Gender, Self-concept and Mathematics and Science Performance of South African Grade 9 Students

DEBRA LYNNE SHEPHERD

Stellenbosch Economic Working Papers: WP11/2017

www.ekon.sun.ac.za/wpapers/2017/wp112017

October 2017

KEYWORDS: gender inequality, STEM, stereotype, performance, self-concept, development
JEL: C21, I21, I24, J16

ReSEP (Research on Socio-Economic Policy)
http://resep.sun.ac.za

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH

SOUTH AFRICA

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

www.ekon.sun.ac.za/wpapers

Gender, Self-concept and Mathematics and Science Performance of South African Grade 9 Students[1]

By Debra Lynne Shepherd[2]

ABSTRACT

Despite improvements over the past decade, South African women continue to be underrepresented in tertiary studies and professional careers in the fields of science, technology, engineering and math. This has implications not only for economic development and growth, but also for social inequality as women continue to have lower access to higher paying employment opportunities. Using data from the Trends in Mathematics and Science Study of 2011, this paper finds that whilst grade 9 girls in the poorest 80% of South African schools experience no difference in domain specific performance, self-concept and motivation, girls in the wealthiest subset of schools are found to significantly underperform in both subjects, as well as possess lower self-concept and motivation, and higher anxiety. Teacher gender and education are shown to correlate with these results; specifically, female teachers with math backgrounds negatively influence girls' performances in wealthy schools. This is argued to be in keeping with stereotype threat theory (Steele, 2003) whereby women that are highly identified with math are subject to greater anxiety and concern over their performance. The relative difference in the performance of girls taught by a female versus a male teacher compared to the performance of boys is smaller when exposed to teachers with education training, suggesting that classroom methodology plays a role in the attainment of girls. Student fixed effects estimation reveals that the teacher characteristics mentioned above play important roles for moderating the relationship between student self-concept and performance of boys and girls.

[1] I gratefully acknowledge funding from the Growth and Economic Opportunities for Women (GrOW) initiative, a multi-funder partnership between the UK's Department for International Development, the Hewlett Foundation and the International Development Research Centre. I also

[2] Lecturer, Department of Economics, Stellenbosch University, Private Bag X1, 7602, Matieland, South Africa. debrashepherd@sun.ac.za

1. Introduction

Many high-paying, high-skilled occupations in developed countries are in the fields of science, technology, engineering and math (STEM), with the South African labour market being no exception. This is not surprising as science and technology have been the engines of profound social change and economic growth (Atkinson and Mayo, 2010). Consequently, they can play a significant role in the economic empowerment of women. However, women continue to be underrepresented in STEM fields. Gender equity in STEM education and jobs is needed not only because the education of women in STEM fields would give access to a larger number of highly-skilled (and productive) labour, thereby meeting the scientific and technological needs of an economy, but also because society can benefit from the insights and knowledge that women possess/ bring to the manner in which science and technology are used.

At the level of tertiary education in South Africa, less than 30% of students enrolled for undergraduate degrees in the fields of engineering and computer science in 2012 were women, whilst about half of science undergraduate enrolments were women; this is compared to 65% of all undergraduate students being women. Whilst there has been some improvement in the representation of women in STEM fields – for example, the proportion of female enrolments for engineering and computer science undergraduate degrees increased from 16.4% to 24.8% between 2007 and 2012, a growth rate of 50% – women continue to be underrepresented, particularly at masters and doctoral levels. Approximately 35% of all professional occupations over the period 2000 to 2014 have been in the fields of science and technology, with women holding approximately 45% of professional science and technology jobs (Statistics SA LF[3] 2000 – 2007; Statistics SA QLFS[4] 2008-2014). Removing nursing occupations from the numerator, approximately 25% to 30% of science and technology (non-nursing) jobs are held by women, indicating that approximately 35% to 40% of professional STEM women are in nursing. This gender gap in STEM occupations can contribute significantly to the overall gender wage gap as well as social inequality (Barres, 2006).

Oakes (1990) points towards three factors that are critical to attainment in STEM fields: *opportunities* to learn science and mathematics; *achievement* in these subjects; and the decision to pursue them (*choice*). Women can lose ground on all three of these factors during their education, such that the STEM gender gap apparent in university and the labour market can begin at the stages of basic education, even when girls are performing at the same level or better than boys in science and mathematics. Overall, gender differences in either mathematics or science achievement tend not to appear until high school levels (c.f. Lindberg,

[3] Labour Force Survey

[4] Quarterly Labour Force Survey

Hyde, Petersen & Linn, 2010), although recent research has shown the gap to emerge as early as the grades 1 and 2 (Levine et al, 2005; Rathbun et al, 2004). Educational foundations, which include social comparisons of performance with peers, can have important implications for teacher expectations, student motivation and future academic achievement (Farkas, 2003).

