + -0.125(0)(1.4) = -0.239$; traditional schools: $\gamma_{\text{boys}} = 0.186(0) + -0.17(2.19) + -0.125(0)(2.19) = -0.373$). The well-being of students who feel highly typical is also more negatively affected in traditional versus liberal schools (i.e., liberal schools: $\gamma_{\text{boys}} = 0.186(4) + -0.17(1.4) + -0.125(4)(1.4) = -0.197$; traditional schools: $\gamma_{\text{boys}} = 0.186(4) + -0.17(2.19) + -0.125(4)(2.19) = -0.727$). However, the well-being difference between schools was once again more pronounced for typical students than it was for atypical ones (i.e., Atypical: $-0.239 - (-0.373) = 0.134$; Typical: $-0.197 - (-0.727) = 0.53$). Even though the effects were somewhat more prominent for homonegative school cultures than for traditional cultures, the patterns were the same for both indicators of school culture.

Adding the control variables in model 2 and 4 dispelled the significant effects of both indicators of school culture and the cross-level interactions. Analyses showed that for both genders, trust in teachers explained the context-dependent effects of typicality on well-being for both homonegative and traditional school climates. For boys, trust in teachers was the only variable that could dispel the context-dependent effects of gender typicality on well-being. For girls, being treated fairly by teachers accomplished the same effect as trust in teachers. Hence, for both genders, positive relationships with teachers mediated the context-effects on well-being.
Table 23.
Multilevel analyses of gender typicality on well-being:
Unstandardized coefficients and standard errors in parentheses

<table>
<thead>
<tr>
<th>Variables</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.873*** (.012)</td>
<td>2.616*** (.023)</td>
</tr>
<tr>
<td>school-level variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homonegativity</td>
<td>-.128*** (.030)</td>
<td>.024 (.054)</td>
</tr>
<tr>
<td>Traditionality</td>
<td></td>
<td>-.17*** (.051)</td>
</tr>
<tr>
<td>School size</td>
<td>0.0 (0.0)</td>
<td>0.0 (0.0)</td>
</tr>
<tr>
<td>SES school composition</td>
<td>-.073 (.097)</td>
<td>-.089 (.081)</td>
</tr>
<tr>
<td>cross-level interactions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typicality X homonegativity</td>
<td>-.093** (.029)</td>
<td>-.062 (.039)</td>
</tr>
<tr>
<td>Typicality X traditionality</td>
<td></td>
<td>-.125** (.048)</td>
</tr>
<tr>
<td>Individual-level variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.188*** (.010)</td>
<td>.110*** (.012)</td>
</tr>
<tr>
<td>Gender pressure</td>
<td>-.104*** (.013)</td>
<td>-.08*** (.013)</td>
</tr>
<tr>
<td>Trust in teachers</td>
<td>.179*** (.014)</td>
<td>.17*** (.013)</td>
</tr>
<tr>
<td>Equal treatment by teachers</td>
<td>.098*** (.01)</td>
<td>.097*** (.009)</td>
</tr>
<tr>
<td>Social integration:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>.265*** (.025)</td>
<td>.263*** (.022)</td>
</tr>
<tr>
<td>High</td>
<td>.44*** (.026)</td>
<td>.423*** (.023)</td>
</tr>
<tr>
<td>Improved</td>
<td>.248*** (.023)</td>
<td>.256*** (.022)</td>
</tr>
<tr>
<td>Variables</td>
<td>Boys</td>
<td>Girls</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Age</td>
<td>-.01</td>
<td>-.03</td>
</tr>
<tr>
<td></td>
<td>(.018)</td>
<td>(.016)</td>
</tr>
<tr>
<td>Sexual orientation:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGB</td>
<td>-.024</td>
<td>-.035</td>
</tr>
<tr>
<td></td>
<td>(.055)</td>
<td>(.05)</td>
</tr>
<tr>
<td>Questioning</td>
<td>-.062</td>
<td>-.105</td>
</tr>
<tr>
<td></td>
<td>(.062)</td>
<td>(.056)</td>
</tr>
<tr>
<td>SES (0 = white collar)</td>
<td>.024</td>
<td>.016</td>
</tr>
<tr>
<td></td>
<td>(.022)</td>
<td>(.02)</td>
</tr>
<tr>
<td>Ethnicity (0 = West-European)</td>
<td>-.062*</td>
<td>-.049*</td>
</tr>
<tr>
<td></td>
<td>(.027)</td>
<td>(.022)</td>
</tr>
<tr>
<td>Individual homonegativity at baseline</td>
<td>.006</td>
<td>.049**</td>
</tr>
<tr>
<td></td>
<td>(.012)</td>
<td>(.02)</td>
</tr>
<tr>
<td>Individual change in homonegativity</td>
<td>-.011</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.013)</td>
<td>(</td>
</tr>
<tr>
<td>Individual gender ideology</td>
<td></td>
<td>-.027*</td>
</tr>
<tr>
<td></td>
<td>(</td>
<td>(.012)</td>
</tr>
</tbody>
</table>

**Variance components**

<table>
<thead>
<tr>
<th></th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>School u_{0k}</td>
<td>.003</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td>(.001)</td>
</tr>
<tr>
<td>Cross-level interaction u_{02k}</td>
<td>.002</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td>(.001)</td>
</tr>
<tr>
<td>Student e_{0jk}</td>
<td>.120</td>
<td>.067</td>
</tr>
<tr>
<td></td>
<td>(.004)</td>
<td>(.004)</td>
</tr>
<tr>
<td>Time e_{0jk}</td>
<td>.156</td>
<td>.134</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>(.003)</td>
</tr>
</tbody>
</table>

*Note. Tests represented here were two-sided. Model 2_{boys}: N_{School}: 57, N_{Student}: 1725.
Model 2_{girls}: N_{School}: 47, N_{Student}: 1722. Model 4_{boys}: N_{School}: 57, N_{Student}: 2562.
Model 4_{girls}: N_{School}: 52, N_{Student}: 2181.
*p < .05. **p < .01. ***p < .001.
Table 24.
Multilevel analyses of pressure for gender conformity on well-being:
Unstandardized coefficients and standard errors in parentheses

<table>
<thead>
<tr>
<th>Variables</th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 5</td>
<td>Model 6</td>
</tr>
<tr>
<td>Intercept</td>
<td>2.833*** (.012)</td>
<td>2.616*** (.023)</td>
</tr>
<tr>
<td><em>school-level variables</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homonegativity</td>
<td>-.132*** (.031)</td>
<td>.019 (.054)</td>
</tr>
<tr>
<td>Traditionality</td>
<td>-.176*** (.051)</td>
<td>.068 (.066)</td>
</tr>
<tr>
<td>School size</td>
<td>0.0 (0.0)</td>
<td>0.0 (0.0)</td>
</tr>
<tr>
<td>SES school composition</td>
<td>-.068 (.097)</td>
<td>-.087 (.081)</td>
</tr>
<tr>
<td><em>cross-level interactions</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pressure X homonegativity</td>
<td>.135*** (.038)</td>
<td>.065 (.059)</td>
</tr>
<tr>
<td>Pressure X traditionality</td>
<td>.175** (.063)</td>
<td>.079 (.075)</td>
</tr>
<tr>
<td><em>Individual-level variables</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender pressure</td>
<td>-.045*** (.013)</td>
<td>-.102*** (.017)</td>
</tr>
<tr>
<td>Gender typicality</td>
<td>.110*** (.011)</td>
<td>.112*** (.01)</td>
</tr>
<tr>
<td>Trust in teachers</td>
<td>.18*** (.014)</td>
<td>.17*** (.013)</td>
</tr>
<tr>
<td>Equal treatment by teachers</td>
<td>.097*** (.01)</td>
<td>.097*** (.009)</td>
</tr>
<tr>
<td>Social integration:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>.267*** (.025)</td>
<td>.263*** (.022)</td>
</tr>
<tr>
<td>High</td>
<td>.439*** (.026)</td>
<td>.421*** (.023)</td>
</tr>
<tr>
<td>Improved</td>
<td>.25*** (.023)</td>
<td>.256*** (.022)</td>
</tr>
<tr>
<td>Age</td>
<td>-.01 (.018)</td>
<td>-.03 (.016)</td>
</tr>
<tr>
<td>Variables</td>
<td>Boys Model 5</td>
<td>Boys Model 6</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------</td>
<td>--------------</td>
</tr>
<tr>
<td><strong>Sexual orientation:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LGB</td>
<td>-.026</td>
<td>-.034</td>
</tr>
<tr>
<td></td>
<td>(.055)</td>
<td>(.05)</td>
</tr>
<tr>
<td>Questioning</td>
<td>-.064</td>
<td>-.111</td>
</tr>
<tr>
<td></td>
<td>(.062)</td>
<td>(.056)</td>
</tr>
<tr>
<td>SES (0 = white collar)</td>
<td>.022</td>
<td>.014</td>
</tr>
<tr>
<td></td>
<td>(.022)</td>
<td>(.02)</td>
</tr>
<tr>
<td>Ethnicity (0= West-European)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individual homonegativity at baseline</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.012)</td>
<td></td>
</tr>
<tr>
<td>Individual change in homonegativity</td>
<td></td>
<td>-.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.013)</td>
</tr>
<tr>
<td>Individual gender ideology</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Variance components</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School $u_{0k}$</td>
<td>.002</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>(.001)</td>
<td>(.001)</td>
</tr>
<tr>
<td>Cross-level interaction $u_{02k}$</td>
<td>.002</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>(.002)</td>
<td>(.003)</td>
</tr>
<tr>
<td>Student $e_{0jk}$</td>
<td>.139</td>
<td>.067</td>
</tr>
<tr>
<td></td>
<td>(.005)</td>
<td>(.004)</td>
</tr>
<tr>
<td>Time $e_{0ijk}$</td>
<td>.160</td>
<td>.134</td>
</tr>
<tr>
<td></td>
<td>(.003)</td>
<td>(.003)</td>
</tr>
</tbody>
</table>

*Note. Tests represented here were two-sided. Model 6$_{boys}$: N$_{School}$: 57, N$_{Student}$: 1975.
Model 6$_{girls}$: N$_{School}$: 54, N$_{Student}$: 3293. Model 8$_{boys}$: N$_{School}$: 57, N$_{Student}$: 2562.
Model 8$_{girls}$: N$_{School}$: 54, N$_{Student}$: 3293.
*p < .05. **p < .01. ***p < .001.
**Context-dependent associations between gender-conformity pressure and well-being.** As can be seen in model 5 and 7 of Table 24, gender-conformity pressure, homonegative and traditional school climates related negatively to well-being. This was true for both boys and girls, indicating that students who experienced high levels of gender-conformity pressure reported lower well-being, and that schools with more heteronormative climates took a negative toll on students’ well-being. However, the cross-level interactions between pressure and both indicators of school culture were only significant for boys. For girls, there was no significant random slope variance for gender-conformity pressure ($p = .055$). This implies that the effect of gender-conformity pressure on girls’ well-being did not vary between schools.

