Stereotype threat in primary and secondary education: the influence of teachers

Demi Siënna Klumper – S3459292

Msc Educational Sciences (Learning in Interaction)

Faculty of behavioral and social sciences

University of Groningen

Index

Abstract

As a society we should strive to educate our children in the best and fairest way possible. Yet, several studies have shown that stereotype threat effects could be a hazard to the academic achievements of pupils from stigmatized groups. Teachers could possibly counteract these stereotype threat effects. The present review study focused on the positive influence teachers have on the academic achievements of children that are susceptible to stereotype threat in primary and secondary education, using three research questions regarding teacher-pupil relationships, teacher behavior and teacher expectations. To synthesize findings from prior studies on stereotype threat effects conducted in primary and secondary education, a systematic review study was conducted. Articles were selected from ERIC, PSYCHINFO and Smartcat and assessed for eligibility following inclusion and exclusion criteria. A data extraction form was used in order to assemble the information from the primary studies. In total, 13 articles met the inclusion criteria. The articles were narratively synthesized by systematically describing and integrating the results of the studies. Results showed that factors that can positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education are the teacher being warm demanding, offering assistance to all students, the quality of the teacher-pupil relationship is strong, warm demanding behavior, naming the stereotype, high expectations, and wise feedback.

1. Introduction

Society has no shortage of stereotypes, even though most people would rather not admit it. Cambridge Dictionary (2022) describes the definition of a stereotype as "a set idea that people have about what someone or something is like, especially an idea that is wrong". One could imagine that being a part of a group that has one or more stereotypes against them, can be detrimental to their beliefs about themselves. Stereotypes are a part of education as well. Overall, as a society we should strive to educate our children in the best and fairest way possible. Yet, stereotypes can make education a lot harder for children. So how can we ensure that children from stereotyped groups, still have the same chances in education compared to their peers? A first step could be for teachers to be aware of the potential stereotypes they hold and combat stereotype threat effects in their classroom, seeing as teachers are a huge part of the school experience of pupils, which will be explained further below. The present review study looks at stereotype threat in primary and secondary education, especially the literature that is available on the role teachers could fulfil in order to prevent and combat stereotype threat effects. Firstly, stereotype threat and research on stereotype threat will be outlined. Secondly, teachers' influence on students' (academic) performance, motivation and learning will be further explained by looking at teacher-pupil relationships, teacher behavior and expectations. These three aspects will be related by linking them together. Stereotype threat and the (positive) influence the teacher can have on combatting stereotype threat effects will be discussed using existing literature on the topic, which is the main goal of the present study.

1.1 Stereotype threat

Stereotype threat is a social-psychological explanation for the observed difference in (academic) performance between different groups (Spencer et al., 1999; Steele & Aronson, 1995). It concerns the negative effects of stereotypes or an emphasized social identity on individuals that are a part of a stigmatized group (Flore & Wicherts, 2015). This can result in negative consequences for their academic performance and academic self-concept (Bowe et al., 2017), seeing as students are more concerned with possibly confirming the stereotype, which draws on mental and psychological resources that causes the student to perform below their potential (Ollrogge et al., 2021). Crucial in this phenomenon is that individuals of a stigmatized group are aware of the stereotype during the task at hand (Agnoli et al., 2021; Flore et al., 2018). Other important moderators suggest that individuals should be strongly identified with the assessed academic domain that concerns the stereotype (Smith & White, 2001), and find the

test (e.g. a mathematics ability test) to be difficult (Keller, 2007; Spencer et al., 1999). Yet, these moderating effects are found not to be as strong as previously assumed in a recent study that investigated whether stereotype threat effects are moderated by performance (Stoevenbelt et al., 2022).

Stereotype threat effects have been researched for several different stigmatized groups, such as for women (e.g., Spencer et al., 1999), ethnic minorities (e.g., Shih et al., 1999; Steele & Aronson, 1995) and children from a low socio-economic status background (e.g., Desert et al., 2009). Other stigmatized groups, such as members from the LGBTQ-community and pupils with a mental illness have not been researched a lot on stereotype threat effects in the context of education. For several researched stigmatized groups, significant effects have been found in the past. Yet, the field of stereotype threat has been identified to suffer from publication bias (Flore & Wicherts, 2015). This means that published findings are more likely to be statistically significant than unpublished ones (Rosenthal, 1979), and there are replications with mixed results (Agnoli et al., 2021; Flore et al., 2018). This could imply that stereotype threat effects are less of a problem than previously assumed. This is an important notion to keep in mind for the results of the present study, even though it could be relevant to negate stereotype threat effects; one could question the influence they truly hold.

1.2 Teachers' influence and stereotype threat

Teachers are generally assumed to have a huge influence on their students. They matter for the motivation, engagement, learning, behavior and psychological support of students (Pianta, 2016). Nouwen and Clycq (2019) argue that teacher-pupil relationships can strongly influence school engagement and academic performance, as well as identification of students with certain subjects in school. Hamre and Pianta (2005) add that the relationship between teachers and students is especially important for students who are already at risk for underperformance, such as pupils from stigmatized groups. Teachers are key actors for children to integrate and experience school as a safe learning environment, especially regarding the respect and acceptance pupils experience from their teachers (Nouwen & Clycq, 2019). Teacher mindset climate can have a significant impact on students' achievements, motivation and persistence as well. By expressing certain beliefs and behaviors, such as giving special treatments to students who are achieving well in a certain subject, teachers could create a fixed mindset for students. This entails that students believe that personal traits are fixed. A fixed

mindset is suggested to be detrimental to students' academic performance and could be linked to students' experiences of stereotype threat (Seo & Lee, 2020). Furthermore, teachers' explicit beliefs and expectations influence students' self-beliefs, academics beliefs, attitudes towards school, the learning process and achievements (DeCuir-Gunby & Bindra, 2022). When teachers express high expectations and belief in students' capabilities explicitly, this correlates with successful academic achievement (Steele, 2010). Teachers' own beliefs and expectations can influence their perceptions, which may result in bias. These biases can be either explicit or implicit. DeCuir-Gunby and Bindra (2022) describe these implicit biases as "the attitudes or stereotypes that affect our understanding, actions, and decisions" (p.1). There are several examples of researches that point out that teachers are biased (Lorenz, 2021; Tiedemann, 2000). According to DeCuir-Gunby and Bindra (2022) these implicit biases create invisible barriers between students and their opportunities and achievements. These biases that some teachers might hold, may relate to stereotype threat, because regarding the stereotype threat phenomenon the decrements in a person's (academic) performance are due to the awareness that the group they are a part of is considered to be less skillful in the assessed domain (Agnoli et al., 2021). In other words, there is some kind of stereotype against them and their performance, in this case in the form of teachers' (implicit) bias, and they are aware of it.

Nouwen and Clycq (2019) point out that teacher-pupil relations are an important factor in explaining stereotype threat effects on student outcomes. They found that pupils in stigmatized groups feel the least respected and accepted by teachers. This is not a good thing, seeing as particularly for them, teacher-pupil relationships play an important role in relation to academic self-concept. Nouwen and Clycq (2019) also found that teacher-pupil relationships play an important role in the occurrence of psychological disengagement, discounting of negative feedback and disidentification with education. Yet, positive student-teacher relationships could help students achieve, engage, regulate their emotions, build social competence and take on academic challenges – if the relationship is of high quality it could mitigate, or at least reduce, stereotype threat (Steele, 2010). Lastly, one of the mediators that is often researched in studies on stereotype threat, is the presence of a positive role model (Flore & Wicherts, 2015). A role model could in theory contradict a stereotype and therefore protect the individual from a stigmatized group from stereotype threat. Research shows moderate evidence that ingroup experts could mitigate the impact of negative stereotypes on stigmatized individuals; ethnic minority teachers have more positive attitudes toward ethnic minority students than ethnic majority teachers, and female STEM professionals and professors (Science,

Technology, Engineering and Mathematics), promote positive attitudes towards STEM, identification with STEM, stronger self-efficacy beliefs and more effort on STEM tests for girls (Ollrogge et al., 2021).