Research on gender differences in brain structure and inherent biological ability has been inconclusive in explaining differences in gender participation in STEM fields (Ceci, Williams & Barnett, 2009). Other research has pointed to the existence of negative stereotypes within society that affect the math and science performance of female students through undermining their confidence (Hill, Corbett & Rose, 2010; Steele and Aronson 1995). Whilst children may view boys and girls as equally able in mathematics, they view adult men as being better at maths than adult women (Steele, 2003); similar tendencies to stereotype the domain of mathematics as a male one have been shown to exist amongst parents (c.f. Frome & Eccles, 1998) and teachers (c.f. Helwig, Anderson & Tindal, 2001).

It is also important to take cognisance of the fact that gender inequalities do not operate within a vacuum, but interact with, for example, race and class. Whilst previous studies have shown that gender gaps in achievement interact with race and SES (e.g. Jencks & Phillips, 1998; Lee & Burkam, 2002), limited research (e.g. Fin & Ishak, 2012; Goni & Bello, 2016; Vyas & Choudhary, 2016) has examined how gender differences in self-concept and motivation differ by socioeconomic status; and this despite the increasing cultural diversity of the school-aged population. It is crucial that we understand how students of different genders and socio-economic backgrounds construct their academic identities, particularly in the context of gender and cultural norms, as well as socialization experiences. Although numerous studies have investigated the effect of various conventional schooling inputs on educational outcomes in South Africa, peer effects and (to a greater degree) school composition and social comparison have been investigated to a limited extent.

This study aims to analyse gender differences in self-concept, motivation and subsequent performance in mathematics and science amongst grade 9 students in South Africa. To this end, the Trends in Mathematics and Science studies (TIMSS) of 2011 is used to construct measures of self-concept and motivation that are then regressed onto TIMSS mathematics and science performance using classroom and student fixed effects estimation. Fixed effects regression models are estimated separately for the wealthiest and poorer subsets of the South African schooling system in order to identify whether or not competence beliefs and motivation/ attitudes play different roles in determining performance depending on the social context of the school environment, and whether a gender gap in STEM achievement exists after controlling for these factors. Finally, the analysis investigates the role of teacher gender and teacher training as a determinant of differences in gender outcomes in performance, as well as a moderator of the effect of self-concept on performance.

The paper proceeds with a review of the theory and existing literature on gender, self-belief (efficacy) and motivation in section 2, followed by a discussion of social comparison theory and stereotype threat in section 3. Sections 4 and 5 describe the data and the constructs to be used in the analysis and the methodological approach, respectively. Section 6 provides a discussion of the empirical results, and section 7 concludes.

2. Gender, self-beliefs and motivation

2.1 Self-(competence) beliefs

Self-beliefs such as self-concept and self-efficacy have received prominence in theories about the motivational sources of individual differences in performance, for example self-regulation theory (Carver & Schreier, 1981), social learning (cognitive) theory (Bandura, 1986) and self-determination theory (Deci & Ryan, 1985). The effect of self-concept on student achievement has received most attention (Valentine, DuBois & Cooper, 2004; Marsh & Craven, 2006). Although known to be distinct constructs, self-concept and self-efficacy are related both in their conception and in their effects on student achievement. Pivotal to Bandura's (1986) social cognitive theory is the concept of self-efficacy, which refers to "the beliefs in one's capabilities to organize and execute the courses of action to produce given attainments" (Bandura, 1997: 3). Central to this theory is the reciprocity between self-efficacy beliefs, classroom structures and social interactions with peers, and the mediating effects that these can have on a student's performance and motivation. For example, 20 years of self-efficacy research has shown that people with augmented self-efficacy beliefs tend to have higher academic achievement than those with lower self-efficacy beliefs (Jinks & Morgan, 1999; Pajares & Schunk, 2001; Usher & Pajares, 2008), as well as better engagement (Schunk & Mullen, 2012).