Nevertheless, boys did experience context-dependent effects of gender-conformity pressure on their well-being. As can be seen in model 5 and 7, the cross-level interaction was positive and similar in size for both indicators of school climates. ($\gamma_{\text{homonegativity}} = .135, p < .001; \gamma_{\text{traditionality}} = .175, p < .01$). Since the effects were comparable, the calculations discussed below for homonegativity can be extrapolated to traditional school climates.

Once again, to assess the effects of the cross-level interactions, we calculate and compare the effects on the extreme endpoints of the scales. That is, we compare the effects on well-being of boys who experienced the lowest amount of pressure (that is, zero) in schools that were either the most progressive (hence, scoring .597 on the homonegativity scale) versus in schools that were the most homonegative (scoring 2.43 on the homonegativity scale) (see Table 20). To do this, we use the formula: $y = b_1*\text{pressure} + b_2*\text{homonegativity} + b_3*\text{pressure}^*\text{homonegativity}$, where the $b$-coefficients can be found in Table 24 and the variable-scores are the extreme-endpoints discussed above (also, see Table 20).

The results of this calculation suggests that the well-being of boys who felt little gender-conformity pressure is more negatively affected in strongly homonegative schools than in more liberal schools (i.e., liberal schools: $\gamma_{\text{boys}} = -.045(0) + -.132(.597) + .135(0)(.597) = -.079$; homonegative schools: $\gamma_{\text{boys}} = -.045(0) + -.132(2.43) + -.135(0)(2.43) = -.321$). However, the well-being of boys who felt high gender-conformity pressure (and hence, scored 3 on the pressure scale) was more positively affected in homonegative than in liberal schools (i.e., liberal schools: $\gamma_{\text{boys}} = -.045(3) + -.132(.597) + .135(3)(.597) = .028$; homonegative schools: $\gamma_{\text{boys}} = -.045(3) + -.132(2.43) + -.135(3)(2.43) = .529$).

Adding the control variables in model 6 and 8 dispelled the significant effects of both indicators of school culture and the cross-level interactions. Analyses showed that for traditional school cultures, three separate variables could
explain the context-dependent effects of gender-conformity pressure on boys’ well-being. That is, adding either social integration, trust in or equal treatment by teachers dispelled the cross-level interaction. Hence, any indicator of positive relationships with peers or teachers mediated the context-effects on boys’ well-being. For homonegative school cultures a combination of variables was needed to account for the cross-level interaction. That is, adding both social integration and equal treatment by teachers to the model explained the context-dependent effects of gender-conformity pressure on boys’ well-being.

**DISCUSSION**

This paper examined in a multilevel fashion the well-being of gender-divergent students, with a specific focus on the ways in which schools’ level of heteronormativity affected the associations between gender divergence and well-being. In contrast to previous research’s focus on school policy, this study accounts for schools’ heteronormativity levels by considering two indicators of school culture: traditional gender attitudes and homonegativity. Additionally, by turning the focus from sexual orientation to gender divergence, attention is drawn towards the broad-ranging impact of heteronormativity on all students. Furthermore, by using multilevel analysis and focusing on the contextual specificity of these associations, possible avenues for change can be highlighted.

**Gender Typicality and School Cultures**

The results show that students tend to feel better in less heteronormative schools. That is, regardless of gender, both students who scored high and low on gender typicality tended to display lower levels of well-being in schools with more traditional and homophobic school cultures. These results underscore the importance of an accepting and gender-equal environment for the well-being of not only LGB or gender-divergent students, but all students.

Contrary to our expectations, the negative impact of heteronormative school cultures tended to be worse for typical students than for atypical students. That is, while both typical and atypical students exhibited lower levels of well-being in more traditional and homophobic school cultures, the difference between typical students in progressive versus heteronormative schools tended to be more pronounced than for atypical students. Or, in other words, the well-being of atypical students was less affected by the school environment than that of typical students. Furthermore, this association was not explained by peer acceptance, as
previous studies suggested (e.g., Goodenow et al., 2006; Smith & Leaper, 2006). Instead, the cross-level interaction between gender typicality and school culture was explained by students’ trust in teachers, or for girls, by their sense of being treated fairly by teachers. This result suggests that, besides peer acceptance, other sources of support in a school environment are of vital importance to students’ well-being. More specifically, this study shows that the relationship between students and teachers is an important factor in explaining variance in students’ well-being between schools. This way, this study adds to previous studies that already hinted at the importance of positive contact with faculty for the well-being of LGB-students (Goodenow et al., 2006; Murdock & Bolch, 2005; Szalacha, 2003). Additionally, the differential impact of school culture on the well-being of gender (a)typical students and the central role of trust in teachers herein were similar for boys and girls. That is, even though small differences in coefficients existed, the observed patterns were confirmed for both genders.

The results imply that the student-teacher relationship, and more specifically the trust that students have in their teachers, is more precarious in heteronormative schools. Other research has shown that LGB students who feel that their teachers were responsive to bullying based on sexual orientation, felt safer at school (Elze, 2003; Goodenow et al., 2006). Similarly, research regarding sexism in the school environment has shown that girls’ perceptions of teachers ameliorate when teachers confront sexist remarks in the classroom (Boysen, 2013). Hence, these studies suggest that teachers who promote safe and bias-free environments in their daily classroom practices, contribute to better student-teacher relationships and feelings of safety. This process might underlie the current findings of reduced trust in teachers in heteronormative schools, though the connection between student-teacher relationships in different school cultures should be investigated more thoroughly in future research.

Interestingly, the student-teacher relationships seemed to be more central in ensuring typical students’ well-being than that of their atypical counterparts. That is, the well-being difference in heteronormative versus more liberal schools was several times larger for typical students than it was for atypical students. A possible explanation for this surprising finding lies in the academic functioning of (a)typical students. That is, research has shown that typical students tend to score higher on academic self-efficacy and autonomous study motivation than their atypical counterparts (Vantieghem & Van Houtte, 2015b; Vantieghem et al., 2014a). Since students with higher academic functioning also tend to have better relationships with teachers, including trusting them more (Adams, 2014; Tschannen-Moran, 2014), this might explain the stronger effects for typical students. That is, because
gender typical adolescents tend to be more academically engaged students, the quality of the relationship with their teachers might be more influential for their well-being than for gender atypical students. Consequently, gender typical students might be more susceptible to the negative influence of poorer student-teacher relationships in heteronormative schools. On the other hand, the quality of the relationship with the teacher might be less influential for disengaged students, who tend to be disproportionately gender atypical students. This results in a seemingly more stable sense of well-being for gender atypical students, regardless of the school context they find themselves in.

**Gender-conformity Pressure and School Cultures**

The analyses for gender-conformity pressure pose a different yet related picture to that of gender typicality. For instance, where the effects of school culture on well-being were similar for gender (a)typical boys and girls, we encountered only significant effects for boys when considering gender-conformity pressure. Hence, while the associations between gender-conformity pressure and well-being were dependent on contextual factors for boys, this was not the case for girls. In other words, girls seemed to experience similar consequences of gender-conformity pressure, regardless of the schools they attended or the peers they interacted with. This is in line with previous research findings stating that pressure for gender conformity tends to be higher among boys (Egan & Perry, 2001), that displaying gender non-conformity is punished more severely among boys than girls (Ueno & McWilliams, 2010; Young & Sweeting, 2004), and that schools’ different gender regimes result in somewhat different constructions of hegemonic masculinity (Connell, 1996; Swain, 2005). This research adds to these previous research findings by demonstrating that more heteronormative schools tend to be environments where gender-conformity pressures are more prevalent and that these heightened pressures affect the well-being of boys.

While the well-being of boys in progressive schools was more comparable regardless of felt pressure, we observed larger differences in well-being levels among boys in heteronormative environments. Contrary to expectations, the absolute level of students’ well-being was not necessarily better in liberal schools. More specifically, while the well-being of boys who experienced little pressure was somewhat lower in heteronormative versus liberal schools, the well-being of boys who experienced high conformity pressure was actually a little higher in heteronormative environments. For a school culture based on homonegativity, this cross-level association was explained by either social integration among peers, being treated fairly by or having trust in teachers. For a school culture based on traditional gender attitudes, this cross-level association was explained by a
combination of social integration among peers and a feeling of being treated fairly by teachers (i.e., one of these variables was not enough to explain the association away, but including both of them in the model did). Hence, these analyses suggest that positive relationships with peers and/or teachers explain why boys who experience high conformity-pressure display higher levels of well-being in heteronormative schools, whereas boys who experience low levels of pressure display lower well-being.

It seems bizarre at first glance that boys who experience a lot of conformity-pressure, would experience more social acceptance in environments that adhere to more rigid gender rules. However, research has demonstrated that children who experience higher levels of gender-conformity pressure, tend to display less cross-gender behavior (Carver et al., 2003; Egan & Perry, 2001). Consequently, it might be that in school environments which perpetuate strict gender norms, boys who adhere to these social rules more completely are rewarded with more social acceptance. Note that the analyses suggest that this social acceptance is not just by peers, but also by teachers. That is, boys who felt higher conformity-pressure felt that their teachers treated them more fairly than boys who did not experience conformity pressure. This suggests that conforming to traditional masculine behavioral norms might not just be rewarded by peers, but similarly by teachers. However, we had no way to assess actual gender-conforming versus cross-gender behavior in the data. Therefore, this proposed mediational pathway of gender-conforming behavior explaining higher social acceptance for boys with high-conformity pressure should be explored in future research.