1.3 The present study

Researching stereotype threat in education is of importance because of several reasons. Agnoli et al. (2021) and Flore and Wicherts (2015) state that it is important for determining when stereotype threat effects emerge, to study the effect in a natural setting such as a classroom, and to identify potential moderators of the (in their case gender) stereotype threat effect on student outcomes. Furthermore, Flore and Wicherts (2015) suspect that stereotype threat is the cause of stigmatized groups to leave certain professional domains. Yet, for some (ethnic) minorities, a problematic home environment is often blamed for poorer academic performance, whereas the structural features within the educational system itself are seldom discussed (Nouwen & Clycq, 2019). Looking at these factors, it is important to research if teachers could prohibit stereotype threat effect from happening, or at least reduce the chances of stereotype threat effect to appear.

The present review study focuses on a systematic analysis of prior studies on stereotype threat effects conducted in primary and secondary education, because researchers suggest that awareness of stereotypes could be found as early as age 6. Research indicates that gender identification emerges from the age of 3 (Katz & Kofkin, 1997). It could be argued that even young children are vulnerable to stereotype threat because of their stable gender identity around the age of 6. Research suggests that the vulnerability to stereotype threat starts around the age of 7, because around this age children are more realistic about their future academic performances (Eccles et al., 1989; Nicholls, 1979). Flore and Wicherts (2015) also suggested that there was a bigger effect size for children (boys and girls) below the age of 13 in a research on stereotype threat effects for girls in math, science and spatial skill tests. Alternatively, other research suggests stereotype threat develops later, during early adolescence (Agnoli et al., 2021). It is clear that stereotype threat effects emerge from a young age, but it is not clear if it is already evident in primary education or only in secondary education. Thus, knowing around what age could be relevant, as this could underline the importance for teachers guiding specific age groups to be aware of their influence on stereotype threat for children.

A systematic literature review was conducted in order to make sense of larger bodies of information. By using a systematic literature review, several relevant research articles are used

to understand the topic at hand instead of just one (Petticrew & Roberts, 2008). It is a replicable method, which makes it a more reliable design than for example a narrative review. The present review study aims to create a comprehensive picture of teachers' influence on stereotype threat, stemming from empirical studies that are published in recent years.

The primary focus of the present study is the role of the teacher in stereotype threat effects in primary and secondary education. This topic will be researched through three research questions. The first question is "Which factors in teacher-pupil relationships can be identified in the literature that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education?", the second question is "Which factors in teacher behavior can be identified in the literature that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education?", the third question is "Which factors in teacher expectations can be identified in the literature that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education?". The three questions assume that teacher behavior, teacher expectations and teacher-pupil relationships are somehow all related to stereotype threat effects in the classroom. It is expected that these three aspects are often connected in the literature. Expectations influence behavior, behavior influences relationships, and relationships influence both behavior and expectations. Nouwen and Clycq (2019) already found that the respect and acceptance pupils experience from teachers is crucial for their identification with education. There are also examples of relationships between teacher perceptions of pupils and teacher behavior (Kranz et al., 1970). Furthermore, it is expected that this literature review will point out empirical studies that support the idea that when teachers hold negative implicit biases, this influences their relationships with pupils and their behavior in the classroom and reinforces stereotype threat for their stereotyped students (DeCuir-Gunby & Bindra, 2022; Seo & Lee, 2020). Vice versa, a teacher who is aware of their potential bias is more likely to protect their students from stereotype threat (Steele, 2010).

2. Method

2.1 Design

A systematic review was used to answer the research questions, this is "a review that strives to comprehensively identify, appraise, and synthesize all the relevant studies on a given topic" (Petticrew & Roberts, 2008, p.19). The present study was a descriptive qualitative study. This type of literature review was chosen, because it is a method of making sense of large bodies of information. Furthermore, it draws on all relevant scientifically sound research instead of a single study to interpret effects (Petticrew & Roberts, 2008).

2.2 Search strategy

The present study focused on articles on stereotype threat and teachers' influence on students in primary and secondary education. The present study focused on published articles from the past 15 years (2007-2022), in order to establish a current vision on the subject at hand and to make the search strategy feasible within the available time. The search terms were: 'Stereotype threat AND teacher-pupil relationships', 'Stereotype threat AND teacher behavior', 'Stereotype threat AND teacher expectations', 'Stereotype threat AND teachers' and lastly 'Stereotype threat AND teachers AND achievement gaps'. The language the articles could be in was English. The types of articles that were included are quantitative as well as qualitative studies.

2.3 Procedure

Data was collected from ERIC, PSYCHINFO and SmartCat. These scientific databases were chosen, because they potentially hold information about stereotype threat and teachers' influence and because they allow for reproducible searches. After searching and finding articles, the title and abstract of the articles were scanned by using the inclusion and exclusion criteria presented below. A flow chart was explicated to visualize the selection process and provide transparency, Figure 1 shows the selection process within this flow chart. The articles that remained were scanned for their methods and results section, looking at relevance for the present study. Studies were excluded if the results presented did not relate teacher-pupil relationships and/or teacher behavior and/or teacher expectations to stereotype threat. A data extraction form was used in order to assemble the information, this can be found in Appendix 1. The list for the inclusion and exclusion criteria is as follows:

- Age of respondents is between 4 and 18 years old
- Articles contain empirical data
- (Quasi-)experimental studies contain an experimental group and a control group

- Articles are peer-reviewed
- Articles contain information about stereotype threat, teacher-pupil relationships and/or teacher behavior and/or teacher expectations
- Articles contain information about the influence of teacher-pupil relationships and/or teacher behavior and/or teacher expectations on the academic performance of students in primary and/or secondary education
- Articles are published between 2007 and 2022
- No grey literature

The amount of unique hits was 997 (duplicates withdrawn), whereas the amount of studies that remained in the end was 13. Within the scientific database the publication date was set in order to swiftly be able to omit articles that were published before 2007. The same was done for the age limit of the respondents and articles that did not undergo peer review. After this was done, the articles that remained were scanned for the more content-related inclusion criteria. As was said, 13 articles remained that were useable for the present review study. An overview of the remaining articles can be found in Appendix 2.

Figure 1

Flow Chart for the Visualization of the Selection Process
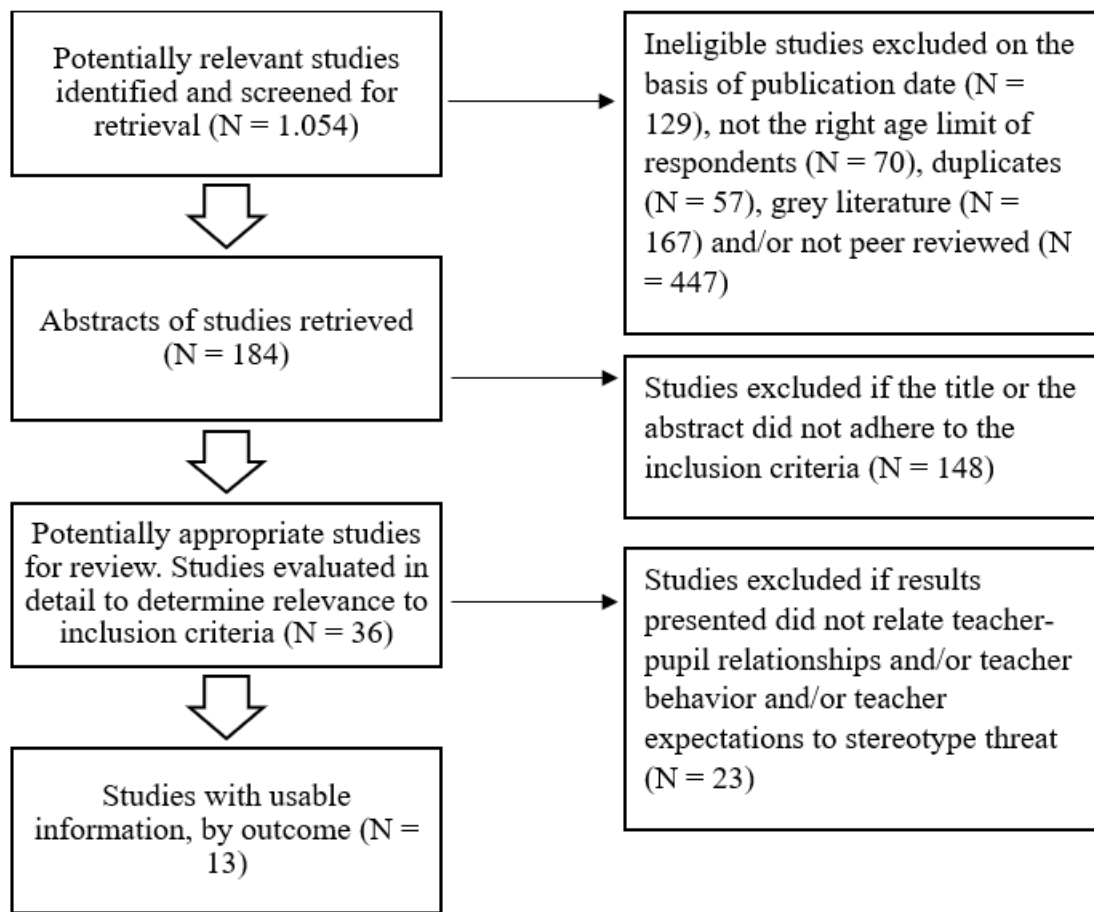
2.4 Data analysis

