A teacher like me: the role of teacher gender representation and gender stereotypes in education

Doornkamp, L.

CHAPTER **FOUR**

Understanding the symbolic effects of gender representation: A multi-source study in education

Previously published as: Doornkamp, L., Groeneveld, S., Groeneveld, M.G., Van der Pol, L.D., & Mesman, J. (2022): Understanding the Symbolic Effects of Gender Representation: A Multi-Source Study in Education, *International Public Management Journal*, DOI: 10.1080/10967494.2021.1991534

ABSTRACT

Symbolic effects of bureaucratic representation assume that outcomes for citizen-clients change in response to the mere presence of bureaucrats with similar backgrounds. The social-psychological mechanisms in clients that may explain these changes are barely examined, though. Based on multi-source data on male and female high school students in the Netherlands, this study empirically tests a theoretical model that links student-teacher gender congruence to students' performance in math and Dutch language through the mediating role of students' gender-stereotypical beliefs and self-perceived abilities. The empirical analyses do not support the hypotheses. For student performance in math, no effects were found, while for Dutch language, a negative gender-representation effect was found for male students. In all, the results point at divergent relations between gender congruence, self-perceived abilities and performance for male and female students across subjects. The null results in this study suggest that symbolic effects of gender representation may be affected by national context and socio-economic status which should be considered as new frontiers for future research on representative bureaucracy.

4.1 INTRODUCTION

Despite the major expansion in studies on representative bureaucracy in the past decades, little is known about the individual-level mechanisms explaining the effects of representation in public organizations (Meier, 2019). Representativeness of public organizations is associated with inclusiveness in decision-making which contributes to the legitimacy and effectiveness of public policy (e.g., Sowa & Selden, 2003). The majority of empirical studies on representation effects in public organizations typically focuses on an aggregate level, modeling increased output for a group of clients based on the percentage of employees in the public organization from that same group. This aggregate focus fails to identify the individual-level mechanisms that drive representation effects (Meier, 2019; Nicholson-Crotty et al., 2016; Zhang, 2019). For decades, research assessments have found the existence of representation effects and subsequently ascribed it to *assumed* mechanisms, mostly situated at the individual level (Nicholson-Crotty et al., 2016).

The study of individual-level mechanisms has intensified in recent years. However, these studies have mostly taken the perspective of the bureaucrat, focusing on the behavior of public officials to explain representation effects (e.g., Bardbury & Kellough, 2011; Keiser et al., 2002; Nicholson-Crotty et al., 2016; Sowa & Selden, 2003; Wilkins & Keiser, 2006; Wilkins, 2007). Only a handful of studies have taken the perspective of the client, explaining representation effects in terms of changes in attitudes and behaviors of the client as a result of the mere presence of a bureaucrat from a similar background (Guul, 2018). Clients matched with a bureaucrat sharing the same characteristics were found to rate higher program satisfaction (Gade & Wilkins, 2013), increase job-seeking efforts (Guul, 2018), perceive bureaucratic actions as more legitimate (Theobald & Haider-Markel, 2008) and as more trustworthy and fair (Riccucci et al., 2014). Although these studies measured changes in attitudes and behaviors in clients, the underlying (un- or subconscious) psychological processes producing these changes remain unknown. In this study, we theorize and empirically test a psychological mechanism on the individual level from the perspective of the citizen-client that so far has been under researched (Figure 4.1). In doing so, we try to disclose one of many possible pathways for symbolic representation, contributing to the understanding of representation effects.

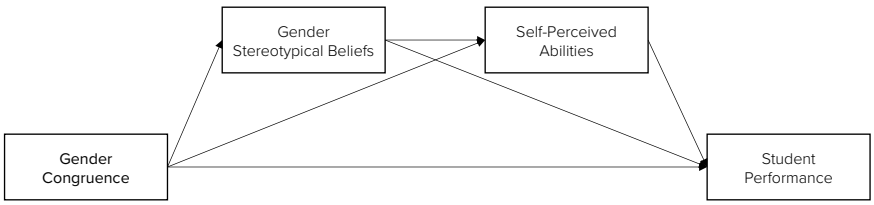

Figure 4.1. Research Model

Like many previous studies on representative bureaucracy (e.g., Dhillon & Meier, 2020; Grissom et al., 2015; Meier & Stewart, 1992; Meier et al., 1999; Nicholson-Crotty et al., 2016; Vinopal, 2018), we empirically test our proposed model in the context of schools, more specifically in Dutch secondary education. We focus on the effects of gender representation on male and female high school students' performance in the subjects of math and Dutch language. Previous studies on this topic, that focused on the effect of female teachers on female students' math performance, found positive effects of gender congruence (Agyapong, 2018; Keiser et al., 2002; Song, 2018; Zhang, 2019). We added to this stream in the literature by means of a survey-experimental study on the effect of the gender of a math teacher on stereotypical attitudes and perceived abilities of male and female students, finding an effect of brief hypothetical exposure to a female math teacher on both male and female students' academic self-concept (Doornkamp et al., 2019). The current study builds on this study, elaborates on its theoretical model and tests new hypotheses based on multi-source data including data on student performance of 329 high school students and their teachers in the Netherlands. In doing so, it acknowledges that empirical research should include real-life and enduring interactions to be able to detect attitudinal and behavioral changes in citizens-clients, if any. Our research is guided by the question: *what psychological processes in students can explain the association between student-teacher gender congruence and student performance?*

This study's contribution is threefold. First, this study empirically tests a theoretical mechanism underlying symbolic representation in which the psychological processes in the citizen-client, not the bureaucrat, are central. Second, in so doing, this study introduces social cognitive theory to representative bureaucracy literature which provides a basis for the study of social-psychological mechanisms in clients to explain representation effects in a variety of bureaucratic contexts. Third, while effects of symbolic representation are often tested through survey-experimental research, if tested at all, this study uses multi-source observational data that improves the ecological validity of the research.

The paper proceeds as follows: the next section elaborates on the theoretical framework. Then the study context as well as its research design and methods are described, followed by the study's results. The paper concludes with the discussion of the findings and outlines avenues for future research.

4.2 REPRESENTATIVE BUREAUCRACY

The concept of representative bureaucracy was coined by Kingsley (1944), who argued that bureaucracies are part of the democratic constitutional state and should therefore comply with democratic principles. Consequently, bureaucracies' demographic composition *should* reflect the demographic composition of the people they serve. Besides the normative argument for representative bureaucracies, a large body of research found that

public organizations with high levels of representativeness perform better than their less-representative counterparts (see Kennedy, 2014; Meier, 2019 for a review of normative and theoretical underpinnings, see Groeneveld & Van de Walle, 2010; Riccucci & Van Ryzin, 2017 for state of the art).