**Overall Effects of Heteronormative School Cultures on Gender-divergent Students**

Overall, this study suggests that the well-being of students tends to be more comparable in progressive schools than in heteronormative schools. In other words, the well-being disparities for different types of students is larger in heteronormative environments than it is in more gender-equal and tolerant environments. This was true for both boys and girls, but results were more pronounced for boys (especially with regards to gender-conformity pressure). This pattern was similarly confirmed regardless of students’ gender typicality or experienced pressure for gender conformity. Hence, this implies that interventions aimed at creating a more tolerant and equal environment would not only be beneficial for LGB or gender-divergent students, but would ameliorate the school experiences of all students, even those who feel gender typical or experience little conformity-pressure.
Additionally, the findings suggest that teachers play an important role in determining the well-being of their students, and that teacher-student relationships tend to be more precarious in heteronormative schools. That is, students’ trust in teachers was lower in heteronormative schools and boys’ feelings of being treated fairly by teachers seemed to be more conditional on displaying “appropriate” gender-conforming behavior. Consequently, faculty who aim to make their schools more accepting and welcoming environments might benefit from promoting positive teacher-student relationships, besides interventions aimed more specifically at challenging the heteronormative culture of their schools.

Moreover, by considering several aspects of gender-divergence (i.e., gender typicality and gender-conformity pressure), we have uncovered processes that are similar for boys and girls, whereas other patterns seem to be gender-specific. Specifically, the results suggested that the context-dependent effects of gender typicality on well-being followed comparable patterns for boys and girls. These findings are in line with other research which found that boys and girls who scored low on self-perceived typicality had equally low levels of autonomous motivation and academic self-efficacy (Vantieghem et al., 2014a). Conversely, the context-dependent effects of gender-conformity pressure on well-being were only encountered among boys in this sample. As discussed above, this is in line with previous research which consistently found stronger gender-conformity pressures among boys (Egan & Perry, 2001) and ethnographic research which uncovered the divergent ways school’s gender regimes contribute to specific norms for masculine behavior (Connell, 1996; Swain, 2005). Additionally, these findings confirm the model proposed by Egan and Perry (Egan & Perry, 2001). That is, Egan and Perry discerned several gender dimensions, including gender typicality and gender-conformity pressure, and stated that these dimensions would function independently from one another (Egan & Perry, 2001). The current analyses seem to confirm this: gender typicality and gender-conformity pressure explained students’ well-being in different ways, and these associations were significant even though the other dimension was included in the controls. This suggests that several aspects of gender influence students’ school experiences in different ways, which is something future research should take into account.

Lastly, the current analyses showed that similar patterns were observed when using either traditional gender role attitudes or homonegativity as indicator of heteronormative school culture. So, even though these scales tapped into a notably different set of attitudes, they tended to have a similar impact on the well-being of gender-divergent students. This confirms the theoretical notion that heteronormativity equally builds upon a press toward heterosexuality as toward
traditional gender binaries. As suggested by Chesir-Teran (2003), future research should include other school aspects of heteronormativity, such as physical-architectural elements, program-policy features, social interaction and other suprapersonal factors—besides the ones considered in this research. If other aspects of heteronormativity influence students’ school experiences in similar ways, a strong case could be made for heteronormativity as a single underlying latent construct. This research is a first step in that direction by showing that two suprapersonal factors with regard to heteronormative cultures result in similar patterns, despite strong divergence in the item content of these scales.

Limitations
Despite several strong aspects in this study’s methodology, it was still subject to some limitations. One, even though we used data from students on three time-points—which heightened the robustness of the findings—no causality claims can be made. That is, we were unable to assess the possible changes in students’ well-being when transferring from a more liberal to a more heteronormative school or vice versa. Such an analysis would require a cross-classified multilevel design, for which we lacked the appropriate data. Rather, the findings apply to average well-being differences in several types of schools. Additionally, despite several demographic controls and a thorough study design, we were unable to completely rule out underlying school characteristics that might have confounded the obtained patterns. More specifically, schools that scored high on heteronormativity tended to be schools that offered vocational tracks with larger proportions of boys, students from lower socio-economic background and ethnic minority descent. Disentangling these different aspects of a school’s make-up proved to be difficult because of the high correlations between these aspects in the Belgian context (for instance, the correlation between school’s homonegativity and the proportion of students from ethnic minority descent was .76, the correlation between schools’ traditional gender attitudes and proportion of students with blue collar background was .65). Hence, we suggest repeating this research in areas where educational track, gender, ethnicity and SES composition are less entwined, to offer a clearer view on the influence of school cultures of heteronormativity on students’ well-being.

Second, the current research considered the associations between schools’ heteronormativity and students’ well-being in a sample of early adolescents. Thus, the current findings cannot be extrapolated beyond this age group. Other research should consider these associations with different age groups to assess the pervasiveness of these patterns. Poteat and colleagues (Poteat et al., 2009) have demonstrated that relational expressions of homonegativity (i.e., the unwillingness
to remain friends) are especially high during early adolescence and tend to be lower among older students. Consequently, we could suspect that the influence of heteronormativity on students’ well-being might be especially pronounced in the age group considered in this study.

Third, this research considered self-perceived gender typicality and felt pressure for gender conformity. Considering other sources might result in different patterns. That is, gender typicality and felt conformity-pressure are self-perceived indicators of gender divergence that might not always be apparent to significant others. It is conceivable that working with peer or teacher reports of non-conformity might result in different well-being patterns for gender-divergent adolescents. There might be considerable discrepancy between self-assessed and third-party-observed gender divergence, since what is most central for felt gender typicality, for instance, tends to differ from person to person (Perry & Pauletti, 2011; Tobin et al., 2010).

Fourth, the between-school variance in well-being was relatively low in this study (more specifically, 3.97% for girls and only 1.63% for boys). While the occurrence of a cross-level interaction under such conditions is especially noteworthy, research which focuses on other outcomes that might display more school variance —such as victimization or school belonging— might prove especially interesting for school-based interventions.

In conclusion, this study showed that heteronormative schools seem to be less optimal environments for all students, regardless of sex, sexual orientation or gender-divergence. More specifically, it appears that student-teacher relationships are more fragile and conditional in heteronormative environments, affecting the well-being of all students attending in adverse ways. All in all, these findings could constitute a strong appeal to schools for tackling heteronormativity, since doing so would not only augment’s students’ well-being, but would simultaneously promote better relationships between students and teachers.
Chapter 11.
Discussion

MAJOR FINDINGS
It has become clear in the last decades that girls are outperforming boys on most educational parameters, and that this trend is occurring to a certain extent in all industrialized western countries (PISA, 2009). A lot of research has gone into explaining this “boy crisis” in education. For instance, the average male underachievement is often framed as a consequence of anti-school cultures among boys’ peer groups (Epstein, 1997, 1998; Francis, 2000; Martino, 1999). While this masculinities-theory-based research has been very illuminating, it tends to remain linked to specific settings because of the research methods employed. Furthermore, the focus on boys has impeded thorough analyses of girls’ situations (Schippers, 2007), and variation within groups of students has frequently been overlooked (Francis, 2000, 2010). By using representative datasets and continuous measures, quantitative research could make more general claims and remedy the neglect of within and between-gender variations. However, quantitative research on the educational gender gap has focused predominantly on boy-girl differences (e.g., Van de Gaer et al., 2006a; Van Houtte, 2004b), thereby losing sight of the nuanced ways in which gender functions. Consequently, in this dissertation, we proposed to use the constructs of gender typicality and gender-conformity pressure to bring a gender focus back to quantitative educational gender gap research. Gender typicality refers to the extent to which people feel themselves to be a good example of their gender category (Egan & Perry, 2001), and as such draws upon intra-individual identity processes. Conversely, gender-conformity pressure refers to the degree to which one experiences a push to uphold gender norms (Egan & Perry, 2001), and as such reflects socializing pressures. Hence, by considering these gender-aspects, both an intra-individual or more social-psychological dimension, as well as an interactional or more sociological level are tapped (Risman, 2004).

This research was motivated by turning the lens on both inter- and intrasexual variation in educational functioning. Additionally, we wanted to re-introduce gender aspects to quantitative educational gender gap research. When considering the first aim, this dissertation showed that there are important ways in which inter- and intragender variation connect. That is, in some respects, boys and girls can be quite similar to each other. This is, for instance, the case with self-perceived gender atypical students who score low on well-being and academic
school functioning irrespective of their sex. More specifically, our analyses suggest that gender atypical boys and girls tend to feel rather bad about themselves, and that this reduced well-being weighs upon their academic self-efficacy and autonomous study motivation. With regards to these equivalent patterns among boys and girls, we could say that atypicality trumps sex.

In other respects, however, differences between genders tend to be more prominent. For instance, when students scored more extremely on felt typicality or gender-conformity pressure, the sex differences in academic functioning were larger as well. More specifically, self-perceived gender typical girls scored the highest of all students on both academic self-efficacy and autonomous motivation. Despite having similar scores on self-perceived typicality, boys scored considerably lower on academic functioning than their female counterparts. Notably, the more extreme students scored on self-perceived typicality, the larger the differences in academic functioning were between boys and girls. This sex-specific working of gender aspects was even more pronounced for gender-conformity pressure. That is, our results showed opposite associations between gender-conformity pressure and academic self-efficacy for boys and girls. When boys experienced higher rates of gender-conformity pressure, their academic self-efficacy was lower. Conversely, once controlled for well-being, a positive effect of gender-conformity pressure on girls’ self-efficacy was revealed.

All in all, our findings show more prominent differences in academic functioning between boys and girls, when they adopted traditional gender expectations more completely. Concerning the link between such sex and gender differentiations, some researchers posit that more egalitarian and prosperous environments would amplify differences between the genders—for instance, in personality traits—, because innate dispositions would have more freedom to fully develop (Schmitt, Realo, Voracek, & Allik, 2008). In contrast, other researchers claim that gender egalitarian societies might reduce the number and size of gender differences, because there would be less sex-specific roles, division of labor and socialization (Hyde, 2014). As Reinisch and colleagues (1997, p. 42) put it: “[…] The more divergent the sex stereotypes within a culture, the more the biologically based sexual dimorphisms in behavior tend to be magnified by postnatal socio-environmental differences. In such societies, sex differences which have no basis in biology also may be generated through highly differentiated sex role expectations”. Our findings show stronger differences between boys and girls when they score highly on gender-aspects. As such, these individual-level results are in line with the idea that stronger societal differentiation between genders might contribute to more actual, observable differences between men and women.
A second aim of the research was to reconsider more subtle gender aspects in quantitative educational research. Our studies show the relevance of including gender besides sex, since gender expectations tend to have a very real impact on early adolescents’ school experiences. Regardless of whether students internalized these gender norms into a sense of identity (as is the case with gender typicality), or experience them as an external pressure (such as school cultures or gender-conformity pressures), gender norms were connected to both well-being and academic school functioning for all students in the sample. Some striking patterns were discovered in the results.