After the selection process, which was visualized in a flow chart, the remaining articles were analyzed using a data extraction form as shown in Appendix 1. This data extraction form outlines the population, details of the (possible) intervention, outcomes of interest and relevant methodological information. By using a data extraction form, consistency and objectivity was ensured as much as possible (Petticrew & Roberts, 2008). The articles were categorized into different types of research and into different stereotypes (gender stereotype threat, ethnic/racial stereotype threat, both or other). The influence on stereotype threat was registered, which created a dataset. The articles were narratively synthesized by systematically describing and integrating the results of the studies. With this narrative synthesis, a textual approach was adopted that provided an analysis of the relationships within and between studies, as well as an overall assessment of the robustness of the evidence (Centre for Reviews and Dissemination,

2009). Conclusions were drawn from a summary of data from the data extraction form and from the narrative synthesis.

3. Results

In total, a number of 13 articles remained after the selection process. There were four studies found that said something specifically about teacher-pupil relationships, six studies mentioned teacher behavior, and ten studies mentioned teacher expectations. Some of the studies mentioned two or more of the elements studied (teacher-pupil relationships, teacher behavior and/or teacher expectations). Two studies took place in primary education, eight studies took place in secondary education, and three studies took place in both primary and secondary education. A matrix of the categorization of studies in type of education and element(s) studied can be found in Appendix 3. For the following paragraphs, the results are sorted first by the type of research design used in the studies. Secondly, the articles are sorted into the type of stereotype threat they researched, either gender stereotype threat, ethnic/racial stereotype threat, both or other. Lastly, the research questions will be answered by using the overview of the studies. Several elements could be identified that have a positive influence on the academic performance of students that suffer from stereotype threat in primary and secondary education. These elements will be connected to either teacher-pupil relationships, teacher behavior and/or teacher expectations.

3.1 Overview of the studies

3.1.1 Experimental studies

Four of the studies were experimental studies. One of them was an experiment based on gender stereotype threat (Master et al., 2014), two of them were based on ethnic/racial stereotype threat (Ollrogge et al., 2022; Yeager et al., 2014) and the last one was based on both gender and ethnic/racial stereotype threat (Seo & Lee, 2021). The experimental studies make it possible to draw conclusions empirically on the causal relationship of teacher influence and stereotype threat effects, with other designs, this is not possible.

The study by Master et al. (2014) focused on gender stereotype threat. Two experiments were conducted, using a three-group posttest only design. In the first experiment, participants were randomly assigned to one of the three experimental conditions: either the stereotype threat condition, no gender difference, or the baseline control condition. Surveys were used to determine how stereotypes about the computer science abilities of girls affect their concerns about being negatively stereotyped. The second experiment had the same type of experimental conditions as experiment 1 and used surveys to determine how teacher gender and stereotype

threat affect the self-reported concerns of adolescents about being negatively stereotyped in computer science courses. The authors concluded that adolescent girls are in fact sensitive to the threatening cues in STEM courses, but that female teachers evoke less of these concerns as compared to male teachers. The teacher gender difference was only significant for girls in the stereotype threat condition. It is important to note that the surveys that were conducted, were self-reported by the adolescent girls. The sample sizes were relatively large (N = 218 for experiment 1, N = 193 for experiment 2). Furthermore, attention and manipulation checks were conducted and Cronbach's alpha indicated a high reliability of the scales used in the questionnaires.

The studies by Yeager et al. (2014) and Ollrogge et al. (2022) examined ethnic/racial stereotype threats. Yeager et al. (2014) used a two-group posttest only design to examine the effects of a strategy to restore trust on minority adolescents' response to critical feedback. They tested an intervention called 'wise feedback', wherein students received critical feedback accompanied by a wise feedback note to lessen mistrust by informing the students that their teacher had high expectations of them and believed in their ability to reach those expectations. Students were either in the placebo control note condition, or in the wise feedback note condition. Results indicated that those high expectations led students to be more likely to revise their essays, which was a significant effect only for African American students, especially those that had low-trust. The wise feedback intervention resulted in better achievements for this subgroup and severed the relationship between chronic mistrust and performance for them. The experiment conducted could be considered as strong empirical evidence for a causal relationship between the intervention and the response students gave. because of several reasons. First of all, the teachers that reviewed the essays were unaware of the condition assignment and coders were independent. There was one control group in experiment 1 and 2, and there were even two in experiment 3. The assignment to each condition was random. Teachers and students were both unaware of the specific details of the experiment. There were no moderator effects found. Furthermore, school trust was measured several times. Also, the experiment was longitudinal, a follow-up was conducted after 2,5 months. Results of the follow-up study further strengthened earlier findings, with results that indicated that the wise feedback intervention increased year-end trust, especially for low-trust African American students.

Ollrogge et al. (2022) used a pretest-posttest control group design to study whether in students of Turkish descent the learning vocabulary is impaired when their teacher activates the

negative stereotype towards them and whether a Turkish-origin teacher can mitigate the possible negative effects of activation of the stereotype. They found that the highest learning gains came from students with Turkish as a family language when they were taught by a Turkish-origin teacher who actually named the stereotype, this went against the hypothesis of the authors. There was a significant interaction found between teacher origin and stereotype. For students with German as a family language, the main effect of activation of the stereotype was non-significant. Strong aspects of this study are that it was experimental, and it held a control group. Yet, the population studied was rather specific, only students of Turkish and German descent living in Germany were studied.

Seo and Lee (2021) studied both gender and ethnic/racial stereotype threats, using an experimental study with a classical randomized controlled trial design. They examined ninth grade students' naturalistic stereotype threat. Therefore, they focused on the control group in the experiment, for they wanted to assess the effects of naturalistic stereotype threats without influence of intervention messages of sorts. They reported several relevant findings. Firstly, a fixed teacher mindset climate had a statistically significant positive association with stereotype threat across all groups. The association was weaker for White girls. There were no significant differences found in the path of coefficients of fixed teacher mindset climate on anxiety across the three groups (Black, Latinx, and White participants). Furthermore, Black and Latino boys experienced the highest level of stereotype threat, then came Black and Latina girls, White boys and then White girls. The study used quite a large sample (N = 6040), used a separate control group and used covariates, namely previous mathematics achievement, mathematics course level, and family socioeconomic status.

All in all, the four studies showed that teachers can be influential in connection to stereotype threat for students in several aspects. First of all, their expectations and mindset could result in better student achievements. Higher expectations lead to better achievements, reduce low-trust in stereotyped groups (Yeager et al., 2014) and when a fixed teacher mindset is avoided, there is no association with stereotype threat (Seo & Lee, 2021). Furthermore, the gender and ethnicity of teachers is also influential (Master et al., 2014; Ollrogge et al., 2022), yet it is important to stay cautious with this conclusion because it is based on a small amount of studies.

3.1.2 Observational studies

There were four observational studies found. Two of which were focused on gender stereotype threat (Picho, 2016; Bowe et al., 2017), one on ethnic/racial stereotype threat (Phillips, 2019) and one of them was placed in the category 'other', and examined the stereotype of being labeled as 'smart' (Olitsky et al., 2010). The latter did mention minority groups, talking about income and racial background. Observational studies do not have the same strength of empirical evidence as experimental studies do, the studies found are mostly retrospective yet some of them are also longitudinal. They might give a better view of teacher-pupil relationships and teacher behavior without there being an intervention, although for one of the studies; the Hawthorne-effect could be a possibility. This is because during the study, one of the participants of the study (a teacher) changed the circumstances in their classroom as a result of the interviews held with students. This makes it evident that participants were aware of being studied.