Although linked to judgments of one's capabilities in a given domain, the construct of self-concept represents a quite different view of self to self-efficacy. Whereas self-efficacy beliefs revolve around questions of "can" (confidence), self-concept beliefs reflect questions of "being" (self-worth) (Pajares & Schunk, 2001). Prior to the 1980s, over 90 percent of studies linking self-concept to achievement reported moderate to weak correlations (Byrne, 1984). However, this was because global rather than domain-specific measures of self-concept were being compared to achievement. When domain-specific (for example, academic) self-concept is compared with achievement in the same domain, the relationship is evidenced to be positive and significant (Marsh, 1993). Academic self-concept can furthermore be defined as "an evaluative self-perception that is formed through experience with, and interpretation of one's school environment" (Guay, et al., 2004: 53).[5]

Gender differences of mathematics self-concept have generally been consistent with traditional gender role expectations and stereotypes, with research showing higher mathematics self-concept scores for males (Eccles et al, 1993; Marsh & Yeung, 1997) and higher mathematics self-efficacy scores for males (Pajares & Miller, 1994; OECD, 2004). As early as school-entry age, children have been shown to make distinct judgments about their abilities in different domains (Eccles, Wigfield et al, 1993). These competency beliefs decline over the course of schooling, although the rate of change differs by gender (Jacobs et al, 2002) and ethnicity (Okeke, 2009). Higher self-perception of mathematics ability in boys has been found to be relatively independent of performance history (e.g., Frome & Eccles, 1998), achievement level, and ability (Holling & Preckel, 2005).

2.2 Achievement motivation

Self-Determination Theory (SDT; Deci and Ryan, 1985) laid down much of the foundation for contemporary research. Based on the different goals that give rise to action, SDT distinguishes between different types of motivation, the most basic distinction being between intrinsic (doing something because it is inherently interesting or enjoyable) and extrinsic (doing something because it leads to an independent outcome, such as good performance). A natural (intrinsic) motivational tendency is critical to cognitive and social development; however, intrinsic motivation does not only exist within individuals, but can exist in relation between individuals or in the activity itself (Skinner, 1953). Because most activities people do are not intrinsically motivated, such as completing non-intrinsically interesting tasks at school, intrinsic motivation does appear to become weaker as individuals move up grade levels.

SDT suggests that extrinsic motivation can vary according to its relative autonomy. For example, a child doing their math homework only out of fear of the parental sanctions that would result from not doing it is very differently extrinsically motivated to a child that believes that completing the same math homework is valuable for getting into university. Deci and Ryan (1985) describe within SDT a process of promoting the internalisation and integration of values and behavioural regulations that motivate students value and self-regulate educational activities. Whilst internalisation is the process of taking in a value (e.g. mathematics is necessary for getting into the university of your choice), integration is the process by which individuals transform this value/ regulation

[5] The term 'academic self-concept' can be characterised by two elements consistent with the Shavelson model: first, academic self-concept reflects *descriptive* (e.g., I like math) as well as *evaluative* (e.g. I am good at math) aspects of self-belief; and second, self-beliefs tend to focus on achievement in school rather than attitudes (Reyes, 1984).

into their own (*identification* and *integrated regulation*). Increased internalisation of values and commitment to these values leads to greater persistence and better engagement (Deci & Ryan, 2000). This can be contrasted to *amovitation* (lack of any intention to act resulting from not valuing an activity and not feeling competent to do an activity), *external regulation* (activities performed to satisfy an external demand) and *introjected regulation* (activites performed to avoid anxiety or to attain ego-enhancements). Studies concerning the different types of extrinsic motivation have shown that more autonomous extrinsic motivation is related to greater engagement, better performance (Miserandino, 1996), lower drop-out (Vallerand & Bissonnette, 1992), higher quality learning (Grolnick & Ryan, 1987) and greater psychological well-being (Sheldon & Kasser, 1995).

Studies conducted since the early 1990s have revealed no clear pattern of gender differences in students' achievement goal orientations and motivation. Where differences have been found, these are usually moderated by ability, race, and classroom context. Several studies in economics have examined the effect of having a female teacher on academic performance, especially in math, and the choice of a math and science major of female students in high school (c.f. Ehrenberg, Goldhaber, and Brewer, 1995; Dee, 2005, 2007; Winters et al., 2013, Antecol et al, 2014). These studies have often found that having a female teacher has positive or no effects on the test scores of girls. Antecol et al (2014), on the other hand, find that having a female teacher lowers the math scores of girls attending primary schools in disadvantaged neighbourhoods; this is eliminated with a stronger math background of the female teacher. The findings of Antecol et al (2014) are in agreement with research in the educational psychology literature that finds that primary school classrooms with more anxious female math teachers as well as a greater endorsement for stereotypes that "boys are goods at math" is related to lower math achievements of girls relative to boys. It is therefore important to understand the role that teacher gender plays as a mechanism linking lower self-concept amongst girls to lower achievement in STEM subjects.