First, gender norms seemed to influence students’ school experiences in consistent ways. For instance, students’ well-being was affected in the same way by school cultures concerning traditional gender ideologies and concerning homonegative ideas. This strengthens the theoretical notion of heteronormativity as a latent construct, being equally evident in traditional views on sexuality and gender. Furthermore, with regard to gender typicality, the results overlapped extensively for both study motivation and academic self-efficacy. Since these outcomes are quite different indicators of academic functioning, this suggests that gender typicality influences academic school functioning in a robust way. All in all, by assessing the influence of gender norms in different ways and on different outcomes, we showed that the uncovered patterns are not coincidental. Rather, our results point to coherent mechanisms in how gender influences students’ everyday experiences.

A second pattern found in the data is that femininity, rather than masculinity, is repeatedly linked to superior school functioning. This was evident in the higher scores on both self-efficacy and autonomous motivation of gender-typical girls, as well as in the diametric associations between gender-conformity pressure and self-efficacy for boys and girls. These results confirm the idea that in the current educational context, study behavior is more compatible with stereotypical feminine characteristics (Coenen, Meng, & Velden, 2011; Elmore & Oyserman, 2012; Jackson, 2003), resulting in a small but consistent female advantage on most school parameters. However, it should be noted that being a girl is not singularly connected to advantageous outcomes. For instance, self-perceived atypical girls scored badly on both well-being and academic functioning. Such findings are an important nuance in the debate on the educational gender gap, since both media and research tend to focus on the average achievement differences between the sexes. Because of this, the wide range of variation within each group as well as the underlying mechanisms contributing to these variations tend to be masked. This
dissertation might contribute to more nuanced accounts of gender differences in the educational domain.

A third pattern emerging from the studies is that living up to gender norms does not singularly connect to advantageous outcomes. For instance, gender typicality was associated with higher levels of controlled motivation. This indicates that boys and girls, who viewed themselves as living up to gender expectations, studied more because they felt pressured to do so. Such external study motivation is connected to suboptimal study behaviors (Walker, Greene, & Mansell, 2006) and poorer school trajectories (Vallerand, Fortier, & Guay, 1997), since students tend to give up once the pressure disappears (Haerens, Kirk, Cardon, De Bourdeaudhuij, & Vansteenkiste, 2010). It remains unclear what the mechanism behind this association between controlled motivation and gender typicality is. One possible explanation is that these students are more prone to internalizing environmental pressures, resulting in both higher rates of typicality and controlled motivation. Another possibility is that these students experience more actual study pressure, due to certain expectations of “typical” boys or girls. Thus, while the origin of this association remains unclear, it is important to underline that gender typicality is not solely connected to advantageous outcomes, such as higher well-being and academic school functioning. Consequently, while self-perceived atypicality has often been linked to adverse consequences (Carver, Yunger, & Perry, 2003; Egan & Perry, 2001), this study shows that self-perceived typicality is not singularly unproblematic either.

Such negative consequences of gender norms emerged even more clearly when considering gender-conformity pressures. When such pressures are assessed at the individual level, we see that they negatively connect to boys’ academic self-efficacy. For girls, the pattern was more complex. That is, at first sight, girls displayed a stable sense of self-efficacy regardless of the intensity of gender-conformity pressure they experienced. However, when we controlled for well-being, gender-conformity pressures connected to a higher sense of academic self-efficacy. Hence, while gender-conformity pressures do not lead to a reduction in girls’ academic self-efficacy at first glance, they do contribute to a deteriorated well-being which, in turn, is connected to reduced academic functioning (Assor, Kaplan, Kanat-Maymon, & Roth, 2005; Bandura, 1997; Lyubomirsky, King, & Diener, 2005). So, while gender-conformity pressures are more intensely associated with adverse consequences among boys, gender-conformity pressures seem to be detrimental for the well-being of any student.

When assessing gender-conformity pressures on an institutional level, as we did by considering heteronormative school cultures, we revealed a similarly
negative connection. That is, students’ well-being was negatively affected in schools where traditional views on masculinity, femininity and sexuality reigned. More specifically, our results show that heteronormative environments were not only detrimental to gender-divergent or LGB-pupils. Rather, gender-equal schools are favorable for all students’ well-being, regardless of sex, gender divergence or sexual orientation. Furthermore, our results unveiled that the quality of relationships with teachers was decisive in accounting for well-being differences across school contexts. That is, student-teacher relationships tended to be more strained in heteronormative environments, which were in turn connected to students’ well-being. Such findings could constitute a strong appeal for taking gender issues and heteronormativity at school seriously. For one, we showed that schools varied in the extent to which a push for heterosexuality and traditional gender binaries was present. Hence, schools are not just a reflection of a society-wide phenomenon, but could be important sites for social action and change. Second, our results imply that schools who apply themselves to promoting gender-sensitive approaches would not only augment students’ well-being, but would simultaneously promote improved relationships between students and teachers (see 11.3.2 for a discussion of ways for the educational community to challenge heteronormativity cultures at school).

Overall, this dissertation uncovered that gender typicality and gender-conformity pressures in a school context function in accordance with previously established mechanisms in gendered behavior. That is, behaving in traditionally masculine or feminine ways is not unequivocally connected to either positive or negative outcomes. Rather, in line with self-determination theory (Deci & Ryan, 1985), the underlying motivation for why people enact certain behavior is crucial. For instance, Good and Sanchez (2010) revealed that traditional gender behavior is connected to positive outcomes when this behavior is volitional. Conversely, when such behavior is motivated by external pressures, traditional gender behavior is connected to lower self-esteem. In line with these findings, our studies showed predominantly positive outcomes for an internalized motivator for gender-enactment (as reflected in gender typicality), while adverse consequences consistently surfaced for gender-conformity pressures on both the individual and institutional level. Hence, this dissertation lends further support to the social-psychological idea that the underlying motivation, rather than the actual behavior, is crucial in determining the consequences of gendered conduct. An important —sociological— nuance in this process is that environmental responses can exacerbate the consequences of such volitional versus enforced conduct. More specifically, this dissertation showed that
in an educational context, the positive consequences of feminine-typed behavior tend to be augmented, while those for masculine-typed conduct are attenuated, and vice versa for adverse outcomes.

LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Explaining the Educational Gender Gap

This dissertation sheds light on how school functioning is connected to gender norms and the well-being that is derived from living up to these norms. By using sex X gender interactions, we showed how embodying traditional notions of masculinity and femininity (as represented in typicality and pressure) contributes to an exacerbation of sex differences in academic functioning. Nevertheless, while it was not the aim of this dissertation, it should be noted that introducing typicality and conformity pressure did not explain the sex differences in school functioning. That is, the coefficients of the sex variable did not decline dramatically when the gender variables were entered into the analyses. Consequently, future research could delve deeper into the differential attainment of boys and girls. Previous research has already demonstrated that this is a complex, multi-faceted problem (e.g., Buchmann, DiPrete, & McDaniel, 2008; Meece, Glienke, & Burg, 2006; PISA, 2009; Voyer & Voyer, 2014), which suggests that combining several angles would be necessary to fully understand this issue. Below, we suggest several angles that could be considered in future research. Note that we follow Risman’s (2004) suggestions to take several gender levels into account, by considering institutional, as well as individual and interactional dimensions. Since previous research has mostly focused on the interactional level, we feel that a focus on intrapersonal and institutional levels might be especially relevant. Additionally, we would advise taking different levels into account simultaneously (as we have done in chapter 10), since this might reveal especially illuminating and policy-relevant patterns.

On an intrapersonal level: Biology. The educational gender gap starts widening in early adolescence (see chapter 2), while gender-differential attainment is much less prevalent in elementary education. This process has been attributed to the gender intensification process in early adolescence, which posits that adolescents experience stricter gender norms resulting in more gender-specific behaviors (Galambos, Almeida, & Petersen, 1990). The origin of these stricter gender norms would be found in sex and gender-related developmental tasks of adolescence, such as physical maturation, consolidation of gender identity, development of sexual interests and managing an academic and occupational future
(Perry & Pauletti, 2011; Tobin et al., 2010). While physical maturation and sexual development are explicitly mentioned as central developmental tasks of adolescence, little research has explored these biologically-driven processes in connection to academic functioning. We could wonder, for instance, to what extent physical maturation itself, including brain development and hormonal changes, impacts upon gender intensification and school performance.

Additionally, the possibility exists that societal reactions exacerbate biological processes, with for instance, increased expectations for gendered adult behavior when adolescents experience a fast pace of physical maturation (Galambos et al., 1990). Likewise with regard to the development of gender identity, researchers are increasingly pointing towards the influence of both biological (including brain, hormonal, and gene variations) and social processes (including parental socialization, social status and environmental responses) (Liao, Audi, Magritte, Meyer-Bahlburg, & Quigley, 2012; Steensma, Kreukels, de Vries, & Cohen-Kettenis, 2013). However, Steensma and colleagues (2013, p. 294) remark that “Despite the acknowledgement that nature and nurture interact, researchers have so far not tried to integrate both aspects”. One of the reasons for this is the sheer complexity of trying to integrate two fields of study, each of which is highly intricate in its own right (Huber, 2007). However, another aspect hampering this integration is the persisting idea that what is rooted in biology is unchangeable, and consequently of little relevance for policy or social research (Hood-Williams, 1996). Research has demonstrated, however, that such ideas are flawed with, for instance, the discovery of the existence of gene X environment interactions (e.g., Kim-Cohen et al., 2006; Vermeersch, T’Sjoen, Kaufman, Vincke, & Van Houtte, 2010).