The representative bureaucracy literature distinguishes between passive, active, and symbolic representation. Passive representation refers to the extent to which the socio-demographic composition of the workforce of a public organization mirrors its clientele. Research on passive representation shows disparity in hierarchical and functional representation for minorities. Moreover, public organizations with higher levels of minority representation were found to serve important democratic values, meet fairness and equality goals, and perform better (Riccucci & Van Ryzin, 2017). In schools, research on passive representation has found that representativeness of the workforce is associated with fewer drop outs (Pitts, 2007), more inclusive school policies (Roch et al., 2010), and better performance for female and ethnic minority students (Agyapong, 2018; Dhillon & Meier, 2020; Keiser et al., 2002, Meier & Stewart, 1992; Meier & Bothe, 2001).

These outcomes of representativeness are generally assumed to originate on the individual level through the interaction between the bureaucrat and the citizen-client. Most studies that focus on this individual level take the perspective of the bureaucrat, explaining representation effects in terms of a bureaucrat acting in concordance with the values or interests of a client because of their similar social origins and shared values. This is referred to as active representation (e.g., Meier et al., 2006; Meier et al., 2005; Park, 2012). For bureaucrats to be able to engage in this minority representative role-taking behavior, bureaucrats need discretionary room and must experience salience of the minority status to the policy area at hand (Keiser et al., 2002).

In education, teachers have discretion in shaping and teaching their classes. Additionally, characteristics such as ethnicity and gender are salient as ethnic minorities and women are historically disadvantaged in education (Keiser et al., 2002). Research on representativeness in education shows that contextual factors such as more class days, longer teaching hours, and smaller student-teacher ratios enhance representation effects assuming active representation plays a role (Dhillon & Meier, 2020; Song, 2018). Minority teachers may spend more time on the students they represent and generate more positive reinforcements (Grissom et al., 2015; Keiser et al., 2002; Zhang, 2019). Moreover, Nicholson-Crotty et al. (2016) demonstrated that minority teachers were more likely to assign gifted services to minority students than their nonminority counterparts.

Symbolic representation shifts the focus from the bureaucrat to the citizen-client. Representation effects are explained in terms of changes in the attitudes and behaviors of the citizen-client as a result of the mere presence of the bureaucrat (Riccucci et al., 2018). Some studies empirically examined and found evidence for changes in attitudes and behaviors of the citizen-client as a result of being matched with a bureaucrat sharing the same characteristics as them (Gade & Wilkins, 2013; Guul, 2018; Riccucci et al., 2014;

Theobald & Haider-Markel, 2008). While studies on representation that were situated in education have not been able to empirically unravel the effects of symbolic and active representation, some studies suggest symbolic representation is likely to occur. A qualitative study in education in China conducted by Xu and Meier (2021) suggests that teachers do not actively pursue the interests of the students they represent and that representation effects in education in China primarily result from symbolic representation. Vinopal (2018) found that parents of minority students were more involved in the presence of a minority teacher than in the presence of a nonminority teacher which could be a result of both symbolic and active representation. Zhang (2019) found a small and partial effect of female teachers on female students' math performance mediated by their gender-stereotypical beliefs. While Zhang (2019) interprets this result as evidence for active representation, this may also reflect effects of both active and symbolic representation. In a previous study in Dutch education we added to this stream in the literature by measuring attitudinal changes in students by means of a vignette experiment that allowed to single out the effect of symbolic representation. Although we found that students' academic self-concept is enhanced in the presence of a female teacher, this was found for both male and female students (Doornkamp et al., 2019).

Although studies on gender representation in education have mostly focused on effects of gender congruence on female students, the logic of representation might also apply to male students. Although Zhang (2019), in his study on gender representation in Chinese middle schools, did not find representation effects for male students in math, English, or Chinese scores, studies on the gender gap in education have found positive effects of same-sex teachers on the academic performance of both female and male students (e.g., Dee, 2005).

Based on the above, the first hypothesis can be formulated:

> **H1:** Student-teacher gender congruence has a positive effect on student performance.

Despite increased attention for symbolic representation in the study of representative bureaucracy, still little is known about *how* symbolic representation actually works. How do attitudinal and behavioral changes emerge in the citizen-client as a result from the presence of a bureaucrat with whom one shares similar socio-demographic characteristics? This is theoretically explored in the next section.

4.2.1 SOCIAL COGNITIVE PROCESSES AS MECHANISM UNDERLYING SYMBOLIC REPRESENTATION EFFECTS

Social-cognitive theory (SCT) emphasizes the role of the environment in a person's knowledge acquisition and personal development; by observing and interacting with others people learn about social norms and structures that influence attitudes, beliefs, and

behaviors. As this study focusses on gender representation (i.e., the influence of the student-teacher gender congruence on students' performance) we will discuss SCT in the context of gender development. Note that SCT can be applied in several contexts and therefore might contribute to the understanding of bureaucratic representation more broadly.

SCT of gender development emphasizes the role of the environment in a person's knowledge acquisition about the sexes to identify males and females, and consequently themselves (Blakemore et al., 2009). By observing and interacting with men and women an individual learns about existing gender norms and rules. Bussey and Bandura (1999) distinguish three ways through which an individual learns: direct information, social consequences to gender related behaviors, and modeling. The last one, modeling, is the most pervasive and powerful as it is the fastest to cognitively process (see Bussey & Bandura, 1999 for an extensive explanation). In short, by observing models (i.e., men and women in the environment) people learn about what is possible and appropriate for men and women. This gender-linked information generates new forms of attitudes, beliefs, and behavior that conform to these gender rules and structures (Bussey and Bandura, 1999).

Following SCT reasoning, symbolic gender representation can be expressed in terms of bureaucrats functioning as a model influencing a client's conception of what is appropriate and possible for him or her as a man or a woman changing his or her attitudes, beliefs, and behavior producing representation effects. Just by being a man or woman, each and every model (including bureaucrats) in the environment of a client can have a greater or lesser impact on a client's gender development. Applied to the school context, teachers can pass along information about what is academically appropriate and possible for male and female students changing students' attitudes, beliefs and behaviors in education.

4.2.1.1 Self-Perceived Abilities

Important to SCT of gender development is the role of models on people's self-perceived abilities, i.e., believe in the capability to master given levels of attainment (Bussey and Bandura, 1999). The concept of self-perceived abilities is often used to grasp self-efficacy beliefs in an educational setting. Self-perceived abilities of students are found to be strongly associated with academic achievement (Marsh et al., 2005). Although the causality in this relation does not work in one direction (self-perceived abilities can predict academic achievement and vice versa), studies show the predictive power of self-perceived abilities in the prediction of academic achievement (Chamorro-Premuzic et al., 2010, Spinath et al., 2006). A large body of research shows that an individual's beliefs and judgements about his or her own capabilities to perform a certain academic task (i.e., self-perceived abilities), can predict actual performance in that task (e.g., Marsh et al., 1988; Chamorro-Premuzic et al., 2010; Greven et al., 2009; Wach et al., 2015). Even when controlled for general mental ability (intelligence), self-perceived abilities contribute to the prediction of academic achievement (Spinath, et al., 2006; Greven et al., 2009; Wach et al., 2015). Moreover, motivational influences such as self-perceived abilities can be stronger predictors of

academic achievement than cognitive influences in some subjects (Weber et al., 2013).