3. Social comparison and stereotype threat

Socialisation and achievement experiences play a pivotal role in the development of gender differences in motivation. Social comparison is concerned with the processes involved in comparing ourselves with others, or thinking about the self in relation to others. This can include comparisons with stereotypes and hypothetical characters (Festinger, 1954; Wood, 1996; Webb-Williams, 2006). Festinger (1954) defined social comparison as a deliberate process of selecting social information to evaluate one's opinions and abilities, as well as to reduce uncertainty with regard to beliefs of self-worth. Social comparisons with peers can be markedly increased by teachers, parents and the students themselves, especially in relation to the domains of school achievement. As mentioned above, social cognitive theory hypothesizes that the cognitions of individuals regarding performance are influenced by social-contextual factors inter alia messages from teachers about task difficulty, perceived abilities of classroom peers, and information about the importance of learning (Bandura, 1986; Dweck & Leggett, 1988). Motivation and self-efficacy and -concept beliefs, therefore, emerge from the interaction between agents within the social contexts of the home, classroom and school (Bandura, 1986; Pajares, 1996; Schunk, 1984; Schunk & Miller, 2002).

Social interactions in schools can take many forms, having complex and varying effects on students' academic and self-efficacy belief. Until recently, the impact of social comparison on performance had not been thoroughly studied, which is surprising given that students tend to compare their own performance to that of their peers rather than to their past performances. Social comparisons are critical to the development of self-concept and self-efficacy beliefs; for example, the Big-Fish-Little-Pond-Effect describes how students form self-concept beliefs in part by comparing their own academic ability with the perceived abilities of other students in their peer group. Both self-concept and self-efficacy research agrees that social-comparative school practices can assist in destroying the self-beliefs of weaker performer, whilst social comparisons can be minimized through individualized classroom structures (Marsh, 1993).

Because gender differences are found so early in development, the home environment plays an important role in the shaping of competency beliefs and interests. At school, children will then have an opportunity to validate, refine, and enact their already learned gender beliefs and behaviour. According to the social cognitive model of Eccles, Adler, Futterman, Goff, Kaczala et al (1983), culture, parents and teachers contribute to shaping gender differences in competency, value beliefs and in several ways, including modelling sex-typed behaviour, communicating different expectations and goals for boys and girls and encouraging different activities and skills. The development of male advantaged in mathematics has also been shown to vary with socio-economic status (Levine et al, 2005), suggesting that gender segregation in STEM occupation may be related to gender dynamics in middle and upper class families.

Research has shown that cultural stereotypes (e.g. boys are better in math and science) influence parents' perceptions of their child's abilities, which in turn affects a child's perception of their own abilities (Parsons, Adler, et al, 1982; Jacobs & Eccles, 1992). Parental influence has also been shown to have enduring influences on achievement and attitudes through its impact a child's career interests and choices (Bleeker & Jacobs, 2004; Jacobs, Chhin & Bleeker,). At the level of the school, socialisation in the form of gender norms and roles may be even more rigid and polarised than what is found in broader society; for example, staffing patterns where men hold roles of leadership whilst women are found in nurturing/ caring roles.

Curriculum and teaching materials can also reinforce gender role lessons. Classroom interactions that follow gendered differentiated patterns (e.g. boys receiving more praise or being called on to answer more difficult questions) also serve to communicate different learning expectations for boys and girls (Brophy & Good, 1974). These gender-differentiated interaction patterns appear to be more pronounced in stereotypically male sex-typed subjects such as math and science (Jones & Dindia, 2004; Kahle & Meece, 1994), but can be moderated by classroom structures and environments (e.g. degree of competitiveness and teacher control). As high school teachers tend to use whole-class instruction and discussion, and boys are more likely to take an active role in these kinds of classroom settings, the emergence of gender differences in performance in favour of boys (or the narrowing of gender differences in performance favouring girls) may be related to a shift in teaching environments.

Value beliefs (the activities that students value) are also evidenced to follow gender and stereotype norms, and decline in relation to achievement values of time. Socialization processes that lead to internalization and acceptance of stereotypes responsible for group differences in competency and value beliefs can therefore have significant effects for engagement and performance (Eccles et al, 1983). However, it needs to be recognised that gender differences can vary across ethnic and socioeconomic groups; along with gendered expectations, black students must also contend with stereotypes of their intellectual inferiority and discrimination (Spencer et al, 1991; Steele, 1992). Black boys, in particular, experience more academic and discipline problems than their female counterparts; this can lead to gender differences in value and self-efficacy beliefs.