Picho (2016) studied gender stereotype threat, yet did also mention some aspects of ethnic/racial stereotype threat. The study objective was the subjective experience of ninth-grade girls who are susceptible to mathematics-related stereotype threat in their authentic learning environments. The study was a phenomenological case study, which is a qualitative process to give the researcher insight in the rich experience of a group of individuals, to better understand them and their experiences (Dabengwa et al., 2023). Picho (2016) found that teacher-pupil relationships associate with the help-seeking behavior of students. Students that are highly susceptible to stereotype threat are often afraid to ask for help, this seems to be linked to status, they report not wanting to appear stupid. These same students also experience a different treatment of students by teachers, which causes them to shut down emotionally, keep a distance from their teacher, and tend to disengage. There were overt behavioral cues that made them feel as though the teacher had low expectations of them, which made them feel discouraged. When students who are highly susceptible to stereotype threat did not feel a connection to their teacher, they had negative attitudes towards them, and were dissatisfied with their pedagogy. The study was a case study of just one high school. Still, triangulation was used in order to select participants. Furthermore, the resulting sample was rather small (8 participants).

Bowe et al. (2017) also studied gender stereotype threat. They examined single-sex and coeducational urban elementary mathematics classes through situated cognitive theory, wherein they looked at achievement differences between genders. Some findings were that one of the teachers found it easier to be a role model in a single-sex classroom, and that it gives better

opportunities to incorporate gender-specific interests into the lessons. This is anecdotal, so whether this could be the case for all teachers in single-sex classrooms is not known. Furthermore, at the single-sex site, both classrooms made significant gains in test scores. This is interesting, because at the pre-test, girls tended to perform lower than the boys on the fifth-grade benchmarks. Boys and girls performed more similarly on the sixth-grade concepts. Whether this result can be contributed to the teacher is not clear. The study was a mixed-methods case study, of which the schools were selected based on volunteering and convenience; there were already established relationships between the researchers and the schools staff. Yet, to rule out some external variables, the context of the schools was chosen to be the same; same district, curriculum, district-wide support system, socioeconomic status of the student population, similar student mobility rates and both did not make annual academic progress the previous academic year. The study was also longitudinal, the test scores were based on several years prior the study and the observations were spaced out through the school year.

Phillips (2019) examined ethnic/racial stereotype threat. It was a mixed-methods case study, in which the researchers looked at engaging teachers at one racially and ethnically diverse public school using the self-system process model of student engagement. Findings indicated that teachers who are seen by students as 'warm demanding', make an effort to make a connection with students, are highly engaging, are available for individualized help during class, offer assistance to all students, use autonomy supportive behavior management, and are responsive to students' discomfort for being singled out in public. These factors are especially important for Black and Latinx students and students from other ethnicities than White. Furthermore, teachers that were considered below average warm demanding, were actually seen as behaving in ways that could be disrespectful, had a controlling style of behavior management and may exacerbate students' discomfort. For minority students this was when they were singled out in public, which would make them uncomfortable because they were especially afraid of making mistakes in public. It made them disengage. The study did not have a large sample size for the interviews conducted (N = 17), was not longitudinal and there is a possible selection bias, because of the guidance counselor who identified student interview participants.

Olitsky et al. (2010) focused on something else than gender or ethnic/racial stereotypes, even though there were mentions of it. They looked mostly at the label 'smart', which was mentioned a lot by students in their research. The study objective was to find the reason for some students in an urban magnet school tot not identify and achieve well in science, even though they find it interesting. Results showed that students at the school in which the research

took place were not feeling as though teachers and staff had high expectations of them. Students without a label as 'smart' avoided participation in class, in fear of getting negative feedback from teachers and peers. This poses a barrier to socially situated learning in science, according to the authors. Receiving a label as 'smart' was needed for social acceptance at the school. Also, students perceived making an effort for their studies as being less smart, because if they were smart, they would already be able to do the task. This hierarchy resulting from the 'smart' label contributed to students not trusting their own abilities in relation to other students and it constrains some of the students from speaking in class. Interestingly enough, most of the students that did not receive the label 'smart' were from low-income, predominantly African American neighborhoods. The study was ethnographic, yet longitudinal, the research was conducted over the course of three years. For the research, student researchers participated in the study and also interviewed other students. The sample size was not very large, at least not for the cogenerative dialogues. Furthermore, the study is very context specific because of the type of research being ethnographic. Therefore, the results are not generalizable to a bigger population. To enhance the validity of the research, different participant researchers were used in the coding to compare descriptions and counts. Also, methodological triangulation was used.

In summary, the studies revealed that teacher-pupil relationships can be very influential for students. They can affect the help-seeking behavior of students and making an effort to connect with students can result in more engagement in the classroom (Phillips, 2019; Picho, 2016). This is especially the case for students from stigmatized groups, such as ethnic minorities. Furthermore, when students from stigmatized groups do not get a 'smart' label, they tend to avoid participation in class, make less of an effort for their studies and it contributes to students not trusting their own abilities (Olitsky et al., 2010). When a teacher shows warm demanding behavior, there will be higher engagement in the classroom, especially for students from stigmatized groups (Phillips, 2019). Furthermore, some teachers viewed single-sex classrooms as having better opportunities to incorporate gender-specific interests into the lessons and they thought it easier to be a role model for the same sex (Bowe et al., 2017).

3.1.3 Cross-sectional studies: interviews/surveys

Five of the studies consisted of either interviews, surveys, or both (Breidebach & Gruber, 2018; Dandy et al., 2015; Edwards, 2018; Haenni Hoti et al., 2019; Nouwen & Clycq, 2019). Two of the studies focused on ethnic/racial stereotype threat, two of them focused on both gender and ethnic/racial stereotype threat, one focused on the stereotype of children raised by grandparents. Interviews and surveys in the included studies are usually from the perspective

of the participant, and therefore are not always generalizable to a larger group. Yet, it can also be valuable to gather the perspective of the participants.

Haenni Hoti et al. (2019) examined ethnic/racial stereotype threat. The study objective was the examination of the impact of various constellations of acculturation orientations of fifth grade immigrant students and their teachers on the students' school adjustment. Results indicated that a better quality of teacher-pupil relationships goes along with a significantly higher school satisfaction for all students. Orientation regarding to acculturation makes a difference for school adjustment as well, whether or not the teacher has a similar orientation or not for students from minority groups. But still, the quality of teacher-pupil relationships is a more dominant predictor of school satisfaction for these students. The better the relationship between teachers and students, the better the academic self-concept of students, aside from the match or mismatch of acculturation orientations and (bi-)national identification. Furthermore, it seems that students with younger teachers had a significantly higher school satisfaction than students with older and more experienced teachers. For the latter group, a mismatch between the students' (bi-)national identification at school and the teacher's perception of the student as an immigrant or native born student could negatively affect the student's academic self-concept. This was not the case for the group of students with young teachers. When there are matching views of the students' (bi-)national identification, they have a better academic self-concept. Also, the most important predictor of the academic self-concept of students was the educational background, not the migration background of students or the matching/mismatching of acculturation orientations. Still, teacher-pupil relationships remained a better predictor of school satisfaction in general. For their study, they used surveys and a multiple regression analysis. The surveys were self-reported by students, but teachers also completed a questionnaire concerning their acculturation orientations and relationship with the students. The teacher surveys were used to gather information about the match or mismatch of acculturation orientations. The sample size was relatively large (N = 1106) and participants were spread out over 60 school classes, 42 schools and eight states in Switzerland. Cronbach's alpha was assessed for some of the items in the questionnaires and it was either good or acceptable (all above .70). Dandy et al. (2015) researched ethnic/racial stereotypes as well. They studies the ethnic group differences in (expectations regarding) academic achievement among Australian students, where they looked specifically at Aboriginals, Anglo-Australian and Asian Australian students. It was found that overall in mathematics, there are higher expectancies for Chinese Australian students as compared to Anglo-Australian and Aboriginal students. Expectancies for

Anglo-Australian students were higher than for Aboriginal students. Overall in English, expectancies for Aboriginal students were consistently lower than for Chinese Australian and Anglo-Australian students. This was the case for both teacher-trainees and experienced teachers. Trainee-teachers also perceived Chinese-Australian and Anglo-Australian students as having more family support than the Aboriginal student. These results align with the expectancies of their own ethnic group, this was primarily the case in mathematics. Asian Australian students had above-average expectancies for themselves, Aboriginal students had very low expectancies. Anglo-Australian students were, again, in the middle. The questionnaires used in this study were self-reported. Furthermore, this was not a longitudinal study and the subgroups were very context specific for Australia.