The gender of a teacher can affect students' self-perceived abilities through their success of having become a teacher as a man or a woman, passing along gender-linked information about what the sexes are capable of in that subject. Social-cognitive theory argues that as soon as an individual is aware of his or her gender, success or failure of someone similar to themselves (man or woman) can influence the individual's beliefs and judgements to master comparable activities (Bussey & Bandura, 1999). As students are aware of their own gender, students can perceive the success of teachers similar to themselves as a representation of their own capabilities in a subject affecting their self-perceived abilities (Bussey & Bandura, 1999). These strengthened self-perceived abilities positively affect their actual performance in a subject (e.g., Wach et al., 2015). This leads to hypothesis 2.

> **H2:** Self-perceived abilities mediate the association between student-teacher gender congruence and student performance.

4.2.1.2 Gender-Stereotypical Beliefs

'Stereotypes reflect general expectations about members of particular social groups' (Ellemers, 2018: .276). Gender-stereotypical beliefs are an individual's ideas about attributes of men and women based on these generalized preconceptions (Bussey & Bandura, 1999). Gender-stereotypical beliefs in individuals emerge through the internalization of repeated modeling of gender-typed behavior in the environment of an individual (Bussey & Bandura, 1999; Spencer et al., 2016).

The gender of a teacher can affect students' gender-stereotypical beliefs by modeling a (contra) gender-stereotypical pattern strengthening or weakening students' gender-stereotypical beliefs connected to a subject. For math and Dutch languages, the subjects of the current study, stereotypes prescribe the capabilities for men and women; men are good math, women are good at Dutch language. Having a male or female teacher for a subject that is generally seen as either masculine (e.g., math) or feminine (e.g., language), provides students with a stereotypical or contra-stereotypical example. By confirming gender-prototypic behaviors (for instance a male math teacher) students' gender-stereotypical beliefs can become stronger, whereas models contradicting gender-prototypic behaviors (female math teacher) can reduce gender-stereotypical beliefs.

These stereotypical beliefs can affect performance through the internalization of stereotypes reflected in self-perceived abilities (Plante et al., 2013). Stereotypes can either facilitate or obstruct optimal performance (i.e., the stereotype 'men are good at math, women are bad at math' can facilitate male students' performance, whereas female students' performance can be obstructed), the negative obstructing effect is assumed to be easier to internalize and more resistant to disconfirmation than the positive facilitating effect of stereotypes (Baumeister et al., 2001). The psychological state in which students believe that they belong to a group stereotypically known for its performance (deficits) affects students'

beliefs and judgements about their own capabilities to succeed (i.e., self-perceived abilities) affecting their performance (Beilock et al., 2010; Spencer et al., 2016). Similarly, research on self-efficacy beliefs and academic achievement propose stereotypical beliefs as an explanation for gender differences in self-efficacy beliefs that cannot be explained based on prior achievement (Marsh et al, 1988; McFarland et al., 2016; Pajares & Miller, 1994; Wach et al., 2015).

Thus, math and Dutch language teachers can affect student performance not only through being perceived as a representation for what is achievable for the sexes, but also through being a negation or a confirmation of the stereotypes connected to the subject affecting students' gender-stereotypical beliefs and self-perceived abilities. Based on the above reasoning, we formulate hypothesis 3.

> **H3:** Stereotypical beliefs mediate the indirect effect of self-perceived abilities in the association between student-teacher gender congruence and student performance.

4.3 STUDY DESIGN AND METHOD

4.3.1 STUDY CONTEXT

This study is part of the longitudinal study Girls in Science running from 2017 until 2022 in the Netherlands.[1] Although the Netherlands ranks in the highest quartile of the Global Gender Gap Index Rating indicating that gender inequality is relatively mild (Global Gender Gap Report, 2020), gender-stereotypical beliefs still play an important role in Dutch society. While Dutch men and women are equal under the law and gender equality has improved over the last decades, Dutch society is still characterized by gender role segregation (Daalmans et al., 2017). For instance, educational and occupational careers of men and women in the Netherlands display gender-stereotypical patterns (Onderwijsraad, 2019). Female students are still underrepresented in technical and mathematical studies and occupations, whereas male students are underrepresented in education and occupations in the healthcare-sector (Onderwijsraad, 2019). Moreover, 60% of Dutch women who are on the labour market work part-time, this is three times more than Dutch men resulting in a pay gap, unequal distribution of caretaking tasks, and more (see OESO, 2019). Hence, we believe gender is a highly salient status characteristic in Dutch education.

4.3.2 DATA COLLECTION

The research project Girls in Science collects data on children and adolescents from 10 to 20 years old to examine gender socialization in the family and school context. This study reports on school-related data from the first wave (collected in 2019-2020) and includes adolescents of on average 15.44 (*SD* = 0.33) and 17.49 (*SD* = 0.53) years old.

Both adolescents and their parents participated in the study. Eligible families in the Western and Middle region of the Netherlands were selected from municipality records. Inclusion criteria were biological parents living on the same address with a first child of 17 years old and a second child of 15 years old with a maximum age difference of 36 months. Exclusion criteria were severe psychical or intellectual handicaps of parent or child, being born outside the Netherlands and/or not speaking the Dutch language. Eligible families were invited in writing (letter and brochure) to participate in the study. Families were able to respond to the invitation by returning an answering card or responding via the website of the research project. In contrast to comparable studies in the field of pedagogy and social psychology, this study's recruitment method allows for a calculation of the response rate. Of the 3,316 eligible families, 6% were willing to participate (N = 199) whereas 14% indicated not to be willing to participate in the study. The rest of the families (80%) did not respond to the invitation. Families may have been reluctant to participate due to the relatively significant effort that participation required from them.

Firstly, father, mother, first and second child of the families were requested to each complete an online set of questionnaires which took 30-40 minutes. Secondly, the family was visited in a two-hour home visit. The home visit included interviews with both children, questionnaires, computer tasks, and different interaction-tasks which were filmed. At the end of every home visit pictures were taken of report cards of recent grades and of the final test of primary school of both children. The extensive data collection resulted in a comprehensive dataset on adolescents and their family and school context.

The adolescents reported on in this study were visited between September 2019 and February 2020. These adolescents were in, or graduated from, secondary education, making this sample most suitable for this study's analysis as the proposed mechanism requires subject specific teachers, rather than general primary teachers. The vast majority of the adolescents (86%) were in secondary education, the remaining 14% graduated from secondary education, which was no more than one and a half year ago, reported on their experiences in secondary education. 62% of the adolescents reported on in this study were in or graduated from the highest educational level of the Dutch secondary school system, 24% were in or graduated from the intermediate level, and 14% were in or graduated from the lowest level.[2] The average national distribution across the educational levels is 22% in the highest level, 22% in the intermediate level, 50% in the lowest level and 6% other (Dutch Report Education Inspectorate, 2019). Our sample is thus relatively highly educated, as are the mothers (average age of 48.64; *SD* = 3.38, at the moment of the home visit) and fathers (50.93; *SD* = 4.68) of the adolescents in our sample. 76% of the mothers and 72% of the fathers have obtained a bachelor's degree at the University of Applied Sciences or higher compared to 33% in the Dutch population (Education in Numbers 2020). In sum, the adolescents in our sample are members of families of a relatively high socio-economic status.