Finally, identity processes play a central role in the development of motivation. As shown in the seminal work of Steele, Spencer and Aronson, members of negatively stereotyped groups (for example, girls in mathematics) may worry that their weak academic achievement could confirm the negative stereotype about their group; the resultant stress leads to worse performance. Even if the individual does not believe the stereotype, the threat of being evaluated in terms of the stereotype could be felt nonetheless. If the psychological environment is able to render the stereotype/ s irrelevant, then it stands to reason that performance amongst these groups would improve. Steele (1997) proposed that individuals who are most affected by stereotype threat can be categorised as being 'highly identified' with the academic domain; for example, a woman who identifies as being good at mathematics. Therefore, stereotype threat is "felt most by people who care or who are invested in the domain where the stereotype threat applies" (Crocker, Major & Steele, 1998).

4. Data and constructs

4.1 Trends in Mathematics and Science Study

The data used were taken from the Trends in Mathematics and Science Study (TIMSS) for 2011. The focus of TIMSS is to assess the mathematics and science knowledge of students in the fourth and eighth grade, respectively (TIMSS International Report, 2011). For South Africa, the TIMSS 2011 survey was carried out in 285 schools, among 11 969 grade 9 students (HSRC, 2011). TIMSS was specifically chosen for the analysis of this paper, as the data is rich in contextual information regarding student attitudes towards school, their teachers and learning.

The sampling and assessment designs for TIMSS pose some complexities for analysis. Specifically, the assessment design is a balanced incomplete block design that is used to increase content-area coverage without simultaneously increasing the assessment time demanded of students. As a result, student achievement in mathematics and science is represented by five plausible values for each student. For purposes of this study, only the first plausible value is used, which is Rasch-scaled to an international mean of 500 and standard deviation of 100. Furthermore, sampling weighting is applied in all model estimation and clustering at classroom level is taken into account for the computation of standard errors.

As interest is specifically in determining the relationship between performance and self-concept and motivation whilst controlling for the cultural and gender diversity of the South African schooling system and classrooms, table 1 summarises descriptive statistics for the sample by school socioeconomic (SES) quintile, the gender of the student, gender composition of the classroom and the gender of the teacher. In line with existing research, large disparities between the average performances of students in the poorest 60 to 80 percent of schools (Q1to4) and the wealthiest 20 percent (Q5) are evident; students in the top quintile of school SES attained average scores close to two international standard deviations higher than students in the 1st quintile. The distribution of girls across schools and classrooms are fairly uniform across all quintiles, except in the case of single-sex schools where representation is largely concentrated within the top quintile of schools. In contrast, approximately a third of all grade 9 mathematics teachers in the top school SES quintile are male, whereas approximately 60 percent of teachers in the remaining school SES quintiles are male. The gender distribution amongst science teachers is fairly uniform across school SES quintiles.

4.2 Measures of academic self-concept and motivation

This paper relies on the use of two key constructs, namely academic self-concept and motivation. These constructs are generated from student responses to a number of survey questions as they appear in the TIMSS 2011 student questionnaire. Students responded on a Likert scale from 1 ("Agree a lot") to 4 ("Disagree a lot"). In order to generate these measures, polychoric factor analysis was used, which serves as an alternative approach to the analysis of discrete data when interest is in the computation of correlations between ordinal variables, such as in the case of Likert scale type responses. The factor weights assigned to each item and their distribution across the three measures are shown in table 2. As can be seen, similar factor weights are assigned to the same item across the domains of mathematics and science.

Table 1: descriptive statistics (mean) for performance and gender distribution of students and teachers, by school SES quintile

	School SES quintile					
	1	2	3	4	5	All
Mathematics score	309	339	347	359	485	368
Science score	261	312	324	350	504	331
Girls	46.6	50.4	48.1	49.2	51.8	49.2
Prop. boys in class	53.4	49.6	51.9	50.8	48.2	50.8
Prop. classes where boys >60%	26.8	18.4	16.1	17.8	22.6	20.4
Prop. classes where boys <40%	13.4	17.8	9.3	15.0	29.0	16.9
Prop. male only classrooms	0	0	3.5	0	9.5	3.2
Prop. female only classrooms	0	0	0	0	6.0	1.6
Male mathematics teacher	60.3	69.7	58.5	52.7	36.4	56.3
Male science teacher	43.2	45.4	44.9	54.9	40.9	45.4
Number of students	2 415	2 373	2 413	2 395	2 373	11 969
Number of classrooms	56	59	57	61	84	317
Number of schools	56	58	55	53	63	285

Note: mathematics and science scores are measured by the first plausible value (scaled to an international M = 500 and SD = 100). School SES quintile is calculated using the average SES of students in each school, where SES is computed using principal component analysis of 11 home possessions/ assets.