Nouwen and Clycq (2019) studied both gender and ethnic/racial stereotype threat. They tried to find out the stereotype threat effects for Flemish third- and fourth-year students in secondary education, specifically regarding teacher-pupil relationships. Firstly, they found a strong positive relation between the GPA and the academic self-concept of pupils in the least negatively stereotyped groups. Also, teacher-pupil relationships play a stronger role in relation to pupils' academic self-concept for pupils in the most stigmatized groups: pupils in vocationally oriented education, and pupils with an immigrant background. When pupils have a negative relationships with their teacher, they are more likely to discount negative feedback from teachers as being unfairly biased. The authors therefore concluded that experiencing low levels of respect and acceptance from teachers strongly increases the occurrence of stereotype threat effects. Furthermore, the authors discovered that having a Turkish or Moroccan background can protect pupils from disidentification from education, but not experiencing positive teacher-pupil relationships has more predictive value. All in all, the authors conclude that teachers' perceptions and expectations play an important role in the engagement and identification with education of pupils, especially for those with a stigmatized group identity in education. Nouwen and Clycq (2019) did look at a very specific subgroup with an immigration background: the research was only linked to Turkey or Morocco and it zoomed in on just the Flemish part of Belgium. There was not a longitudinal design. Also, the surveys were self-reported by pupils. There may have also been a Hawthorne effect, where the pupils know that they are being studied on a specific presumed relationship. Breidebach and Gruber (2018) also looked at both gender and ethnic/racial stereotype threat. Their study objective consisted of the influence of pupil and teacher characteristics on the diagnostic competence of teachers in the context of educational disadvantage in Germany. It was found that both migration background

and gender can influence the assessments of teachers. Female pupils are generally assessed higher in learning behavior and pupils without a migration background have more accurate assessments than those with a migration background. Teacher gender was also of influence, female teachers tended to rate migrants slightly better than male teachers did and rated male and female pupils the same in terms of learning behavior. Only experienced teachers assessed students with a migration background impartially, so years of teaching experience is influential as well. The assessments in the study were made by teachers. For the study, teacher bias was taken out of the equation and environmental factors were considered. A correlational survey study was used, with descriptive comparison. The population consisted of 11th grade students and teachers from Germany, yet some students were over 18. The study was still incorporated here because students with a migration background can often be a bit older than their classmates. The size of the achieved sample was 156 students, but the majority was female (N = 119), which makes it harder to draw conclusions from a balanced sample.

Edwards (2018) focused on another stereotype threat, namely the children raised by grandparents. The study objective was to make a comparison of children's perceptions of their teachers' views of their school performance and children's general feelings about school, between children raised by grandparents and children raised by parents and foster parents. It was a cross-sectional survey design with a relatively large sample (N = 6.550). Edwards (2018) found that children raised by grandparents did believe they were viewed differently by teachers, but that this is not the case in reality. Children raised by grandparents did not have negative feelings about school when compared to their peers. Sampling was random and observations were independent, there was a robust dataset.

All in all, it becomes clear from the presented studies that a better quality of teacher-pupil relationships is associated with a higher school satisfaction and a better academic self-concept for students, especially those in the most stigmatized groups, even when there is a mismatch of acculturation orientations and (bi-)national identification (Haenni Hoti et al., 2019; Nouwen & Clycq, 2019). Having matching views of students' (bi-)national identification does help for a better academic self-concept (Haenni Hoti et al., 2019), same as with a GPA for students from the least stigmatized groups (Nouwen & Clycq, 2019), but it is not as strong a predictor as teacher-pupil relationships are. When there is a negative teacher-pupil relationship, the pupil is more likely to discount negative feedback as unfairly biased and low levels of respect and acceptance from teachers can go along with stereotype threat effects (Nouwen & Clycq, 2019). Furthermore, higher expectations from teachers go along with better engagement

and identification with education of pupils (Nouwen & Clycq, 2019). It seems that those expectations align with students' expectations of their own ethnic group, and teachers do tend to have lower expectations of certain ethnic groups (Dandy et al., 2015). Furthermore, children raised by grandparents believed they were viewed differently, yet in their opinion, this was not the case (Edwards, 2018). Lastly, migration background and gender can influence teachers' diagnostic competences of learning behavior (Breidebach & Gruber, 2018).

3.2 Factors in teacher-pupil relationships that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education

A few elements of teacher-pupil relationships that have a positive influence on the academic performance of students that suffer from stereotype threat in primary and secondary education could be identified in the literature discussed in the previous paragraph: being 'warm demanding', making an effort to connect with students, being highly engaging, being available for individualized help during class, offer assistance to all students, having autonomy supportive behavior management, being responsive to students' discomfort for being singled out in public (Phillips, 2019).

There were also studies that mainly emphasized the importance of high quality teacher-pupil relationships in the academic performance of students, and reported positive effects: it positively affects help-seeking behavior of students, which is something that students susceptible to stereotype threat tend to avoid (Picho, 2016), plays a strong role in relation to the academic self-concept of pupils in the most stigmatized groups (Haenni Hoti et al., 2019; Nouwen & Clycq, 2019), and goes along with a significantly higher school satisfaction for students susceptible to stereotype threat as well as students who are not (Haenni Hoti et al., 2019).

When teacher-pupil relationships are negative, there were also negative effects reported: pupils are more likely to discount negative feedback from teachers as unfairly biased and low levels of respect and acceptance from teachers increases the occurrence of stereotype threat effects (Nouwen & Clycq, 2019).

3.3 Factors in teacher behavior that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education

A few elements of teacher behavior that have a positive influence on the academic performance of students that suffer from stereotype threat in primary and secondary education

could be identified in the reviewed studies: (for certain ethnic groups) having the stereotype named by a teacher from the same origin (Ollrogge et al., 2022), being 'warm demanding', making an effort to connect with students, being highly engaging, being available for individualized help during class, offer assistance to all students, having autonomy supportive behavior management, and being responsive to students' discomfort for being singled out in public (Phillips, 2019).

There were also elements of teacher behavior found that could negatively influence the academic performance of students that suffer from stereotype threat in primary and secondary education: treating students differently and having overt behavioral cues that indicate low expectations (Olitsky et al., 2010; Picho, 2016), behaving disrespectful to students, having a controlling style of behavior management, exacerbate students' discomfort (Phillips, 2019), and (not) labeling students as 'smart' (Olitsky et al., 2010).

These elements could cause negative effects for marginalized students: shutting down emotionally, keeping a distance from the teacher, disengaging, discouragement, having negative attitudes towards their teacher, being dissatisfied with their pedagogy (Picho, 2016), avoiding participation in class, being in fear of negative feedback, and having a barrier to learning (Olitsky et al., 2010).

3.4 Factors in teacher expectations that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education

There were also elements regarding teacher expectations specifically, that have a positive influence on the academic performance of students that suffer from stereotype threat in primary and secondary education: high teacher expectations lead to students revising their essays and receiving 'wise feedback' led to them achieving better (Yeager et al., 2014). Potential positive effects of high teacher expectations are more engagement and identification with education of pupils, which is especially the case for students from a stigmatized group (Nouwen & Clycq, 2019).

Furthermore, elements of teacher expectations that could negatively influence the academic performance of students that suffer from stereotype threat in primary and secondary education were: a fixed teacher mindset (Seo & Lee, 2021), low expectations (Dandy et al., 2015), and a mismatch between the students' (bi-)national identification at school and the teachers' perception of the student as an immigrant or native born student (Haenni Hoti et al., 2019).

These elements could cause negative effects as well: the low expectancies of teachers could align with the expectancies of students from their own ethnic group (Dandy et al., 2015), and a mismatch in identification could lead to a lower academic self-concept (Haenni Hoti et al., 2019).

Some elements that were found were not directly related to teacher behavior, but were influential teacher characteristics. A few examples are the age of the teacher (Haenni Hoti et al., 2019), the experience of the teacher (Breidebach & Gruber, 2018), the gender of the teacher (Bowe et al., 2017; Breidebach & Gruber, 2018; Master et al., 2014), and the origin of the teacher (Breidebach & Gruber, 2018; Ollrogge et al., 2022). Gender was mentioned to be influential, because for the same gender it might be easier for a teacher to be a role model (Bowe et al., 2017).