Only the adolescents with scores on all relevant variables are included in the study resulting in N = 329 (42% male). Although these adolescents are nested in families, intra-

class correlation coefficients show that the educational data of siblings are independent (ICC < 0.05 for both subjects). The independence indicates that grades in Dutch language and math of adolescents within one family differ as much as grades from adolescents belonging to different families. As the design effect is smaller than 2 (deff = 1.03), the analyses can be performed using general linear modeling (Peugh, 2010).

4.3.3 MEASURES

This study makes use of measures included in the online questionnaire as well as measures included in the structured interview and computer task conducted during the home visits.

Gender Congruence

The independent variable indicating student and teacher having the same gender, is a binary variable with 1 = gender congruence and 0 = no gender congruence. Students indicated their gender in the online questionnaire.[3] The gender of their teachers in math and Dutch language were asked during the structured interview in the home visit (e.g., 'What is the gender of your math teacher?'). The students expressed their teachers' gender as either male or female. Only a few students had both a male and female teacher for a subject in one year or, when graduated, forgot (the gender of) their teacher; these cases are regarded as missing data (N = 10).

The distribution of male and female teachers across math and Dutch language is skewed, but similar for female and male students. 75% of the students in our sample has a female Dutch language teacher, in respect to 25% with a male Dutch language teacher. In math, 38% of the students has a female math teacher, whereas 62% has a male math teacher. The distribution of female and male teachers over the subjects in our sample is comparable with the distribution in the Netherlands. Dutch language teachers in the Netherlands are disproportionately female (72%). For math, the distribution is slightly more balanced with 58% male and 42% female teachers (numbers based on open educational data 2018: DUO, 2020). The skewed distribution is not expected to introduce bias, because teachers do not select their students (or vice versa).

Student Performance: Grade in Math and Grade in Dutch Language

Grade for math and grade for Dutch language are the calculated averages ranging from 1.0 to 10.0 obtained from photographed final exams (14%), secondary school reports (58%), or screenshots from online grade administration systems (28%). For some students, grades are missing because reports or screenshots were not available. These cases are regarded as missing data (N = 17). Additionally, reports with international grades (scale 1 to 7) (N = 6), letters instead of numbers (N = 6), or assessments in text (N = 2) were regarded as missing data. Also, reports that stem from one and a half year before the home visit (N = 4), wrongly photographed reports (N = 1), reports without grades for math (N = 9) or Dutch language (N = 4), or both (N = 2) were regarded as missing data.[4]

Self-Perceived Abilities

The first mediator in the analyses is students' beliefs and judgements about their capabilities in math and Dutch language. The self-perceived abilities are measured using three subject-specific items completed on a 5-point Likert-scale (*how good do you think you are at* [...]) with 1 = very bad until 5 = very good. The full scale consists of twelve items representing beliefs and judgements about students' capabilities in four subjects (Math, Dutch language, humanities and science). The scale is an translation of the math, first language and science items used in earlier research on self-perceived abilities of adolescents (Chamorro-Premuzic et al., 2010; Spinath et al., 2006), and is complemented with items for humanities.

The subscale for math includes the items 1) multiplying and dividing, 2) calculating with capacity measures, and 3) estimating the outcome of a sum with large numbers. The subscale for Dutch language includes 1) text comprehension, 2) grammar, and 3) writing a text. The variables 'self-perceived abilities in math' and 'self-perceived abilities in Dutch language' are constructed by calculating the average of the items for each respondent. Cronbach's Alpha is 0.83 for math and 0.64 for Dutch language.

Gender-Stereotypical Beliefs

The second mediator is assessed by a computerized version of the arts-science Implicit Association Test (IAT) (Nosek et al., 2002a).[5] The IAT was built with E-prime 3.0 based on the task on the Harvard project Implicit demonstration website and the Nosek et al. (2002a) paper. The IAT was conducted on laptop computers and reaction time and accuracy were automatically recorded. Each adolescent's level of implicit stereotypes was determined by calculating differences in reaction time and accuracy in congruent and incongruent blocks resulting in a d-score (using scoring algorithm by Greenwald et al., 2003). A positive d-score represents stronger implicit gender-stereotypical beliefs (e.g., math is for males, Dutch language for females). Negative d-scores represent contra-stereotypical beliefs (e.g., math is for females, Dutch language for males). To reduce possible order effects of the presentation of congruent and incongruent blocks, two versions of the IAT were used, one in which the congruent block was administrated first and one in which the incongruent block was administrated first (Nosek et al., 2005). Adolescents starting with the congruent blocks had significantly higher differences scores (p=0.000). We do not expect that this will introduce any bias as adolescents were randomly assigned to the version starting with congruent blocks or the version starting with incongruent blocks. The randomization ensured that the different gender/teacher combinations in math and Dutch language included about the same number of students who started with a congruent version as students who started with an incongruent version.

4.3.4 ANALYTICAL STRATEGY

Previous research on gender stereotypes in education has mostly focused on math, assessing whether female students suffer from the stereotype 'men are better at math than women' (e.g., Steele, 2003). Some studies include language education to examine the role of the stereotype assuming that 'women are better with language than men' (Plante et al., 2013; Wach et al., 2015). In this study, we will include both subjects (math and Dutch language), estimating the effect of gender congruence through gender-stereotypical beliefs and self-perceived abilities for male and female students in math and Dutch language. As the effect of gender congruence on student performance through gender-stereotypical beliefs and self-perceived abilities is expected to vary for male and female students in math and Dutch language, the hypotheses testing will be split in four separate analyses for male and female students in math and Dutch language. In math, it is expected that female math teachers *reduce* female students' gender-stereotypical beliefs, strengthening their self-perceived abilities and increasing their performance, whereas for male students, male math teachers are expected to *reinforce* gender-stereotypical beliefs strengthening the students' self-perceived abilities, and increasing their performance. In Dutch language, it is expected that female Dutch language teachers *reinforce* female students' gender-stereotypical beliefs, strengthening their self-perceived abilities and increasing their performance, whereas for male students, male Dutch language teachers are expected to *reduce* gender-stereotypical beliefs strengthening their self-perceived abilities, increasing their performance.