The sample (Cronbach's alpha) reliabilities of the self-concept and motivation composites for each subject are shown in table 3. There is some variation in the internal consistency of particularly the self-concept and amotivation constructs across school SES. Closer inspection of item-test and item-rest correlations within school SES quintiles reveals that items 8, 9 and 15 in table 2 do not, as expected, correlate negatively with the remaining self-concept items in quintile 1 schools. This suggests that students in the poorest schools are possibly receiving poor signals/ feedback regarding their performance such that they can effectively evaluate their own performance against that of their peers, or conversely, performance is generally very low in this context such that assessment of own performance in one subject against another and in comparison with peers is impractical. A correlation analysis of the three constructs (see bottom panel of table 3) indicates, as expected, positive correlations between self-concept and motivation and negative correlations with amotivation.

Table 2: Academic self-concept and motivation item distribution and weighting results from polychoric factor analysis

		Construct					
		Self-concept		Motivation		Amotivation/ anxiety	
	Item	Math	Science	Math	Science	Math	Science
1	I enjoy learning ______.			0.70	0.77		
2	I wish I did not have to study ______.					0.72	0.73
3	______ is boring.					0.70	0.73
4	I learn many interesting things in ______.			0.64	0.61		
5	I like ______.			0.74	0.82		
6	It is important to do well in ______.			0.66	0.75		
7	I usually do well in ______.	0.74	0.74				
8	______ is more difficult for me than for many of my classmates.	-0.38	-0.37				
9	______ is not one of my strengths.	-0.50	-0.43				
10	I learn things quickly in ______.	0.72	0.73				
11	______ makes me confused and nervous.					0.50	0.60
12	I am good at working out difficult ______ problems.	0.68	0.71				
13	My teacher thinks I can do well in ______ lessons with difficult materials.	0.57	0.67				
14	My teacher tells me I am good at ______.	0.72	0.71				
15	______ is harder for me than any other subject.	-0.46	-0.39				
16	I think learning ______ will help me in my daily life.			0.75	0.81		
17	I need ______ to learn other school subjects.			0.55	0.67		
18	I need to do well in ______ to get into the university of my choice.			0.73	0.83		
19	I need to do well in ______ to get the job I want.			0.75	0.83		
20	I would like a job that involves using ______.			0.64	0.81		

Table 3: Cronbach's alpha and factor correlation across constructs

	School SES quintile					
	1	2	3	4	5	All
Math self-concept	0.52	0.69	0.72	0.77	0.90	0.76
Science self-concept	0.62	0.71	0.74	0.75	0.89	0.75
Math motivation	0.82	0.81	0.78	0.78	0.84	0.81
Science motivation	0.84	0.84	0.85	0.86	0.91	0.87
Math amotivation/ anxiety	0.56	0.63	0.63	0.63	0.74	0.64
Science amotivation/ anxiety	0.61	0.69	0.69	0.68	0.77	0.69

	Self-concept	Motivation	Amotivation
Self-concept	1.00 (*1.00*)		
Motivation	0.54*** (*0.61****)	1.00 (*1.00*)	
Amotivation	-0.45*** (*-0.45****)	-0.46*** (*-0.38****)	1.00 (*1.00*)

Note: correlations for science constructs shown in parentheses. *** p = 0.01, ** p = 0.05, *, p = 0.10.

5. Methodological approach: classroom and within-student fixed effects estimation

In order to make inferences of causality, self-concept research should include a sufficiently large and diverse sample. Furthermore, domain-specific self-concept and achievement should be measured at least twice so that a reciprocal effects (non-recursive) model that allows for feedback between multiple endogenous variables can be estimated (Pajares & Schunk, 2001). Whilst the TIMSS data does satisfy the condition of a large and diverse sample, the cross-sectional nature of the data prevents the estimation of a non-recursive model. However, evidence does suggest that self-concept may play a stronger causal role at higher-grade levels (Skaalvik & Hagtvet, 1990). Therefore, this paper models the relationship between self-concept, motivation and academic performance as a recursive one, with the self-concept and motivation constructs assumed to be exogenous.