Lastly, in one of the studies the students from a stigmatized group actually assumed low expectations from teachers, even when this was not the case. Still, it did not result into them feeling negatively about school (Edwards, 2018).

4. Discussion

4.1 Conclusions

In the present study, a systematic literature review was conducted to research which factors in teacher-pupil relationships, teacher behavior and teacher expectations can be identified that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education. This was done in order to organize the information about teachers and stereotype threat in existing literature and to contribute to information about what works to reduce stereotype threat effects occurring in education. The systematic literature review indicated broadly a few factors that are of influence concerning teacher-pupil relationships, teacher behavior and/or teacher expectations. Based on the systematic literature review conducted, some of these elements appeared only relevant for children from stigmatized groups, some of these elements appeared relevant for all students.

The first research question aimed to make clear which factors in teacher-pupil relationships can be identified that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education (Haenni Hoti et al., 2019; Nouwen & Clycq, 2019; Phillips, 2019; Picho, 2016). Results of the present study indicated that one of these factors is that the teacher is warm demanding and thereby generates social capitol which makes students want to participate in class. This is especially the case for students from stigmatized groups. Furthermore, another factor is the teacher offering assistance to all students, even ones who behave disrespectfully, is a positive factor as well. Lastly, the quality of the teacher-pupil relationship has to be strong in order to generate positive effects on school satisfaction and academic self-concept. This is especially important for students from stigmatized groups as well.

The second research question aimed to make clear which factors in teacher behavior can be identified that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education (Olitsky et al., 2010; Ollrogge et al., 2022; Phillips, 2019; Picho, 2016). Two factors could be identified. Firstly, as was indicated in the answer to the first research question, warm demanding behavior can generate positive effects on the academic performance of students susceptible to stereotype threat. It leads to better engagement. Teachers that displayed warm demanding behavior were friendly, declared their affinity for their students and took an interest in their lives, let students get to know them, were available for individualized help during class, had an autonomy supportive behavior management style, made it possible to make mistakes, clarify their questions, give students the

right to opt-out, and use jokes and simple questions to refocus the attention. Secondly, another factor that positively influenced the academic performance of students susceptible to stereotype threat was the teacher naming the stereotype threat, but only when the teacher was of the same origin as the student.

The third research question aimed to make clear which factors in teacher expectations can be identified that positively influence the academic performance of students that suffer from stereotype threat in primary and secondary education (Bowe et al., 2017; Breidebach & Gruber, 2018; Dandy et al., 2015; Edwards, 2018; Haenni Hoti et al., 2019; Master et al., 2014; Nouwen & Clycq, 2019; Ollrogge et al., 2022; Seo & Lee, 2021; Yeager et al., 2014). An important factor is that when teachers have high expectations in general, expressing this either verbally or through their behavior, this has a positive influence on students. For minority students, this increased their motivation. Furthermore, an intervention concerning 'wise feedback' was effective as well. It reduced mistrust in school and performance for students susceptible to stereotype threat.

In conclusion, the aforementioned elements (the teacher is warm demanding, offers assistance to all students, the quality of the teacher-pupil relationship is strong, warm demanding behavior, naming the stereotype, high expectations, and wise feedback) can contribute to better achievements for students that suffer from stereotype threat in primary and secondary education, although the strength of the empirical evidence of many of the included studies was limited. For most of these elements, the population researched was a stigmatized group. The stereotype that was studied most was the ethnic/racial stereotype, usually concerning students with an immigrant background.

Looking back at the possible minor influence stereotype threat effects actually hold because of the publication bias in the research field, as was mentioned in the introduction, the importance of the results of the present review study can be questioned (Agnoli et al., 2021; Flore et al., 2018; Flore & Wicherts, 2015; Rosenthal, 1979). Still, when there are possible negative effects of stereotype threat and there is a positive influence teachers can have to negate this – there is no harm in applying the behaviors found to be effective.

Furthermore, before the present review study it was expected that the three aspects that were examined (teacher-pupil relationships, teacher behavior, teacher expectations) are connected in the literature (Nouwen & Clycq, 2019; Kranz et al., 1970). This was the case in some of the literature, four studies made a connection between two or more of the aspects

(Haenni Hoti et al., 2019; Olitsky et al., 2010; Phillips, 2019; Picho, 2016). It was also expected that the review would point out empirical studies that support the idea that when teachers hold negative implicit biases, this influences their relationships with pupils and their behavior in the classroom and reinforces stereotype threat for their stereotyped students (DeCuir-Gunby & Bindra, 2022; Seo & Lee, 2020). This was not explicitly mentioned, but it can be concluded that teacher expectations are researched most, seeing as ten of the found studies mentioned expectations. Therefore, it seems that in the literature found teacher expectations can be seen as influential.

Furthermore, most of the studies found were conducted in secondary education. This could indicate that stereotype threat effects emerge later than primary education age between 4 and 12, going against the previous hypothesis that it could already emerge as early as the age of 7 (Eccles et al., 1989; Nicholls, 1979). Yet, it could also indicate that there is less research available on certain age groups. Further research is necessary in order to establish this.

4.2 Limitations

The present study contains a few limitations. Firstly, the limitations of the articles found will be discussed. Secondly, the limitations of the present study itself will be discussed.

First of all, some of the studies found talk about 'strong effects' or indicate causal relationships, yet they are not experimental studies (Haenni Hoti et al., 2019; Nouwen & Clycq, 2019; Olitsky et al., 2010). It is important to remain critical of such claims when a study is not experimental and therefore cannot rule out contextual influences. Secondly, some of the studies were examining very specific populations, such as students of Turkish and German descent living in Germany (Ollrogge et al., 2022) or a case study from just one high school (Picho, 2016). Also, some sample sizes in studies were relatively small, ranging between 8 to 33 participants (Bowe et al., 2017; Olitsky et al., 2010; Picho, 2016). The specific populations and small sample sizes make it harder to generalize the results found towards broader populations. Furthermore, some of the surveys that were incorporated were self-reported by participants (Haenni Hoti et al., 2019; Master et al., 2014; Nouwen & Clycq, 2019; Seo & Lee, 2021). This means that the results of those studies were subjective, largely focusing on perceived stereotype threat rather than stereotype threat effects. Yet, it is important to note that this information could actually be valuable because of the subjectiveness, which is sometimes part of the aim of a qualitative study. This is especially the case for the present study, because it studies stereotype threat, in which case personal experiences of participants are very important. Similarly, there

was one phenomenological study and one ethnographic study. In these type of studies, the researcher sometimes takes part in the study as well. The results found in these studies are very context specific and usually cannot be generalized toward the population. This is especially the case for the phenomenological study, seeing as ethnographic research typically tries to define the collective experiences of the community while phenomenological research highlights the diverse experiences of members of such a community (Jamali, 2018). Yet, these results might be very specific, and the circumstances might change too, so remaining critical of the results is of importance here as well. Furthermore, interviews were apart of some studies too. This means that some of the results were anecdotal. This was especially the case for one study, where there were just four teachers interviewed and some of the conclusions were directly made from what they had said, even though it might have been the opinion of just a single teacher (Bowe et al., 2017).

Second of all, the present study contains some limitations too. One of the limitations is that some of the elements identified in the answers to the research questions were supported by just one or two studies. Further research would be needed in order to confirm these elements as influential factors in teacher-pupil relationships, behavior and expectations for the academic achievements of students susceptible to stereotype threat. The amount of studies used in the present study is not a lot either, many articles found did not align with the inclusion and exclusion criteria. A follow-up study could expand the inclusion criteria (such as not just looking at relationships, behavior and expectations but teacher influence in general), use more search terms, search through more databases and/or include more languages (instead of just English) in order to get more elaborate results. A final limitation could be that only one researcher analyzed the articles found, this may have led to an objectivity problem. Yet, there was a supervisor that provided feedback on the data extraction and the analysis and peer feedback was provided at one point in the process.