Human biology is not a static thing, guiding human beings on fixed paths to determined endpoints. Rather, biological plasticity thoroughly influences the human body, with displays of high flexibility and variation in answer to environmental conditions (Baumeister, 2000). This implies that biology and culture can come together to create new expressions of human diversity. Consequently, as long as it remains unclear where sex stops and gender begins, and —considering the research unveiling biology X environment interactions— where both intersect and interact, it seems pertinent to similarly uncover biological influences on gender and educational research. Interesting questions in this regard are the extent to which —school— environments can reduce or exacerbate innate dispositions by, for instance, framing certain behaviors as adaptive versus unfit for either sex.
On an interpersonal level: Peers. Second, this research showed that a sense of pressure leads to a decline in school engagement, especially among boys. This is in line with previous ethnographic research that demonstrated how boys’ peer cultures often revolve around school opposition (Connell, 1989; Martino, 1999). A strength of such qualitative research is that it has a level of detail and richness that quantitative research can rarely replicate. Consequently, while ethnographic researchers have more often considered intersections between class, race, sexual orientation, and gender on school engagement (e.g., Mac an Ghaill, 1994; Morris, 2012; Pascoe, 2007), few quantitative studies on study cultures, including our own, reach this level of complexity. Nevertheless, some quantitative researchers are starting to consider intersectionality more thoroughly by, for instance, using cluster analysis (e.g., Huyge, Van Maele, & Van Houtte, 2015a). Other researchers (e.g., Titkova, Ivaniushina, & Alexandrov, 2013) are employing a combination of network and multilevel analyses to reach a more detailed and complex understanding of gendered peer cultures. Hence, we would advise quantitative researchers to combine the power of large datasets with advanced analytical tools to achieve a more nuanced understanding of the complex social reality of contemporary school life.

On an interpersonal level: Teachers. Schools are complex environments where several social actors come together. While research usually focuses on students, it is equally relevant to consider other social agents who influence students’ school experiences. Principal among these actors is the teacher. We showed that student-teacher relationships are crucial to students’ well-being and more precarious in heteronormative environments (see Chapter 10). Future research could delve more deeply into teachers’ role in the gendered school experiences of students. For instance, a lot of such teacher-focused research has tried to discover whether the feminization of the teaching profession has contributed to the educational gender gap (e.g., Carrington & McPhee, 2008; Derks & Vermeersch, 2001; Einarsson & Granström, 2002; Marsh, Martin, & Cheng, 2008; Timmerman & Van Essen, 2004). So far, there is no conclusive evidence to suggest that female and male teachers employ qualitatively different teaching styles, let alone treat boys and girls so differently as to account for the educational gender gap (Timmerman & Van Essen, 2004). Nevertheless, it remains true that very few teachers are male (Van Woensel, 2007). Consequently, the possibility remains that the feminization of the faculty has contributed to the overall feminine image of the school institute, and that a lack of academically-engaged male role models might negatively contribute to boys’ identification with school. So far, there is no
empirical literature to support or deny this claim, and hence, this could be assessed in future research.

So, while teachers’ gender seems to have little influence on the way they are perceived by students (Carrington & McPhee, 2008), student’s gender does influence teachers’ perceptions and behaviors. For instance, many teachers tend to see girls as ideal students (Jones & Myhill, 2004; Leaper & Brown, 2014; Van Houtte, 2007), while boys are reprimanded more often than their actual behavior warrants (Consuegra, Halimi, & Engels, 2015). Heyder and Kessels’ research (2015) suggests that such teacher biases are rooted in stereotypes concerning the mismatch between masculinity and school. Future research should replicate these findings with other samples, investigate the consequences of these biases and ways to combat them.

**On an institutional level: Schools.** Besides individual agents, such as students and teachers, it is important to consider the impact of institutions as well. A question that seems to crop up time and again, especially in the media, is whether single sex schools would be more effective. Years of research has led to contradictory results, with some studies backing single sex education and others defending mixed schools. An important part of this confusion seems due to a lack of consideration for student make-up, with single sex schools often being elite schools consisting of more privileged students (Bigler, Hayes, & Liben, 2014; Sikora, 2014). Current consensus seems to center on the notion that the benefits of single sex schooling have not been proven (Bigler et al., 2014; Pahlke, Hyde, & Allison, 2014). Conversely, mixed schools would be more beneficial, because studies have shown that gender segregation contributes to gender-typed behavior among students (Martin, Fabes, & Hanish, 2014) and reinforces gender stereotypes (Glasser, 2012). More nuanced studies have considered student composition in percentages, rather than a black-and-white comparison of single sex versus mixed schools. For instance, Van Houtte (2004a) showed that boys tended to do better in schools with more girls, because the overall study culture was more academically oriented. In a similar vein, Demanet and colleagues (2013) demonstrated that students showed higher levels of school involvement in schools where more girls attended, lowering the overall levels of school misconduct. These studies seem to suggest that schools with a majority of girls are beneficial for both boys and girls. Given the almost 50-50 division of men and women in society, this seems an unobtainable goal, with mixed schools being the next best option. When school composition is opened up to consider other factors besides sex, Legewie and Diprete (2012) unveiled that the gender gap in reading achievement tended to be
larger in schools with lower SES-contexts. Consequently, since school compositions (both in SES or gender) are hard to change, future research might consider how academic school cultures can be stimulated regardless of the demographic make-up of a school.

In this dissertation, we considered school cultures by showing how heteronormative school cultures impact on students' well-being. We might wonder whether heteronormativity might influence students’ achievement as well, and because gender norms are stricter for boys (Herek, 2002; Young & Sweeting, 2004), whether their achievement might be more affected than girls’. Wilkinson and Pearson (2009) already demonstrated that heteronormativity at school is related to higher rates of school failure among LGB students, but it remains to be seen whether similar associations can be found among heterosexual students.

Besides school cultures and school composition, pedagogy is an influential aspect of the organizational structure as well. It has been suggested that schools are organized in a way that favors women, with pedagogies geared towards passive listening, and by placing importance on communication, cooperation, deadlines and neatness for homework (Martino & Kehler, 2006; Reints, 2013). Conversely, few school activities cater to “masculine” learning styles involving active exploration, problem-solving oriented approaches, experimentation, or outside activities. Little research has considered to what extent such claims are founded in reality, and the current findings seem to be somewhat contradictory. One study found, for instance, that girls rated themselves higher than boys on 8 out of 11 assessed school competencies, such as cooperating, communicating, planning and independent work (Coenen et al., 2011). The authors went on to note that schools had focused more on these competencies throughout the years, further increasing the gap between boys and girls. Conversely, another study found no significant differences between 16-year-old boys’ and girls’ preferences for activities, nor did either gender achieve better on more stereotypically masculine or feminine activities (Reints, 2013). Nevertheless, the authors did note that the study material contained few “masculine” activities. In conclusion, it remains unclear to what extent claims about “feminine pedagogies” can be confirmed, and importantly, to what extent boys and girls actually experience and benefit differently from such pedagogies. Future research should address this gap.
The Roots of Gender Typicality

The gender typicality scale developed by Egan and Perry (2001) is one of the best contemporary measures of self-assessed masculinity and femininity (see Chapter 4). Nevertheless, some issues with this instrument remain. For instance, the scale uses an identification approach, since previous identity-approaches (such as the BSRI or PAQ) proved flawed (Vantieghem, Vermeersch, & Van Houtte, 2014). Because of this identification approach, however, researchers are left with the question: what does gender (a)typicality mean to the respondent? Little is known about the meanings attached to the instrument of gender typicality (Corby, Hodges, & Perry, 2007). We suggest that this dearth of research might be partly due to the theoretical framework underlying the construct. That is, multifactorial theory states that people would build their overarching sense of gender identity in an idiosyncratic way (Spence, 1984). Consequently, what is most central for experienced gender identity might differ from person to person. Such a theoretical premise constructs felt gender identity as a thoroughly individual characteristic, which is almost impossible to research. From a sociological point of view, however, it seems plausible that people belonging to certain social groups attribute meaning in a rather consistent fashion. That is, in sociology, culture is defined as a fairly stable set of shared beliefs, assumptions and meanings (Parsons, 1951; Van Houtte, 2005), and there seems little reason to expect gender to be exempt from these processes. Hence, future research should investigate what meanings people attach to gender typicality, and consider systematic variations in these meanings according to respondents’ sexual orientation, socio-economic background, ethnicity, age, region, and so on. Qualitative research seems especially apt to study such variations in meaning. Nevertheless, some quantitative research has already unveiled ways in which social groups differ in responding to these instruments. Wilson and Leaper (2015) showed in a US context, for instance, that men and women from different ethnic groups tended to score differently on typicality and conformity-pressure. Furthermore, these measures connected to self-esteem in different ways for these social groups. Similarly, Huyge and colleagues (2015a) showed that Belgian LGB-students tended to self-identify primarily as gender atypical, while middle-class, white, high-achieving boys tended to score high on gender typicality and rather low on gender-conformity pressure. Such studies show that interpretations of what it means to be a boy or girl might vary in consistent fashions among certain social groups. Note that these studies follow an intersectional approach, which posits that different social statuses (such as ethnicity, gender, class, disability, age, and so on) not only have additional effects, but intersect to produce specific life histories and meanings (West & Fenstermaker, 1995). Additionally, such findings are in line with
central premises from masculinities theory. That is, Connell (1995) differentiated racialized, classed and sexualized versions of doing masculinity in her typology of hegemonic versus subordinated, marginalized and complicit masculinity. Hence, we propose that future research, especially qualitative research, could uncover shared meanings attached to gender typicality to help illuminate current research findings.

Other Aspects of Gender
Motivated by theoretical considerations and previous research (as discussed in Chapter 1), we focused on gender typicality and pressure for gender conformity as central gender aspects in this research. This does not mean, however, that other aspects of gender could not similarly contribute to an understanding of inter- and intrasexual variation in school functioning. Egan and Perry (2001), for instance, developed measures to examine gender contentedness and intergroup bias, while other researchers proposed a focus on interests (Lippa, 1998). Besides these intra-individual factors, gender attitudes could be considered as well. Limited research has already considered the influence of gender ideology on students’ school belonging (Huyge, Van Maele, & Van Houtte, 2015b) or the changing relationship between masculine school behavior and homonegativity (McCormack & Anderson, 2010). Furthermore, social identity theory and the identity-based motivation model underline the importance of salience, stating that social identities are only influential when they become salient in a given context (Oyserman & Destin, 2010; Stets & Burke, 2000b). Consequently, it seems relevant to consider which contextual cues contribute to the salience of gender identity. Some studies suggest that important cues for contextual salience are, for instance, the gender ratio in the people present (Abrams, Thomas, & Hogg, 1990), gender-based forms of class organization or gender-specific language (Bigler, 1995; Hilliard & Liben, 2010). While theoretically plausible (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), it remains unclear whether symbolic connotations of certain environments—such as schools—as masculine or feminine might have similar influences. Moreover, little is known about individual variation in salience (Stryker & Serpe, 1994). It seems plausible that gender identity is more readily available for activation in certain people, especially when gender is more central in their overall self-concept (Hogg, Terry, & White, 1995; Stets & Burke, 2000b). Factors contributing to the individual salience of gender identity should be researched and incorporated in future research.