The following equation is estimated using ordinary least squares:

$$y_{ic,t} = \alpha + \beta F_{ic,t} + \delta SC_{ic,t} + \gamma M_{ic,t} + \rho AM_{ic,t} + \theta \boldsymbol{X}_{ic,t} + \vartheta_{c,t} + \varepsilon_{ic,t} \quad (1)$$

where y_{itc} is the test score for student i, in subject t and in classroom c. F is a dummy variable representing whether the student is a girl, and SC, M and AM are the self-concept, motivation and amotivation constructs, respectively. $\boldsymbol{X}_{itc}$ is a vector of student i's specific characteristics, including language spoken at home, exposure to English and home socio-economic status. ϑ_{tc} represent classroom fixed effects. Given the potential for error correlation across students in the same class, all standard errors are corrected to reflect classroom clustering.

There are two important identification issues that may cause bias in conventional ordinary least squares estimation. First, unobserved teacher and student traits that are correlated with student test scores may further bias the conventional OLS estimates. The classroom fixed effects are included such that unobserved differences (common treatments) across classrooms and schools are dealt with.[6] Furthermore, controlling for classroom specific fixed effects exploits for the natural variation in student self-concept and motivation and cohort compositions within classrooms in order to identify the relationship between these factors and performance in math and science. Second, students may not be assigned randomly across and within schools (c.f. Clotfelter, Ladd & Vikdor, 2006; Kane, Taylor, Tyler & Wooten, 2011). If better-performing students with higher self-concept and motivation select into, for example, quality schools or classrooms taught by better-educated/ higher-quality teachers, the coefficients of interest will be overstated. Common practice for dealing with this issue is to control for students' prior achievement (c.f. Hanushek & Rivkin, 2010).

A within-student, between-subject estimation procedure would be able to deal with non-random sorting and subject-invariant student and teacher/ classroom unobservables. However, any student and classroom characteristics that are invariant across subjects, such as student gender and the gender composition of the class, cannot be included in the model. Separation of the full grade 9 sample by school SES and a suitable set of student and home level characteristics as controls in the model may correct for non-random sorting of students across schools. Restricting the model further to single-class schools eliminates within school sorting.[7]

As there are two observations for each student (math and science), within-student variation in performance and self-concept and motivation can be exploited to identify δ, γ and ρ after accounting for subject-invariant unobservable student and home background traits. Following Clotfelter et al (2007), the preferred model now takes the following form:

$$(2) \qquad y_{ic,s} - y_{ic,m} = \alpha_s - \alpha_m + \delta(SC_{ic,s} - SC_{it,m}) + \gamma(M_{ic,s} - M_{ic,m}) + \rho\,(AM_{ic,s} - AM_{ic,m}) + \varepsilon_{ic}(s) - \varepsilon_{ic}(m)$$

where φ_i refers to a set of student specific fixed effects. As student prior achievements are not available, the assumption is made that the student fixed effect captures any overall ability or achievement level, and furthermore that ability is independent of subject. Finally, the effects of student gender, self-

[6] 80 percent of the schools surveyed only had one grade 9 classroom, with the remaining 20 percent of schools had predominantly 2 classrooms. Therefore, the classroom fixed effect is also a school effect.

[7] However, single-class schools are not randomly distributed across the school system.

concept, motivation and amotivation are assumed to be subject-invariant, for example, $\beta F_{ic,s} = \beta F_{ic,m}$. It can be shown that equation (2) is equivalent to:

$$(3) \qquad (y_{ic,t} - \bar{y}_i^*) = \delta(SC_{ic,t} - \overline{SC}_i^*) + \gamma(M_{ic,t} - \overline{M}_i^*) + \rho\,(AM_{ic,s} - \overline{AM}_i^*) + (\varepsilon_{i,t} - \bar{\varepsilon}_i^*) + (\varepsilon_{ic,t} - \varepsilon_{i,t})$$

where the variables with asterisks are student-specific means, $(\varepsilon_{i,t} - \bar{\varepsilon}_i^*)$ refers to a student-specific error term that varies across subject and $(\varepsilon_{ic,t} - \varepsilon_{i,t})$ refers to a subject-specific error term that varies with unobservable student (and classroom) characteristics. A student's achievement in subject t (in classroom c) is therefore measured relative to the average of their achievement based on both tests. Similarly, teacher and classroom characteristics will be measured relative to the average classroom and teacher characteristics of that student.