Table 4.1. Descriptive Statistics

		Mean (SD)			Correlations Math (p-value)				Correlations Dutch language (p-value)		
		Full Sample N = 329	Male Students N = 138	Female Students N = 191	GC	Grade	SPA	GSB	GC	Grade	SPA
Math	GC	0.47	0.61	0.37							
		(0.500)	(0.490)	(0.485)							
	Grade	6.828	6.690	6.927	-0.085						
		(1.176)	(1.191)	(1.158)	(0.124)						
	SPA	3.565	3.911	3.316	0.055	0.167**					
		(0.937)	(0.807)	(0.946)	(0.324)	(0.002)					
	GSB	0.317	0.297	0.332	-0.066	0.058	0.017				
		(0.402)	(0.408)	(0.398)	(0.234)	(0.297)	(0.754)				
Dutch	GC	0.51	0.21	0.73	-0.063	0.025	-0.143**	0.052			
		(0.501)	(0.409)	(0.446)	(0.257)	(0.652)	(0.009)	(0.346)			
	Grade	6.607	6.339	6.801	-0.032	0.421**	-0.059	0.055	0.079		
		(0.854)	(0.826)	(0.824)	(0.563)	(0.000)	(0.284)	(0.324)	(0.151)		
	SPA	3.431	3.302	3.524	-0.008	0.029	-0.041	-0.034	0.034	0.368**	
		(0.775)	(0.783)	(0.758)	(0.878)	(0.598)	(0.464)	(0.537)	(0.539)	(0.000)	
	GSB	0.317	0.297	0.332	-0.066	0.058	0.017	1.000**	0.052	0.055	-0.034
		(0.402)	(0.408)	(0.398)	(0.234)	(0.297)	(0.754)	(0.000)	(0.346)	(0.324)	(0.537)

Note: GC = Gender Congruence, SPA = Self-Perceived Abilities, GSB = Gender Stereotypical Beliefs ** p<0.01, * p<0.05, + p<0.1

4.4 RESULTS

Table 4.1 presents the means and standard deviations for the full sample and split by adolescent gender. In addition, correlations for math and Dutch language for the full sample are presented. Significant correlations were only found for self-perceived abilities and grades in both math and Dutch language.

The main results for the first hypothesis, that, student-teacher gender congruence has a positive effect on student performance, are presented in Table 4.2. ANOVA analyses were conducted to assess variation in math and Dutch language grades for the different combinations of the student-teacher gender dyad.

Table 4.2. Differences in Grades for Student-Teacher Gender Combinations (N = 329)

	Mean Math Grade (SD)	Mean Dutch language Grade (SD)
Male student x Male teacher	6.639 (1.261)	6.025 (0.806)[a]
Male student x Female teacher	6.768 (1.081)	6.423 (0.815)[b]
Female student x Male teacher	6.991 (1.154)	6.779 (0.801)[c]
Female student x Female teacher	6.820 (1.166)	6.809 (0.835)[c]
Significance Between Groups	$p = 0.203$	$p < 0.001$

Note: means with different superscripts are significantly different from each other, $p < 0.05$

For math, Table 4.2 reveals no statistically significant differences in grades for the various student-teacher combinations. Contrary to what was expected, student math performance is not related to having a teacher of the same gender.

A different pattern applies for Dutch language. Table 4.2 shows statistically significant variation in Dutch language grades for student-teacher gender combinations. Post-hoc analyses were conducted to test the hypothesized association. The analyses reveal that male students with a male Dutch language teacher achieve 0.399 ($p = 0.021$) lower average grades in Dutch language compared to male students with a female teacher. Gender congruence thus has a negative effect for male students' performance in Dutch language. For female students, no such effect was found. Female students' Dutch language performance is not associated with having a female Dutch language teacher. Additionally, the analyses show that female students achieve higher grades in Dutch language than male students regardless of the gender of their teachers. In all, hypothesis 1 is not supported by the data.

Despite the absence of a direct effect, we test the hypothesized indirect effects as Hayes (2018) argues that a correlation between X and Y as precondition for a mediation effect is not necessarily needed. To test the first step in the proposed mechanism explaining the association between student-teacher gender congruence and student performance,

self-perceived abilities is added to the model. The hypothesized mediation effect of self-perceived abilities is assessed using the single mediation model (model 4) in PROCESS. PROCESS is an OLS regression-based macro for SPSS which incorporates bootstrapping techniques for estimating indirect effects (Hayes, 2018).[6] We report bootstrap estimates of indirect effects derived from 10,000 bootstrap samples along with 95% CIs, as 'there is relatively little added value to increasing bootstrap samples above 10,000' (Hayes 2018: 103). We run the model for math and Dutch language for males and females separately, resulting in four analyses of which the findings are presented in Table 4.3. The first and third models present the results of regressing student-teacher gender congruence on students' self-perceived abilities. The second and fourth models present the results of regressing self-perceived abilities on students' grades controlled for student-teacher gender congruence. The indirect effect of student-teacher gender congruence on students' grades through self-perceived abilities are also included.

Table 4.3. The Effect of Gender Congruence on Grades through Self-Perceived Abilities

Male Students				
	Math		**Dutch language**	
	(1) SPA	**(2) Grade**	**(3) SPA**	**(4) Grade**
Constant	3.790 (0.109) p = 0.000	6.170 (0.508) p = 0.000	3.309 (0.075) p = 0.000	5.275 (0.287) p = 0.000
Gender Congruence	0.198 (0.140) p = 0.160	-0.160 (0.209) p = 0.446	-0.033 (0.164) p = 0.841	-0.387* (0.161) p = 0.017
Self-Perceived Abilities		0.158 (0.127) p = 0.216		0.347** (0.084) p = 0.000
Indirect effect		0.031 (0.038) -0.019 0.127		-0.012 (0.059) -0.134 0.105
Observations	138	138	138	138
R^2	0.014 p = 0.160	0.014 p = 0.385	0.000 p = 0.841	0.147** p = 0.000
Female Students				
	Math		**Dutch language**	
	(1) SPA	**(2) Grade**	**(3) SPA**	**(4) Grade**
Constant	3.394 (0.086) p = 0.000	5.843 (0.309) p = 0.000	3.603 (0.105) p = 0.000	5.370 (0.288) p = 0.000
Gender Congruence	-0.211 (0.141) p = 0.136	-0.100 (0.168) p = 0.554	-0.109 (0.123) p = 0.380	0.072 (0.126) p = 0.567
Self-Perceived Abilities		0.338** (0.086) p = 0.000		0.391** (0.074) p = 0.000
Indirect effect		-0.072 (0.050) -0.182 0.017		-0.043 (0.050) -0.148 0.046
Observations	191	191	191	191
R^2	0.012 p = 0.136	0.081** p = 0.000	0.004 p = 0.380	0.129** p = 0.000

Note: Standard errors in parentheses. SPA = self-perceived abilities. Size of bootstrap sample for calculations of the indirect effect = 10,000, 95% confidence interval. ** p<0.01, * p<0.05, + p<0.1

Table 4.4. Double Mediation of Gender Stereotypical Beliefs and Self-Perceived Abilities in the Relation Between Gender Congruence and Students' Grades