4.3 Recommendations

With the present study, there is an integral summary of important factors in teachers' relationships with pupils, their behavior and their expectations which holds information that can be communicated towards teachers or trainee-teachers. Therefore, this systematic literature review can give them insights into important competences to focus on, which they can use in their teaching practices. Furthermore, educational policies can accommodate teachers in striving to reach the elements found in the present study. Trainee-teachers can be better prepared for educational practice, being aware of the possible stereotype threat effects they would help

create if they are not mindful of their behavior. Therefore, it is recommended to create awareness and accommodate (trainee-)teachers in becoming teachers that are mindful of stereotype threat. Also, if possible, it would be recommended to have a diverse teaching staff, seeing as the gender, ethnicity and experience of the teacher are, at least in the perception of stigmatized groups, influential factors as well (Bowe et al., 2017; Breidebach & Gruber, 2018; Haenni Hoti et al., 2019; Ollrogge et al., 2022; Master et al., 2014).

The elements identified in the present study are mostly based on just one or two studies. For example, the teacher being warm demanding with the according behavior as described before (Phillips, 2019), having the stereotype named by a teacher from the same origin (Ollrogge et al., 2022), and the wise feedback intervention (Yeager et al., 2014) were all elements based on just a single study. In order to make sure that these are indeed the elements in teacher-pupil relationships, teacher behavior and expectations that are of influence on the achievements of students who suffer from stereotype threat, it is recommended that more research is done to confirm those effects, especially regarding these specific behaviors and/or interventions. Furthermore, most of the studies found focused on secondary education. More research in primary education is necessary in order to establish if stereotype threat effects emerge this early on. It would also be rather interesting to have an experimental study regarding the positive influence teachers can have in negating stereotype threat effects, to see if there are more elements to be found that could positively influence the academic achievements of children susceptible to stereotype threat. Currently, the studies that mention teachers when examining stereotype threat effects, were looking at the stereotype in general – not at the influence of the teacher specifically. More research on the specific role of the teacher in stereotype threat effects is needed to shed light on this matter, in order to find more specific results instead of generic.

5. References

Agnoli, F., Melchiorre, F., Zandonella Callegher, C., & Altoè, G. (2021). Stereotype threat effects on Italian girls' mathematics performance: A failure to replicate. *Developmental Psychology, 57*(6), 940-950. https://doi.org/https://doi.org/10.1037/dev0001186

Bowe, A. G., Desjardins, C. D., Covington Clarkson, L. M., & Lawrenz, F. (2017). Urban Elementary Single-Sex Math Classrooms: Mitigating Stereotype Threat for African American Girls. *Urban Education, 52*(3), 370–398.

Cambridge Dictionary (2022). S*tereotype definition: 1. a set idea that people have about what someone or something is like, especially an idea that is. ... Learn more.* https://dictionary.cambridge.org/dictionary/english/stereotype

Dabengwa, I. M., Young, S., & Ngulube, P. (2023). Rigour in phenomenological and phenomenography studies: A scoping review of library and information science research. *Library & Information Science Research*, *45*(1), 101219. https://doi.org/10.1016/j.lisr.2022.101219

DeCuir-Gunby, J. T., & Bindra, V. G. (2022). How does teacher bias influence students?: an introduction to the special issue on teachers' implicit attitudes, instructional practices, and student outcomes. *Learning and Instruction, 78.*

Desert, M., Preaux, M., & Jund, R. (2009). So Young and Already Victims of Stereotype Threat: Socio-Economic Status and Performance of 6 to 9 Years Old Children on Raven's Progressive Matrices. *European Journal of Psychology of Education, 24*(2), 207–218.

Eccles, J. S., Wigfield, A., Constance, A., Miller, C., Reuman, D. A., & Yee, D. (1989). Self-esteem: Relations and changes at early adolescence. *Journal of Personality, 57*, 283–310.

Flore, P. C., Mulder, J., & Wicherts, J. M. (2018). The influence of gender stereotype threat on mathematics test scores of Dutch high school students: A registered report. *Comprehensive Results in Social Psychology, 3*(2), 140-174. https://doi.org/10.1080/23743603.2018.1559647

Flore, P. C., & Wicherts, J. M. (2015). Does stereotype threat influence performance of girls in stereotypes domains? A meta-analysis. *Journal of School Psychology, 53*, 25-44. https://doi.org/http://dx.doi.org/10.1016/j.jsp.2014.10.002

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure? *Child Development, 76*(5), 949–967. Retrieved from http://www. jstor.org/stable/3696607

Jamali, H. R. (2018). Does research using qualitative methods (grounded theory, ethnography, and phenomenology) have more impact? *Library & Information Science Research*, *40*(3-4), 201-207.

Katz, P. A., & Kofkin, J. A. (1997). Race, gender, and young children. In S. S. Luthar, J. A. Burack, D. Cicchetti, & J. R. Weisz (Eds.), *Developmental psychopathology: Perspectives on adjustment, risk, and disorder* (pp. 51–74). Cambridge, UK: Cambridge University Press.

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths performance. *British Journal of Educational Psychology, 77*, 323-338. https://doi.org/10.1348/000709906X113662

Kranz, P. L., & And Others. (1970). *The Relationships Between Teacher Perception of Pupils and Teacher Behavior Toward Those Pupils*.

Lorenz, G. (2021). Subtle Discrimination: Do Stereotypes among Teachers Trigger Bias in Their Expectations and Widen Ethnic Achievement Gaps? *Social Psychology of Education: An International Journal*, *24*(2), 537–571.

Ollrogge, K., Frühauf, M., Mros, T., Böttger, J., Höhne, E., McElvany, N., Zander, L., & Hannover, B. (2022). Can I Dismiss the Stereotype -- as My Teacher Did? Influence of Stereotype Activation and an Immigrant Teacher on Student Learning. *Social Psychology of Education: An International Journal, 25*(4), 745–766

Nicholls, J. G. (1979). Development of perception of own attainment and causal attributions for success and failure in reading. *Journal of Educational Psychology, 71*, 94–99.

Nouwen, W., & Clycq, N. (2019). The Role of Teacher-Pupil Relations in Stereotype Threat Effects in Flemish Secondary Education. *Urban Education*, *54*(10), 1551–1580.

Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons.

Pianta, B. (2016). Classroom processes and teacher-student interaction: Integration with a developmental psychopathology perspective. In D. Cicchetti (Ed.), *Developmental psychopathology: Vol. 4. Risk, resilience, and intervention* (3rd ed., pp. 770–814). Hoboken, NJ: Wiley.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Seo, E., & Lee, Y. (2021). Stereotype threat in high school classrooms: How it links to teacher mindset climate, mathematics anxiety, and achievement. *Journal of Youth and Adolescence, 50*(7), 1410–1423. https://doi-org.proxy-ub.rug.nl/10.1007/s1096402101435-x

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science, 10*(1), 80–83. https://doi.org/10.1111/1467-9280.00111

Smith, J. L., & White, P. H. (2001). Development of the Domain Identification Measure: A tool for investigating stereotype threat effects. *Educational and Psychological Measurement, 61*, 1040-1057. https://doi.org/10.1177/00131640121971635

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28. https://doi.org/https://doi.org/10.1006/jesp.1998.1373

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of personality and social psychology, 69*(5), 797. https://doi.org/https://doi.org/10.1037/0022-3514.69.5.797

Steele, C. (2010). *Whistling Vivaldi: And other clues to how stereotypes affect us*. New York, NY: Norton.

Stoevenbelt, A. H., Flore, P. C., Schwabe, I., & Wicherts, J. M. (2022). The uniformity of stereotype threat: Analyzing the moderating effects of premeasured performance. *Intelligence, 93*, 101655.

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

Appendix 1

Empty Data Extraction Form

Data to be extracted	*Notes to reviewer*
Title of study	
Author	
Year of publication	
Population	
Setting	Town/region and country of intervention
Time	Time when study took place (rather than date of publication)
Study objective clearly stated?	
Study methodology (or methodologies) used	
Inclusion of sufficient data to assess validity of conclusions?	Including statistical tables; or direct quotes from respondents
Data source	(e.g. interviews, survey, etc.)
Size of achieved sample	Also include size of total population if stated
Length of follow up	

Teacher-pupil relationships	
Mention of (aspects related to) teacher-pupil relationships	
Relation between teacher-pupil relationships and student achievements or achievement gap	
Relation between teacher-pupil	

relationships and stereotype threat	

Teacher behavior	
Mention of (aspects related to) teacher behavior	
Relation between teacher behavior and student achievements or achievement gap	
Relation between teacher behavior and stereotype threat	
Relation between teacher expectations and stereotype threat	

Teacher expectations	
Mention of (aspects related to) teacher expectations	
Relation between teacher expectations and student achievements or achievement gap	

Stereotype threat	
Mention of (aspects related to) stereotype threat	

Method of measurement	
Separate control group?	