While gender-related attitudes provide insight into individual motivations for behavior, these motivations usually remain invisible and unknowable to others in real-life situations. Actual gendered behavior, conversely, is the base on which most interactions and onlookers’ assumptions hinge. For instance, in the empirical
chapters, several explanatory pathways were suggested, but could not be confirmed due to a lack of data. Quite notable among these were gendered performances at school, such as observable “gender-appropriate” conduct, teachers’ responsiveness to bullying (see chapter 10) or —disruptive— class behavior. Other explanatory mediational pathways such as frustration of the need for autonomy (see chapter 8), beliefs regarding appropriate school behavior, and teacher expectations could be researched to further strengthen the proposed pathways presented in the empirical chapters. Consequently, besides gender-related attitudes, it would be interesting to include actual —gendered— behavior in future research as well.

Adolescent Development
Our studies uncovered some findings which are interesting from a developmental standpoint. Among these is the finding that 7th grade seems to be a crucial year for the socialization into gender-specific study cultures, and the observation that associations between students’ well-being and schools’ level of heteronormativity seem to be robust across middle school. Nevertheless, developmental changes were not the main focus of this dissertation. There are two reasons for this, which are both of a practical nature. First, the data used in this research only had a limited number of time-points. More specifically, it covered three time-points and spanned only two school years. Because of this, research questions remained limited to middle school, and strong developmental transitions were less likely to be uncovered. Second, because this research was part of an ongoing project, data only became available in the course of the study. This caused serious time constraints for running and writing up longitudinal analyses.

Consequently, future research could delve more deeply into the connection between gender-differential achievement and developmental evolutions in adolescence. For instance, we could wonder whether the influence of gender-conformity pressure on school functioning would drop off after mid-adolescence. Seeing as gender-conformity pressures tend to be highest in middle childhood and early adolescence, and decline as people grow older (Carver et al., 2003; Horn, 2007), it could be that its influence on academic functioning dies out by the end of secondary education. Conversely, we could wonder whether gender typicality retains its influence throughout the life course. Since gender typicality is an indicator of an essential social identity (Elmore & Oyserman, 2012), it could be hypothesized to remain a strong behavioral motivator throughout people’s lives. Some studies have already connected gender typicality to self-efficacy, study choice (Leaper & Van, 2008) and sensation seeking (Saxvik & Joireman, 2005) among college students, thereby showing that, at least among emerging adults, gender
typicality tends to remain an influential construct for gendered and academic behavior.

**School Functioning and Achievement**

In this dissertation, central outcomes were academic self-efficacy, study motivation and well-being. The expectancy-value model (Wigfield & Eccles, 2000) and previous research demonstrated the importance of these variables for broader school functioning (Bandura, 1997; Deci, Vallerand, Pelletier, & Ryan, 1991), as well as achievement (Hattie, 2008). This dissertation revealed some interesting connections between gender aspects and these educational outcomes. Nevertheless, not all proposed relations in the conceptual model have been tested. More specifically, the association between gender-conformity pressure and study motivation, as well as the association between school context and academic functioning have not been researched. Consequently, to further strengthen the robustness of the discovered patterns, future research could complete these relations within the conceptual model.

To further assess the influence of gender aspects on early adolescents’ school experiences, other indicators of school functioning could be considered. Especially interesting in this regard is the influence on hard end points, such as GPA, dropout, grade retention or certificates. There are some indications that achievement-related variables would indeed be influenced by the gender aspects considered here. More specifically, Van Maele and colleagues (2014) showed that gender typicality assessed at the start of 7th grade predicted the chance of positive versus negative evaluations at the end of the school year. In line with the research discussed in this dissertation, students who scored lower on typicality had a higher chance of negative evaluations (i.e. having to repeat the year or change track). Analyses showed that this association was explained by a reduced sense of school belonging among these students. This echoes the research findings presented here, showing that self-perceived gender atypical students displayed lower school functioning largely because of a reduced sense of well-being. Other research could investigate whether similar explanations hold for other endpoints.
IMPLICATIONS

For the Scientific Community
From a scientific perspective, this dissertation has some relevant implications for other researchers. First, we showed that it is possible to research gender issues in a quantitative way in a school setting. More importantly, we demonstrated that it is feasible to do this in a non-binary way by going beyond simple boy-girl differences, and without losing sight of the nuanced ways in which gender operates. Consequently, we strongly recommend future researchers use quantitative methods to take gender, beyond biological sex, into account. Doing so will shed light on important variations within each gender, and open up the black box regarding why we find certain differences between boys and girls.

Second, we would like to highlight that different aspects of gender unveil different facets of the complex gendering of everyday life. Our research revealed, for instance, that gender typicality and gender-conformity pressure influence students’ experiences in compatible, though unique ways. This confirms that gender is a complex, multifaceted construct, influencing our lives in diverse and nuanced ways. Consequently, researchers who wish to explore the impact of gender on contemporary society would benefit from considering more than one indicator.

Similar suggestions can be made for research concerning heteronormativity. This research revealed that homonegativity and traditional gender ideology are equally valid indicators of heteronormativity at school. Chesir-Teran (2003) already suggested studying several indicators of institutional heteronormativity, including architectural elements, policy features, social interaction and suprapersonal factors (the latter includes organizational cultures, as considered in our research). If other researchers reveal that additional indicators of heteronormativity weigh upon students’ experiences in the same way, a strong case could be made for heteronormativity as a single latent construct. Furthermore, previous research has focused predominantly on LGB students (e.g., Chesir-Teran & Hughes, 2009; Goodenow, Szalacha, & Westheimer, 2006; Szalacha, 2003; Toomey & Russell, 2013; Wilkinson & Pearson, 2009). Our research revealed that heteronormativity impacts upon all students’ well-being, regardless of sex, gender-divergence or sexual orientation. Consequently, it seems important to extend this research to broader populations, and investigate if these patterns hold true in different regions and for different outcomes.
For the Educational Community

This research revealed several issues regarding gender at school. First, atypical students, both boys and girls, tend to feel rather bad about themselves, which impedes their school functioning. Second, study-related activities seem to be more compatible with feminine identities than masculine ones. This construction of studying as feminine contributes to anti-school cultures among boys. Third, gender-conformity pressures lead to reduced well-being among all students. Furthermore, such conformity-pressure are central in socializing students into gender-specific study cultures, which are especially detrimental to boys.

We developed an overarching program to combat heteronormativity at school (discussed below), which might help to reduce these issues. This program suggests several ways in which the educational community can reduce traditional gender binaries at school. Importantly, schools would benefit from implementing as many suggestions as possible, because only when a holistic change is made can real changes in students’ school experiences be expected (Roberts & Marx, 2015; Szalacha, 2003). While the following suggestions are relevant for any level of schooling, our empirical results do suggest that the start of secondary education might be especially crucial, insofar as we found that students are socialized into gender-specific study cultures through gender-conformity pressures in the course of 7th grade. Consequently, we would especially recommend educational practitioners in middle school to adopt the following suggestions.

Policy. The first action is located at the policy level, by taking a critical look at school regulations. One important part of these is the anti-bullying policy. Explicitly including gender-based bullying in these policies helps to communicate the importance of these issues to both students and staff (Hong & Garbarino, 2012; Toomey, McGuire, & Russell, 2012). This also promotes acceptance of students who defy traditional gender norms, and offers them a safer and more comfortable school environment (Poteat et al., 2014a). Such interventions could be crucial in heightening the well-being of gender atypical students, since students defying traditional gender norms typically report more school-based victimization (Stoudt, 2006; Toomey et al., 2012). Simultaneously, it might help to reduce high gender-conformity pressures at school, since gender-based teasing and name-calling is often used as a way to police “unacceptable” behavior, especially among boys (Collier, Bos, & Sandfort, 2013; Pascoe, 2007).

Nevertheless, a significant number of students remain unaware of school policies, and policies are not always consistently enforced (Chesir-Teran & Hughes, 2009; Szalacha, 2003). Because of this, steps should be taken to ensure the actual
implementation of these anti-bullying policies. Teachers in particular play a pivotal role in this respect. Research has shown that teachers often fail to intervene when confronted with homonegative bullying (Elze, 2003), yet other studies have demonstrated that when teachers consistently intervene in sexual-orientation based bullying, LGB students tend to feel safer at school (Elze, 2003; Goodenow et al., 2006). Similarly, girls’ perceptions of teachers ameliorate when teachers confront sexist remarks in the classroom (Boysen, 2013). Therefore, it is advisable to organize workshops and clear guidelines for teachers on how to handle —gender-based— bullying.

Besides policies concerning bullying, school regulations regarding dress and accessories are seen as important symbolic signifiers of traditional gender binaries (Connell, 1996; Pascoe, 2007). Especially striking in such regulations are sex-specific proscriptions, which consider certain items of clothing as more problematic for one gender than the other. Common examples of these are long hair, skirts and earrings for boys, while for girls, there are often restrictions on necklines, skirt or sleeve length. Note that these proscriptions concern exclusions of traditionally feminine items for boys, whereas the underlying issue with girls’ clothing seems to revolve around decency and sexuality (Whisner, 1982). Underlying such regulations are views on “appropriate” performances of masculinity and femininity. Such ideas are transmitted to students, telling them that femininity is unacceptable in boys, whereas girls displaying certain parts of their body are indecent. If schools are serious about tackling heteronormativity, such sex-specific regulations can easily be changed. Nevertheless, doing so demands the courage to challenge the core of societal heteronormativity, that is, issues appearing so thoroughly commonsense that they seem ridiculous to question.