Male Students	Math			Dutch language		
	(1) GSB	(2) SPA	(3) Grade	(4) GSB	(5) SPA	(6) Grade
Constant	0.304 (0.056)	3.715 (0.120)	6.148 (0.506)	0.285 (0.039)	3.392 (0.088)	5.138 (0.299)
	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Gender Congruence	-0.011 (0.071)	0.201 (0.140)	-0.151 (0.208)	0.060 (0.085)	-0.016 (0.163)	-0.402* (0.160)
	p = 0.882	p = 0.153	p = 0.470	p = 0.487	p = 0.924	p = 0.013
Gender Stereotypical Beliefs		0.249 (0.168)	0.379 (0.250)		-0.292+ (0.164)	0.251 (0.162)
		p = 0.140	p = 0.132		p = 0.077	p = 0.125
Self-Perceived Abilities			0.133 (0.128)			0.367** (0.084)
			p = 0.298			p = 0.000
Observations	138	138	138	138	138	138
R^2	0.000	0.030	0.031	0.004	0.023	0.162**
	p = 0.882	p = 0.125	p = 0.242	p = 0.487	p = 0.204	p = 0.000
Female Students	**Math**			**Dutch language**		
	(1) GSB	(2) SPA	(3) Grade	(4) GSB	(5) SPA	(6) Grade
Constant	0.360 (0.036)	3.425 (0.106)	5.855 (0.322)	0.320 (0.055)	3.575 (0.114)	5.361 (0.290)
	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000	p = 0.000
Gender Congruence	-0.075 (0.060)	-0.218 (0.142)	-0.102 (0.169)	0.017 (0.065)	-0.110 (0.124)	0.072 (0.126)
	p = 0.211	p = 0.127	p = 0.548	p = 0.795	p = 0.374	p = 0.571
Gender Stereotypical Beliefs		-0.085 (0.173)	-0.030 (0.205)		0.088 (0.139)	0.036 (0.141)
		p = 0.626	p = 0.886		p = 0.528	p = 0.797
Self-Perceived Abilities			0.338** (0.086)			0.390** (0.074)
			p = 0.000			p = 0.000
Observations	191	191	191	191	191	191
R^2	0.008	0.013	0.081**	0.000	0.006	0.130**
	p = 0.211	p = 0.293	p = 0.001	p = 0.795	p = 0.558	p = 0.000

Note: GSB = Gender Stereotypical Beliefs, SPA = Self-Perceived Abilities ** p<0.01, * p<0.05, + p<0.1

For math, the findings are presented in the first two models. Model 1 reveals that student-teacher gender congruence does not have an effect on self-perceived abilities in math for either male or female students. Similar to the findings on student performance but contrary to what was expected, student-teacher gender congruence is not associated with students' self-perceived abilities in math. Model 2 presents the findings of regressing both student-teacher gender congruence and self-perceived abilities in math on math grades. For male students, no significant effects were found for either gender congruence or self-perceived abilities. Hence, both the gender of their teachers and self-perceived abilities are not associated with male students' math grades. For female students, model 2 shows a significant effect of self-perceived abilities on math grades suggesting that stronger self-perceived abilities in math are associated with higher math grades.

In accordance with the insignificant effects of student-teacher gender congruence and self-perceived abilities on math grades, the confidence intervals of the indirect effects

include zero for both males (BootLLCI -0.019, BootULCI 0.127) and females (BootLLCI -0.182, BootULCI 0.017). This confirms that self-perceived abilities do not mediate the association between gender congruence and grades in math for either male or female students.

For Dutch language, the findings are included in models 3 and 4. Model 3 shows no effect for gender congruence on either male or female students' self-perceived abilities in Dutch language. In contrast to what was expected, having a teacher of the same gender is not related to students' self-perceived abilities in Dutch language. Model 4 presents the findings of regressing both gender congruence and self-perceived abilities on Dutch language grades. For male students, significant effects of both variables are found. Contrary to the expectation, controlling for self-perceived abilities, student-teacher gender congruence predicts lower Dutch language grades for male students. Self-perceived abilities are positively associated with male students' Dutch language grades, controlling for gender congruence. For female students, model 4 only reveals a positive significant effect of self-perceived abilities indicating that, controlled for student-teacher gender congruence, stronger self-perceived abilities predict higher grades in Dutch language. An effect for gender congruence on female students' Dutch language grades was not found.

For both male and female students, the bootstrap confidence interval of the indirect effect of gender congruence on student performance through self-perceived abilities in Dutch language includes zero (male students BootLLCI -0.134, BootULCI 0.105, female students BootLLCI -0.148, BootULCI 0.046), indicating that self-perceived abilities do not mediate the association between gender congruence and performance in Dutch language for either male or female students. Therefore, hypothesis 2 should be rejected.

To examine the role of gender-stereotypical beliefs, in a last step gender-stereotypical beliefs is added to the model. The serial multiple mediator model is assessed with 'Model 6' in PROCESS (Hayes, 2018). Again, bootstrap estimates of indirect effects derived from 10,000 bootstrap samples along with 95% CIs are reported. Table 4.4 presents the findings of regressing student-teacher gender congruence on gender-stereotypical beliefs (model 1 and 4), regressing both gender congruence and stereotypical beliefs on self-perceived abilities (model 2 and 5), and regressing gender-stereotypical beliefs and self-perceived abilities on students' grades controlled for student-teacher gender congruence (model 3 and 6). The mediation effects of gender-stereotypical beliefs and self-perceived abilities, as well as the double mediation effect and the total effect of the model are presented in Table 4.5.

Table 4.4 shows a similar pattern as Table 4.3, the addition of gender-stereotypical beliefs to the model does not change any of the previously found relations. Model 1 and 4 show that, contrary to what was expected, gender congruence is not associated with students' gender-stereotypical beliefs. Having a teacher of the same gender is not related to students' beliefs about what is capable for the sexes in either math or Dutch language. Model 2 shows that gender-stereotypical beliefs are not related to students' self-perceived abilities in math. Model 5 shows a negative effect, on the 0.10 level, of gender

stereotypical beliefs on male students' self-perceived abilities in Dutch language. Stronger gender stereotypical beliefs of male students (beliefs about what is capable for the sexes in math and Dutch language) seem to be negatively associated with their beliefs about their own capabilities in the subject. For female students, no such effect was found. Lastly, model 3 and 6 reveal that gender-stereotypical beliefs are not associated with students' performance in math or Dutch language, for either male or female students.

Table 4.5 summarizes the indirect effects of student-teacher gender congruence on students' grades through gender-stereotypical beliefs, self-perceived abilities and both gender-stereotypical beliefs and self-perceived abilities. Table 4.5 reveals no indirect effect of student-teacher gender congruence on students' grades through students' gender-stereotypical beliefs for either male or female students in either math or Dutch language. Moreover, Table 4.5 shows no evidence for a serial mediation effect of student-teacher gender congruence on students' grades through gender-stereotypical beliefs and self-perceived abilities. Hence, hypothesis 3 should be rejected.

The results reported above are the findings based on the basic models. To control for effects of student age and educational level of the students, all models have also been tested with these two variables as covariates. The addition of the covariates reveals a significant negative effect, on the 0.01 level, of male students' gender stereotypical beliefs on their self-perceived abilities in Dutch language (basic model: p = 0.077, model with covariates: p = 0.009). All other effects and patterns found are similar to the findings of the basic model. No other conclusions are to be drawn regarding the hypotheses.