Validity	
Overall impression of internal validity (low, medium, high)	Assessment based on the quality of the sampling and response and the treatment of confounding factors
External validity	In some circumstances, the results of a study may be so dependent on the local context that it must be considered to have low external validity regardless of the quality of the study
Conditions affecting external validity	

Overall summary of study	

Appendix 2

Table with Reviewed Articles

Number	Study	Country	Study objective	Research design	Sample	Overall results	Specific results for specific groups
1	Nouwen & Clycq (2019)	Belgium	Teacher-pupil relationships and stereotype threat effects	Correlational survey study	6.244 pupils in the third and fourth year of secondary education (surveys)	Negative teacher-pupil relationships leads to students discounting negative feedback as unfairly biased Teachers' perceptions and expectations play an important role in the engagement and identification of pupils	For pupils in the most stigmatized groups (pupils in vocationally oriented education, and pupils with Turkish or Moroccan backgrounds): positive relation GPA and academic self-concept disconnects, and teacher-pupil relationships play a stronger role in relation to academic self-concept

2	Picho (2016)	United States of America	The subjective experience of ninth-grade girls susceptible to mathematics-related stereotype threat in their authentic learning environments	Phenomenological study, observational case study	2.300 girls from ninth grade, 8 girls from ninth grade selected for in-depth interviews	Girls with high susceptibility for stereotype threat experienced low levels of support from and a negative relationship with their teachers, which resulted in them disconnecting from the learning material and being less inclined to reach out for help	The stereotype girls were most aware of being held against them was concerning their ethnicity
3	Haenni Hoti et al. (2019)	Switzerland	Examination of the impact of various constellations of acculturation orientations of immigrant		1.106 fifth graders, 60 teachers spread out over 60 school classes, 42 schools, from eight German	Congruency between the attitudes regarding cultural assimilation and diversity leads to better school satisfaction. It was	Immigrant students with a minority orientation whose teachers value cultural diversity showed a lower academic self-

#	Author	Country		Design	Sample		
			students and their teachers on the students' school adjustment (in 5th grade).		speaking federal states in Switzerland	evident that the quality of teacher-student relationships was a bigger influence than the congruency of attitudes regarding cultural assimilation	concept. If there was a mismatch between the students' (bi-)national identification at school and the teacher's perception of the student's status as immigrant or native-born, this could have an adverse effect on the academic self-concept of these students.
4	Breidebach & Gruber (2018)	Germany	The influence of pupil and teacher characteristics on the diagnostic competence of	Correlational survey study	156 students from six 11th grade classes, assessed by 18 teachers	Expectations of teachers do not always align with the actual grades of students	The overall results are especially accounting for female students and students with a migration background

			teachers in the context of educational disadvantage in Germany				
5	Edwards (2018)	United States of America	A comparison of children's perceptions of their teachers' views of their school performance and children's general feelings about school, between children raised by grandparents and children raised by	Cross-sectional survey	6.550 children in grades 6-8	Overall results align with results for the specific group (children raised by grandparents)	Children raised by grandparents believe teachers view their school performance negatively when compared with children living with both parents.

			parents and foster parents				
6	Philips (2019)	United States of America	Engaging teachers at one racially and ethnically diverse public school using the self-system process model of student engagement.	Mixed-methods case study	Tripod survey: 1257 students Interviews: 17 students Two class sessions per teacher participant were observed; six teachers, 12 observations in total.	Teachers that score high on warm demanding in the survey and of whom students also indicated warm demanding behavior in interviews, have more engagement of students in the classroom	For Black, Latinx and students from other ethnicities (as opposed to white students), responsiveness is of great importance. One-on-one help is especially helpful for them.
7	Dandy et al. (2015)	Australia	Studying the ethnic group differences in academic achievement	Survey study	55 experienced teachers, 144 training teachers, 516 school students	Overall results align with results for the specific group	Asian students were expected to perform better and expend greater effort in mathematics than

			among Australian students, looking specifically at Aboriginals, Anglo-Australian and Asian Australian students.				Aboriginal and Anglo-Australian students. When looking at Aboriginals and Anglo-Australian students, Anglo-Australian students were expected to perform better.
8	Bowe et al. (2017)	United States of America	Examining single-sex and coeducational urban elementary mathematics classes through situated cognitive theory, looking	Mixed-methods case study	32 student interviews at the single-sex site, 13 student interviews at the coed site, one structured interview with each of the classroom	Teacher buy-in, congruency in students' perceptions about their learning environment, teacher's role in fostering positive academic identities, and exacerbated	Especially African American girls could profit from single-sex classrooms.

			at achievement differences between genders.		teachers (4) and principals (2) at both sites	social identities are important contextual factors within these environments that go beyond the mere separation of the sexes.	
9	Yeager et al. (2014)	United States of America	Examining the effects of a strategy to restore trust on minority adolescents' responses to critical feedback	Longitudinal experimental study	Sutdy 1: 44 seventh-grade students in three social studies classrooms at a suburban public middle school on the intermediate performance range. Study 2: new cohort of 44 students form the	Overall results align with results for the specific group.	Mistrusting African American students benefited the most from critical feedback in the experimental group. Receiving high expectations from their teacher had a significant effect on their performance in the revised essays.

					same three seventh-grade social studies teachers' classrooms as study 1, from the subsequent year. Study 3: 76 students from a medium-sized urban public high school.		
10	Ollrogge et al. (2022)	Germany	Whether in students of Turkish descent learning vocabulary is impaired when the teacher activates the negative	Experimental survey study	182 students with German or Turkish as a family language	Overall results align with results for the specific group.	Turkish-origin students learned significantly more under stereotype activation with the Turkish-origin teacher than in all other conditions.

			stereotype towards them and whether a Turkish-origin teacher can mitigate negative effects of the activation of the stereotype.				
11	Olitsky et al. (2010)	Not clear, presumably United States of America	Finding the reason for some students in an urban magnet school to not identify and achieve well in science, even though they find it interesting.	Longitudinal ethnographic study	33 students	Structural issues such as the school's selection process, the discourses perpetuated by teachers, administrators, and peers regarding "who belongs" at the school, and negative stereotype	Does not apply

						threat posed obstacles for students by highlighting rather than mitigating the inequalities in students' educational backgrounds.	
12	Master et al. (2014)	United States of America	How stereotypes about girls' ability in computer science affect their concerns about being negatively stereotyped / How teacher gender and	Experimental survey study	Experiment 1: 218 students Experiment 2: 193 students	Female teachers evoke fewer concerns from girls about negative stereotypes in computer science than male teachers. This is especially the case when threatening cues, or cues that signaled the strong	Does not apply

			stereotype threat affect adolescents' self-reported concerns about being negatively stereotyped in computer science courses			likelihood that they would be judged through the lens of negative gender stereotypes, were present. The result was broadly the same for both experiments.	
13	Seo & Lee (2021)	United States of America	Examining ninth grade students' naturalistic stereotype threat The present research focuses on participants in the control group because it	Experimental survey study	6.040 students	Overall results align with results for the specific group.	Black and Latinx students experience higher levels of stereotype threat in high school mathematics classrooms than do their White peers. When students perceive that their teachers have created

aims to assess the effects of naturalistic stereotype threats that were not influenced by intervention messages

fixed mindset climates, they experience greater stereotype threat. Stereotype threat, in turn, negatively predicts Black and Latino boys and White girls' later achievement via anxiety.

Appendix 3

Matrix of Studies Categorized by Type of Education and Type of Element Studied

	Primary education	Secondary education	Primary and secondary education
Teacher-pupil relationships		Nouwen & Clycq (2019)	
Teacher behavior		Ollrogge et al. (2022); Master et al. (2014)	
Teacher expectations	Bowe et al. (2017)	Breidebach & Gruber (2018); Seo & Lee (2021)	Edwards (2018); Dandy et al. (2015); Yeager et al. (2014)
A combination of two or more of the above	Haenni Hoti et al. (2019)	Picho (2016); Phillips (2019); Olitsky et al. (2010)	