**Curriculum.** A second point is the curriculum. A gender-inclusive curriculum normalizes diversity in gender expressions, helping students to accept people defying traditional gender notions both at school and in broader society (Poteat et al., 2014a). Furthermore, providing more possibilities to identify with the curriculum encourages school engagement among gender-nonconforming and LGBT students. Moreover, by actively promoting acceptance of diversity and by breaking through traditional gender notions, conformity-pressures among students may be lowered as well. Consequently, studies have shown that inclusive curricula contribute to safer school environments for gender non-conforming students (Toomey et al., 2012) and less school victimization among LGB students (Chesir-Teran & Hughes, 2009).
Schools can promote a gender-inclusive curriculum in several ways (Poteat et al., 2014a). A first step is to critically consider the illustrative material (çavaria, 2011a; RoSa, 2010). For instance, which people are represented in handbook pictures, and do these give a fair representation of societal diversity? This should not just be considered from a gender perspective, but include other categories of diversity. For instance, are people of color, women, disabled people, people of different ages and in several family relations present in the illustrative material? Besides the mere presence of diversity, the way people are represented is essential as well. An analysis of teaching materials revealed that men were presented in a professional setting three times more often than women, while women were more often represented in passive situations (çavaria, 2011b). Conversely, men are only rarely presented as care-givers or in other norm-breaching professions.

Besides the illustrative material, language-use in exercises is also important (çavaria, 2011a; Poteat et al., 2014a; RoSa, 2010). For instance, a mathematical problem might include the description of a non-nuclear family, or might feature people with non-native names. It is similarly important to use neutral language in the formal communication of the school, by for instance, addressing letters to the parents and guardians, rather than “mom and dad”.

Next to these symbolic representations, actual curriculum content is crucial. For instance, it is important to highlight the contributions of women, people of color and LGBT’s in the course material of mathematics, chemistry, history, literature, art and so on (çavaria, 2011a; Poteat et al., 2014a; RoSa, 2010). The work of, for instance, Marie Curie or Oscar Wilde is not only highly relevant, but also easily integrated in the regular curriculum. Similarly, when the holocaust is discussed, the liquidation of disabled, gay and Roma people can be considered alongside the extermination of the Jewish community. Similarly, it is important to normalize gender and sexuality as topics of discussion, for instance by including these in sex education as well as religion and morality classes.

**Classroom practices.** Third, how teachers behave in front of the class is crucial. It does not suffice that the curriculum breaks through gender-norms, teachers should integrate this in their classroom practices as well. For instance, a lot of teachers consider girls to be better students (Beaman, Wheldall, & Kemp, 2006; Jones & Myhill, 2004). Such generalized expectations are not without consequences, since they help to create self-fulfilling prophecies and contribute to biases in what teachers —fail to— notice. For instance, research has shown that boys are subject to disproportionate rates of discipline (Consuegra et al., 2015), while struggling girls go unnoticed (Benjamin, 2003; Jones & Myhill, 2004). In
trying to keep boys on-task, teachers often interact more with boys by asking individual questions and allowing them to answer more frequently (Beaman et al., 2006; Einarsson & Granström, 2002; Younger, Warrington, & Williams, 1999). Girls, on the other hand, are punished less often when talking to friends (Consuegra et al., 2015), or get away with rule-breaking by exaggerating “good girl” behavior or flattering the teacher (Ohrn, 1993; Younger et al., 1999).

Most teachers are unaware of any biases in their teaching and believe they treat all students equally (Garrahy, 2001; Younger et al., 1999). Unfortunately, gendered teaching is notoriously hard to change, because it is a largely unconscious process, and the stereotypes on which these biases are based are deeply entrenched in our minds. Nevertheless, some things might be done to help teachers recognize gender biases in their classroom practices. First, a seminar discussing gender-sensitive teaching and the gender gap in education might highlight intrasexual variations in attainment and classroom behavior. Such a seminar can make these topics discussable among the teaching staff, and might help to create the necessary support for the policy and curriculum changes discussed above. Such seminars can also contribute to an open mind-set towards self-improvement and continuous learning, which might promote school practices such as asking colleagues into the classroom to observe and give feedback.

A more easily changeable, yet influential practice in everyday classroom and school organization is gender-mixing. That is, research has shown that gender segregation contributes to gender-stereotyped behavior among students (Martin et al., 2014) and using gender as a divisional marker heightens its salience and relevance in the eyes of students (Bigler, 1995; Hilliard & Liben, 2010). Consequently, educational practitioners should avoid using gender as an organizational tool for group work assignments, competitions, or indeed, school organization.

Additionally, teacher training in autonomy-supportive class methods might be beneficial in combating the educational gender gap. That is, our studies showed that boys tended to score lower on autonomous motivation than girls (see chapter 9). We suggested that this might be due to a vicious circle, in which boys’ disruptive class behavior —triggered by masculine anti-school cultures— evokes more controlling methods among teachers. Such controlling teaching methods in turn reduce autonomous motivation among students (Assor et al., 2005; Reeve, 2009). By training teachers in autonomy-supporting methods and handling disruptive class behavior, such vicious circles might be broken.
Concluding remarks, or what not to do. It should be noted that some of the suggestions above will probably be more beneficial to some students than others. More specifically, autonomy-supportive teaching methods are likely to have a greater impact on boys, simply because they are more subject to controlled teaching (Consuegra et al., 2015). Similarly, gender atypical and LGB students stand more to gain than other students with school policies and curricula that challenge heteronormativity. Nonetheless, all suggestions mentioned above remain beneficial to the entire school population by providing a supportive and accepting school environment. For instance, autonomy-supportive teaching methods are beneficial for girls as well as boys (Assor et al., 2005; Reeve, 2009), and as shown in chapter 10, gender-equal school cultures promote higher levels of well-being for all students. Hence, while some suggestions might benefit certain groups of students more than others, simply because they are in more disadvantaged positions, all suggestions contribute to more optimal school environments for the whole student population.

This is important, because interventions which have tried to combat the educational gender gap by proposing separate schools, specific course materials or pedagogies geared towards boys, tend to have some negative side effects. First, schools should strive to allow all students to achieve their full potential. Strategies predominantly focused on one subgroup risk harming the educational opportunities of other groups. Furthermore, such interventions are often rooted in stereotypical ideas about gender differences (Carrington & McPhee, 2008; Martino & Kehler, 2006). Certain topics or learning styles are assumed to be predominantly masculine or feminine, leading to suggestions to focus on sports, cars, music, and games to pique boys’ interest. And because boys would not be able to sit still or pay attention for a long time, things should be presented in active or competitive ways (Reints, 2013). Such suggestions reconfirm societal stereotypes and promote a view of essential differences between the sexes in the process. Clearly, such interventions ignore the vast amounts of intrasexual variation in interests and behavior (Martino, 1996). That is, girls might very well be interested in “boyish” things, something such interventions often fail to consider. Similarly, for each boy that might conform to the previously cited stereotype, there is another one who holds completely different interests or displays very different behaviors. In our evidence-based opinion (Bigler et al., 2014; Heemskerk, van Eck, Kuiper, & Volman, 2012; Lindsay & Muijs, 2006), schools and students would be better served by an approach focused on inclusivity and diversity. This involves offering each student an optimal learning environment, suited to his or her individual needs and
challenges, rather than working from stereotypical ideas on what boys and girls would or should need.

It should be noted that combatting heteronormativity at school is not a simple task (Tharinger, 2008). A school is not a social vacuum. Students, staff, and parents bring with them to school beliefs about how men and women differ and how they should differ. Such ideas are indications of societal-wide paradigms, and it would be naive to think that schools can singlehandedly change a whole society. Nevertheless, schools remain important agents for social change, and are singularly well suited to broaden students’ minds, confront them with new ideas, and shape them into active, well-informed and constructive citizens. In our opinion, programs aimed at reducing heteronormativity at school could be a valuable part of this mission, considering that such programs are aimed at promoting equal rights and opportunities for all, without discrimination on the basis of sex or sexual orientation—a valuable goal in any democratic state. Helping people become accepted for who they are and grow to their full potential seems an important goal, and yes, even a duty for schools and society as a whole to fulfill. We can only hope that our work, presented here in this dissertation, might be a small contribution to this goal.
References


---

206


Hull, G. T., Bell Scott, P., & Smith, B. (1982). All the women are white, all the Blacks are men, but some of us are brave: Black women’s studies. Old Westbury, NY: The Feminist Press.


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. doi: 10.1080/02671522.2014.970226


(SST)’s key assumptions about sex-distinct mating mechanisms. *Sex Roles, 64*(9-10), 629-643.


Reeve, J. (2009). Why teachers adopt a controlling motivating style toward students and how they can become more autonomy supportive. *Educational Psychologist, 44*(3), 159-175.


Summary

The discovery that girls are consistently outperforming boys throughout the school career, called the educational gender gap, has garnered a lot of attention from both media and researchers. Most of this research has stayed stuck on average achievement differences between boys and girls, while the intrasexual variation and the more subtle functioning of gender have been neglected. Consequently, in this dissertation, we focused on uncovering inter- and intrasexual variation in the school functioning of boys and girls by using a gender focus.

First, spurred by insights from general gender theories, masculinities theory, social identity theory and multifactorial theory, we decided to tap the influence of gender norms by considering gender typicality and pressure for gender conformity. Gender typicality refers to the extent to which people feel similar to others in their gender category, while gender-conformity pressure taps the extent to which people experience pressure to uphold gender norms. The concept of gender typicality is situated on the intra-individual level, while gender-conformity pressures are located more on the interactional level. By using these two concepts, we employed insights from both the socio-psychological and sociological field.

Second, we operationalized school functioning in line with the expectancy-value model by considering study motivation, academic self-efficacy and subjective well-being. By using several indicators of school functioning, a more robust overview of the impact of gender norms on students’ experiences could be given. Furthermore, because the educational gender gap tends to widen at the start of secondary education, we assessed these associations in a sample of Flemish early adolescents.

First, we considered the influence of gender typicality on academic self-efficacy and study motivation. The results overlapped extensively for autonomous study motivation and academic self-efficacy, suggesting that gender typicality has a robust influence on academic school functioning. More specifically, gender typical girls score the highest of all students on both academic self-efficacy and autonomous motivation. Typical boys score notably lower than their female counterparts, while atypical students score the lowest on cognitive school functioning. These findings are in line with previous research suggesting that stereotypical feminine characteristics (such as being tidy, punctual, communicative, compliant and passive) are more compatible with a student role than stereotypical masculine characteristics. Furthermore, self-perceived atypical boys and girls score equally low
on academic school functioning. This reduced academic functioning is largely explained by subjective well-being, which suggests that gender atypical students tend to feel rather bad about themselves, and that this weighs upon their academic functioning. These findings are consistent with ethnographic studies that showed how adolescents who do not live up to gender expectations are frequently teased and bullied by peers. Our results add to these findings by directly connecting these welfare issues among gender-divergent adolescents to school functioning. Additionally, since atypical boys and girls score equally low on cognitive school functioning, the results show that these processes work in a similar fashion for boys and girls. Hence, despite most research’s focus on boy-girl differences, considering more nuanced gender aspects in educational research can uncover general rather than sex-specific processes.