Table 4.5. Indirect Effects of Gender Congruence on Students' Grades

Male students						
	Math			Dutch language		
	Effect Size (SE)	LLCI	ULCI	Effect Size (SE)	LLCI	ULCI
Total Mediated Effect	0.022 (0.051)	-0.071	0.136	0.003 (0.065)	-0.127	0.132
Mediation through GSB	-0.004 (0.033)	-0.079	0.062	0.015 (0.025)	-0.031	0.072
Mediation through SPA	0.027 (0.036)	-0.020	0.120	-0.006 (0.062)	-0.132	0.116
Double mediation	-0.000 (0.004)	-0.010	0.008	-0.006 (0.011)	-0.031	0.016
Female students						
	Math			Dutch language		
	Effect Size (SE)	LLCI	ULCI	Effect Size (SE)	LLCI	ULCI
Total Mediated Effect	-0.069 (0.054)	-0.184	0.030	-0.042 (0.051)	-0.151	0.049
Mediation through GSB	0.002 (0.021)	-0.039	0.053	0.001 (0.009)	-0.020	0.020
Mediation through SPA	-0.074 (0.050)	-0.185	0.014	-0.043 (0.050)	-0.150	0.048
Double mediation	0.002 (0.006)	-0.007	0.018	0.001 (0.004)	-0.008	0.009

Note: Direct effect refers to X>Y controlled for M1 and M2, Total Mediated Effect is the sum of the indirect effects of X>Y. Size of bootstrap sample for calculations of the indirect effect = 10,000. Standard deviations of the 10,000 bootstrap estimates of the indirect effect between parentheses (SE). LLCI = bootstrapped lower level confidence interval; ULCI = bootstrapped upper level confidence interval, 95% confidence interval.

4.5 DISCUSSION AND CONCLUSION

Despite the fact that empirical research on representative bureaucracy and performance has expanded extensively in recent years, to date individual-level mechanisms explaining the effects of representation on performance in public organizations have hardly been examined. In particular, individual-level mechanisms explaining representation effects from the perspective of the citizen-client are subject to speculation. This study contributes to representative bureaucracy theory by examining a psychological process that may underlie symbolic representation effects. We empirically tested our model in the school context and hypothesized an effect of student-teacher gender congruence on students' performance via students' gender-stereotypical beliefs and self-perceived abilities in the subjects math and Dutch language. The model was tested on multi-source data on high school students in the Netherlands collected in the first wave of a longitudinal study among families in the Netherlands.

We separated our analyses for math and Dutch language and for female and male students. By doing so, our proposed model was tested four times; for male and female students in math and male and female students in Dutch language. Our study provides no evidence for the hypothesized positive representation effects and the theorized mechanism underlying symbolic representation in any of the four cases.

For math, we did not find that gender representation was associated with either male or female students' math performance, nor did we find evidence for a role of students' self-perceived abilities or gender-stereotypical beliefs in the hypothesized associations. We did find that self-perceived abilities can predict math scores for female students, but not for male students, suggesting different dynamics of self-perceived abilities in math for male and female students.

Also for female students in Dutch language, our results showed no evidence for a relation between female Dutch language teachers and female students' Dutch language performance, self-perceived abilities, or gender-stereotypical beliefs. The results did show that stronger self-perceived abilities of female students in Dutch language are related to higher grades in the subject.

For male students in Dutch language, our study showed surprising findings. In contrast to the positive gender-representation effects we expected, we found a *negative* gender-representation effect for male students in Dutch language. The results showed that male students with a male Dutch language teacher had lower grades than male students with a female Dutch language teacher. The proposed psychological processes described in our model cannot explain this negative effect: neither male students' ideas about males' and females' capabilities in Dutch language, nor male students' judgements about their own capabilities in the subject were associated with having a male Dutch language teacher. The results did show that stronger gender stereotypical beliefs can predict lower self-perceived abilities and that stronger self-perceived abilities of male students in Dutch language can

predict higher grades in the subject.

The null-findings and negative gender-representation effect in Dutch language for male students contradict the evidence for effects of gender representation on student performance from previous studies (Agyapong, 2018; Dhillon & Meier, 2020; Keiser et al., 2002; Song, 2018; Zhang, 2019), and positive representation effects in general (Riccucci & Van Ryzin, 2017). While our study is not the first to find negative representation effects (Doornkamp et al., 2021; Keiser et al., 2002), previously found negative effects have been assumed to be 'spillover effects', suggesting that, to a certain degree, the majority may benefit from minority representation effects as well (Riccucci et al., 2018). However, as the positive representation effects for female students (the minority gender in our study) are absent in our study, the negative gender-representation effects for male students seem to stand on their own. This, and the null-findings with respect to most hypotheses, lead us to explore other explanations.

Let us first more closely consider the national context in which this study has been conducted. In representative bureaucracy theory, it is argued that for gender-representation effects to occur gender has to be salient to the policy at hand (it has to be a men's issue or a women's issue) (Keiser et al., 2002). The psychological processes theorized in our model also depend on this notion (the role of gender-stereotypical beliefs and self-perceived abilities). Gender is assumed to be salient to education as a consequence of enduring disadvantages of women compared to men, as becomes apparent from gender differences in achieved educational level (Evans et al., 2020). Historically, women have been discriminated against based on gender stereotypes in which they are assumed to have lower academic capabilities (e.g., Hausmann, 2014; Prentice & Carranza, 2002; Szymanowicz & Furnham, 2011). Even now, educational access for girls and women is not self-evident everywhere in the world. Probably, it is no coincidence that in recent studies effects of gender representation in education have been found in countries in which gender inequality in educational attainment still exists: Ghana, India, South Korea, and China rank in the third quartile of the Global Gender Gap Index Rating; subindex Gender Gap in Educational Attainment (Agyapong, 2018; Dhillon & Meier, 2020; Song, 2018; Zhang, 2019; see Global Gender Gap Report, 2020). However, educational attainment in the Netherlands, the country in which this study has been conducted, is equal for men and women (see Global Gender Gap Report, 2020). Although the educational careers of Dutch male and female students display gender-stereotypical patterns (see Onderwijsraad, 2019), the role of gender (and gender stereotypes) might be less salient and explicit to Dutch boys and girls in high schools than assumed, causing the drivers of gender-representation effects to diminish (or even disappear) (cf: Fay et al., 2020; Meier, 2019).

More importantly, the sample in our study consists of adolescents mainly from Dutch families with a higher SES compared to the nation's average. Stereotypical gender roles and behaviors are more pronounced in segments of the population with lower SES compared to segments with a higher SES (e.g., Trusty et al., 2000). As a result, the role of gender might

be less salient and explicit for the students in our sample not only because of relatively high gender equality in the Dutch national context, but also because of gender roles being less pronounced in high-SES families, decreasing the likelihood of gender-representation effects.