Second, we assessed the influence of gender-conformity pressures on students’ academic self-efficacy. The results showed opposite associations between gender-conformity pressure and academic self-efficacy for boys and girls. When boys experience higher rates of gender-conformity pressure, their academic self-efficacy tends to suffer. Conversely, girls’ academic self-efficacy is much less affected by their level of conformity-pressure. Only when controlled for well-being is a positive effect of gender-conformity pressure on girls’ self-efficacy revealed. These findings are in line with previous research, which showed that anti-school cultures are more prevalent among boys, while girls’ peer groups tend to have more positive attitudes toward school. Importantly, 7th grade seems to be a crucial time-point for socialization into these gender-specific study cultures. That is, the associations between gender-conformity pressure and self-efficacy are not present at the start of 7th grade, but arise in the course of the school year. Consequently, interventions trying to disrupt these gender-specific study cultures might benefit from focusing on the first year of secondary education. Furthermore, despite the especially detrimental effect on boys’ school functioning, reducing gender-conformity pressures at school would be beneficial for all students due to its adverse impact on well-being.

In the last study, we zoomed out from the individual and interactional level to focus on the institutional level by considering school cultures. More specifically, we assessed the influence of heteronormative school cultures on the well-being of students. The results showed that schools where traditional views concerning gender and sexuality reign, are detrimental to the well-being of any student, regardless of sex, gender-divergence or sexual orientation. Moreover, this study
suggests that the quality of the relationships with the teacher is decisive in accounting for these well-being differences across school contexts. This implies that teacher-student relationships are more strained in heteronormative school environments, leading to the observed well-being deficit in these schools. We have suggested several ways in which the educational community might challenge such heteronormative cultures. Adopting a gender-inclusive curriculum, enforcing anti-bullying policies, paying attention to gender-mixing and gender-biased teaching in the classroom, and taking a critical look at gendered school regulations could go a long way in making schools more welcoming environments for all students.

All in all, this dissertation demonstrates the relevance of considering gender aspects in educational research besides sex. The studies show that gender norms tend to have a very real impact on students’ well-being and academic school functioning, regardless of whether these norms were internalized into a sense of identity, or were experienced as an external pressure. Interestingly, this research suggest that traditional masculine or feminine characteristics are not unambiguously connected to adverse or beneficial outcomes. Rather, the underlying motivation seems crucial in determining the consequences of gendered behavior. Volitional behavior (as evidenced by gender typicality) is largely connected to higher well-being and superior academic school functioning, while gender-conformity pressures are consistently linked to school functioning deficits. An important sociological nuance in this largely socio-psychological process is that environmental responses can augment or attenuate the consequences of volitional versus externally-motivated conduct. More specifically, the studies showed that the positive consequences of feminine conduct are boosted in the Flemish educational context, while the adverse consequences of traditional masculine enactment seem to be augmented. In line with other research, this suggests that traditional femininity seems more compatible with a school context, while a mismatch exists between traditional masculine enactments and school expectations.
Samenvatting

De bevinding dat meisjes beter presteren dan jongens gedurende de volledige schoolloopbaan, de zogenaamde genderkloof in onderwijs, heeft veel aandacht gekregen van zowel de media als onderzoekers. Het meeste van dit onderzoek is blijven steken bij de constatatie van prestatieverschillen tussen jongens en meisjes, terwijl intraseksuele variatie en de meer subtiele werking van gender werden verwaarloosd. In dit doctoraat hebben we ons daarom, met behulp van een gender-focus, toegespitst op het blootleggen van inter- en intraseksuele variatie in het schoolfuncteren van jongens en meisjes.

Ten eerste hebben we besloten om, gebaseerd op inzichten vanuit algemene gendertheorieën, masculinities theorie, social identity en multifactoriële theorie, de invloed van gendernormen te onderzoeken via gender-typicaliteit en druk tot genderconformiteit. Gender-typicaliteit verwijst naar de mate waarin mensen zichzelf vergelijkbaar vinden met anderen in hun gendercategorie. Druk tot genderconformiteit verwijst naar de mate waarin mensen druk ervaren om gendernormen te handhaven. Het concept van gender-typicaliteit situeert zich op het intra-individuele niveau, terwijl druk tot genderconformiteit zich eerder bevindt op een interactie-niveau. Door het gebruik van deze twee begrippen hanteren we inzichten uit zowel de sociale psychologie als de sociologie.

Ten tweede werd, in lijn met het “expectancy-value” model, schoolsfunctioneren geoperationaliseerd als studiemotivatie, academische self-efficacy en subjectief welbevinden. Door verschillende indicatoren van schoolsfunctioneren te gebruiken, konden we een betere inschatting maken van de invloed van gendernormen op de schoolervaringen van leerlingen. We onderzochten deze verbanden tussen gendernormen en schoolsfunctioneren bij jonge Vlaamse adolescenten, aangezien de genderkloof in het onderwijs verbreedt bij de aanvang van het secundair onderwijs.

Wanneer we de invloed van gender-typicaliteit op academische self-efficacy en studiemotivatie nagaan, zien we dat de resultaten grotendeels overlappen. Dit suggereert dat gender-typicaliteit een robuuste invloed heeft op academisch functioneren. Gender-typische meisjes scoren het hoogst op zowel autonome motivatie als academische self-efficacy. Typische jongens scoren aanzienlijk lager dan hun vrouwelijke evenknieën, terwijl gender-atypische studenten het laagst scoren op academisch functioneren. Deze bevindingen komen overeen met vroeger onderzoek dat suggereert dat stereotiep vrouwelijke kenmerken (zoals ordelijkheid,

Ten tweede hebben we de invloed nagegaan van druk tot genderconformiteit op academische self-efficacy. De resultaten toonden tegenovergestelde verbanden voor jongens en meisjes. Bij jongens neemt het gevoel van self-efficacy namelijk af wanneer zij meer druk tot genderconformiteit ervaren, terwijl de self-efficacy van meisjes daarentegen veel minder beïnvloed is door genderconformiteitsdruk. Pas wanneer we controleren voor welbevinden, zien we een positief effect van druk tot genderconformiteit op de self-efficacy van meisjes.

Deze bevindingen zijn in lijn met voorgaand onderzoek waaruit bleek dat anti-schoolcultures meer voorkomen bij jongens, terwijl meisjes doorgaans een positievere houding aannemen ten opzichte van school. Het is belangrijk op te merken dat het eerste jaar middelbaar een cruciale periode lijkt voor de socialisatie in deze geslachtsspecifieke studieculturen. De associaties tussen druk tot genderconformiteit en self-efficacy waren namelijk niet aanwezig bij de start van het eerste jaar middelbaar, maar verschenen in de loop van dit schooljaar. Interventies gericht op het aanpassen van dergelijke geslachtsspecifieke studieculturen zouden er dus goed aan doen zich te focussen op het eerste jaar van het secundair onderwijs. En hoewel voornamelijk het schools functioneren van jongens lijdt onder druk tot genderconformiteit, heeft een dergelijke druk een nadelige invloed op alle leerlingen door de negatieve invloed op welbevinden.
In de laatste studie namen we schoolculturen in beschouwing, waarbij we dus uitzoomden van het individuele en interactie-niveau naar het institutionele niveau. We hebben in het bijzonder gekeken naar de invloed van heteronormatieve schoolculturen op het welbevinden van leerlingen. De resultaten tonen aan dat scholen waar traditionele visies op gender en seksualiteit overheersen, schadelijk zijn voor het welbevinden van alle leerlingen, ongeacht hun geslacht, seksuele geaardheid of variatie in gender-typicaliteit en druk tot genderconformiteit. Bovendien blijkt uit de resultaten dat de relatie met de leerkracht een doorslaggevende factor is voor de schoolverschillen in welbevinden. Dit suggereert dat leerkracht-leerling-relaties meer precair zijn in heteronormatieve scholen, wat op zijn beurt leidt tot inferieur welbevinden op deze scholen. In dit doctoraat hebben we verschillende manieren voorgesteld waarop de onderwijsgemeenschap dergelijke heteronormatieve culturen kan bestrijden. Het opnemen van een gender-inclusief curriculum, handhaven van anti-pest maatregelen, aandacht schenken aan gender-bias en organisatie in de klas, en een kritisch oog werpen op het schoolreglement zouden belangrijke stappen kunnen zijn om scholen meer veilige omgevingen te maken voor alle leerlingen.

Kortom, dit doctoraat toont het belang aan van niet enkel geslacht, maar ook gender mee op te nemen in onderzoek. Onze studies tonen namelijk dat gendernormen een reële impact hebben op het welbevinden en academisch functioneren van alle leerlingen, ongeacht of deze normen werden geïnternaliseerd tot een identiteit, of ervaren werden als een externe druk. Interessant is dat de resultaten suggereren dat traditioneel mannelijke of vrouwelijke kenmerken niet eenduidig verbonden zijn met negatieve of positieve gevolgen. Integendeel, de onderliggende motivatie van het gegenderde gedrag lijkt doorslaggevend te zijn. Wanneer gedrag intrinsiek gemotiveerd wordt (zoals het geval is bij gender-typicaliteit), gaat dit grotendeels samen met hoger welbevinden en beter academisch functioneren. Druk tot genderconformiteit, daarentegen, is consequent gekoppeld met negatieve gevolgen. Een belangrijke sociologische nuance in dit sociaalpsychologisch proces is dat de omgevingsreacties de gevolgen van vrijwillig versus afgedwongen gedrag kunnen versterken of afzwakken. De studies toonden namelijk dat de positieve gevolgen van stereotiep vrouwelijk gedrag versterkt worden in de Vlaamse onderwijscontext, terwijl de negatieve gevolgen van traditionele mannelijkheid toenamen. Dit suggereert, in overeenstemming met voorgaand onderzoek, dat traditionele vrouwelijkheid meer compatibel lijkt met een schoolse context, terwijl er een mismatch bestaat tussen traditionele mannelijkheid en schoolverwachtingen.