So far, in representative bureaucracy theory salience of minority status to the policy at hand has been discussed in the context of active representation; minority bureaucrats need to experience salience of their minority status to the policy at hand for their willingness to actively represent. Our study points to the role of salience in the context of symbolic representation; for symbolic representation effects to occur the minority citizen-client has to experience their minority status to be salient in the context of the bureaucratic interaction. This might inspire scholars to apply social cognitive theory to the study of symbolic representation in a variety of bureaucratic interactions in order to explore the role of modeling effects in contexts that may vary as to the saliency of social norms related to a minority status.

Moreover, the national and socioeconomic context in which interaction between the bureaucrat and citizen-client takes place, seem to play an important part in the degree to which salience is experienced and representation effects are likely to occur. Future research on bureaucratic representation might want to study the role of national and socioeconomic context in experienced salience of minority status to the policy at hand as a condition for symbolic representation to occur.

The negative gender-representation effect for male students in Dutch language found in our study could not be explained by the proposed psychological mechanism including gender-stereotypical beliefs and self-perceived abilities. The negative representation effect might have been caused by other mechanisms from the perspective of the student (symbolic representation), or by mechanisms from the perspective of the teacher (active representation). From the perspective of the student, teachers similar to themselves might strengthen pressures to succeed. Stress and pressure to succeed can create a psychological state in which negative thought is promoted and optimal academic performance is impeded (Autin & Croizet, 2012).

From the perspective of the teacher, teachers' grading might be influenced by the gender of a student. Research on gender grading bias shows that male students are systematically graded more strictly by teachers than female students, resulting in a negative grading bias against boys (Bonnesronning, 2008; Cornwell et al., 2013; Falch & Naper, 2013; Lindhal, 2007; Lavy, 2008; Matějů & Smith, 2014; Protivínský & Münich, 2018). For Dutch language, this grading bias in combination with the stereotype that male students would perform worse in Dutch language than female students could (unconsciously) influence teacher grading practices. In particular, male teachers might be stricter in grading male students in Dutch language as family studies indicate that men hold stronger explicit stereotypical attitudes about gender than women (Blakemore & Hill, 2008; Endendijk et al., 2013; Tenenbaum & Leaper, 2002). Also, social role theory proposes that fathers are more likely to socialize

children, especially boys, in accordance with prevailing gender stereotypes in society (Eagly et al., 2000), which may also apply to male teachers. And although the teacher may show contra-stereotypical behavior (male Dutch language teacher), this teacher may still believe the gender stereotype connected to their subject is true (Xu & Meier, 2021). To receive a better understanding of (negative) representation effects, future research might want to persist to explore the (unconscious) psychological processes at the individual level underlying both symbolic and active representation.

The findings of this study should be interpreted in light of its limitations. In addition to the study's national context and the relatively high SES of the families included in the sample, the composition of the sample as to the gender of teachers across subjects led to statistical limitations. The skewed distribution of teachers in Dutch language and math resulted in small subsamples of male students with male Dutch language teachers and female students with female math teachers, whereas, theoretically, they were key to our study. Although bootstrapping techniques account for irregularities of sampling distributions and increase power for the calculation of indirect effects, this may have weakened the statistical power and, hence, the likelihood of Type 2 errors. This may have driven the null findings, too.

Furthermore, the educational context in which our study is conducted makes a strong case for active representation because of teachers' discretion and salience of minority status in education (Keiser et al., 2002). Therefore, in such a context, both symbolic and active representation may occur simultaneously and should be seen as complementary processes. Hence, it is difficult, if not impossible, to disentangle effects of symbolic and active representation in the empirical analyses. That said, social cognitive theory provided us with strong theoretical argumentation for the mechanism through which symbolic representation occurs. More importantly, as student-teacher gender congruence is an exogenous variable and teachers and students do not have any chance to influence the student-teacher-match, no systematic biases were expected.

Despite these limitations, our study provides some important implications for the study of symbolic effects of gender representation in education specifically. First, when studying psychological mechanisms explaining symbolic effects of gender representation, scholars should consider differences in mechanisms for male and female students across subjects. Our results show divergent relations between the same variables for male and female students in math and Dutch language.

Second, contrary to our expectations, our results show no evidence for effects of gender-stereotypical beliefs. We used a measure of implicit gender stereotypical beliefs, while previous studies have relied on survey measures of gender stereotypical beliefs (Doornkamp et al., 2019; Zhang, 2019). It may be relevant to explore the role implicit and espoused gender stereotypes may play in symbolic effects of gender representation. This could contribute to further theorizing about the importance of unconscious bias compared to opinionated beliefs in such psychological processes.

Finally, this study's null effects have drawn attention to the characteristics of our

sample, emphasizing the need for comparative research on representation effects across countries. Additionally, scholars in representative bureaucracy should consider and further theorize on the role of SES and its intersection with gender (and race), both in education and in other public services. Moreover, this study calls for further exploration of (unconscious) psychological mechanisms grounded in the human interactions producing representation effects. The reverse effects and suggested conditions for symbolic effects of gender representation to occur point at avenues for further exploration of psychological mechanisms and their conditions in the effects of symbolic representation.

4.6 NOTES

1. The longitudinal study has been discussed with a data protection officer and has been approved by Leiden University ethics committee.

2. In the Netherlands, secondary education exists of three educational levels: VMBO (lowest level; prepares for vocational education, four years; age 12 till 16), HAVO (intermediate level; prepares for applied sciences, five years; age 12 till 17) and VWO (highest level; prepares for university, six years; age 12 till 18).

3. As representative bureaucracy theory uses the term '*gender* congruence', for consistency purposes we also use the term 'gender' to indicate 'sex'. The questionnaire included a non-binary option for respondents to indicate their sex. This option was not used. Consequently, we treated the gender congruence concept as a binary concept although we are aware gender is not binary.

4. To assess whether the missing data is random, ANOVA analyses were conducted comparing students with and without missing grades for the main variables in our model. The analyses show that the groups with and without missing grades do not differ significantly.

5. The IAT measures the association of female and male attributes with the school related concepts 'arts' and 'science'. During the IAT, participants are requested to sort words into groups by pressing keys. In congruent blocks, female names (e.g., 'Julia') and feminine words (e.g., 'Dutch language') need to be sorted in the 'language and culture' group. In incongruent blocks female names and masculine words (e.g., 'Math') need to be sorted in the 'nature and technology' group. Reaction time and accuracy in sorting these names and words are administrated and used to calculate participants' implicit gender stereotypes associated in the school context.

6. Bootstrapping is a resampling process in which an empirically derived representation of the sampling distribution is generated. The indirect effect for each taken random sample from the original sample is calculated constructing the confidence intervals. As a result, irregularity of sampling distribution is better respected and inferences are more likely to be accurate. Especially in small samples, the bootstrap technique is

useful as the bootstrapping increases statistical power (Hayes, 2018).

4.7 ACKNOWLEDGEMENTS

The authors would like to thank Kenneth J. Meier and Dimiter Toshkov for their constructive and insightful feedback on previous versions of this article.