Model (3) will provide unbiased estimates of δ, γ and ρ if neither $(\varepsilon_{i,t} - \bar{\varepsilon}_i^*)$ nor $(\varepsilon_{ic,t} - \varepsilon_{i,t})$ are correlated with the demeaned variables. Taking $\varepsilon_{i,t}$ to represent the student's ability in subject t, the term $(\varepsilon_{i,t} - \bar{\varepsilon}_i^*)$ would equal zero if student ability does not differ across math and science, and therefore no statistical issues would arise. The second error term, $(\varepsilon_{ic,t} - \varepsilon_{i,t})$, accounts for the effects on student achievement of unobservable student and classroom characteristics, such as global self-esteem and learning strategies. This term will not bias the coefficients of interest if these characteristics are randomly distributed among students and classrooms. It cannot be proven conclusively that the above-mentioned conditions hold, and therefore the analysis conducted in this paper cannot be interpreted as completely free from bias.

One consequence of fixed effects estimation is that any student and classroom characteristics that are invariant across subjects cannot be included in the model. As student gender is fixed within a student, student fixed effects estimation prevents estimation of the relationship between gender and performance; however, it is possible to estimate and compare the results from separate models for male and female students.

6. Empirical Results

6.1 Distribution of achievement, self-concept and motivation

Means and cluster robust standard errors for girls' and boys' achievement (test scores), self-concept, motivation and amotivation/ anxiety in mathematics and science across school SES quintile are indicated in table 4 below. Whilst no significant gender difference in mathematics and science performance is observed for the full sample of schools, a significant difference in mathematics scores of 33 points (approximately 0.4 quintile 5 standard deviations) in favour of boys is observed for Q5 schools, and similarly a significant (at the 10 percent level) difference of 33 points for science; no significant gender difference is evidenced for the remaining Q1to4 schools.

With regards to self-concept, motivation and amotivation, again significant gender differences in favour of boys are found for Q5 schools. Therefore, the findings from the international literature only appear to hold for the wealthiest subset of the South African schooling system. In addition to the gender differences across school SES quintile, it is interesting to point out that, amongst all school quintile and gender groupings, girls attending Q5 schools have the lowest expected mathematics and science self-concept whilst their male counterparts have the highest. Girls attending Q5 schools are also the least motivated and most amotivated of the student groups.

Table 4: Mean domain specific achievement, self-concept, motivation and amotivation/ anxiety of grade 9 boys and girls, by school SES quintile

	School SES							
	Quintile 1-3		Quintile 4		Quintile 5		All	
	girls	boys	girls	boys	girls	boys	girls	boys
				Mathematics				
Test score	327.6	325.3	368.7	364.7	477.8	509.4**	354.0	352.1
	(3.10)	(2.88)	(5.53)	(5.48)	(6.93)	(10.23)	(4.03)	(4.60)
Self-concept	0.09	0.12	-0.20	0.03**	-0.39	0.09***	-0.03	0.10**
	(0.03)	(0.03)	(0.05)	(0.05)	(0.07)	(0.07)	(0.03)	(0.02)
Motivation	0.07	-0.04	0.06	0.20*	-0.45	-0.18***	-0.01	-0.01
	(0.03)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.03)	(0.03)
Amotivation	-0.05	-0.01	-0.03	-0.10	0.27	0.04**	0.00	-0.02
	(0.03)	(0.02)	(0.05)	(0.05)	(0.05)	(0.05)	(0.03)	(0.03)
				Science				
Test score	297.0	291.9	363.1	353.5	495.3	528.2*	333.8	328.0
	(4.80)	(4.27)	(7.73)	(6.95)	(7.56)	(10.18)	(5.73)	(6.04)
Self-concept	0.04	0.06	-0.06	0.09	-0.50	-0.08**	-0.05	0.05
	(0.03)	(0.03)	(0.08)	(0.05)	(0.09)	(0.07)	(0.03)	(0.03)
Motivation	0.08	0.10	-0.06	0.08	-0.62	-0.31***	-0.04	0.05
	(0.03)	(0.03)	(0.08)	(0.05)	(0.09)	(0.07)	(0.04)	(0.02)
Amotivation	0.03	0.10	-0.11	-0.08	0.10	-0.01	0.01	0.06
	(0.04)	(0.03)	(0.06)	(0.05)	(0.07)	(0.05)	(0.03)	(0.03)

Note: mathematics and science scores are measured by the first plausible value (scaled to an international M = 500 and SD = 100). School SES quintile is calculated using the average SES of students in each school, where SES is computed using principal component analysis of 11 home possessions/ assets. Mathematics and science self-concept is computed using polychoric factor analysis on 7 items, motivation is computed using polychoric factor analysis on 7 items, and amotivation/ anxiety is computed using polychoric factor analysis on 3 items. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. *** p = 0.01, ** p = 0.05, *, p = 0.10.

6.2 *Classroom fixed effects results*

Tables 5 and 6 summarise the model results for math and science performance of the full sample of students, respectively. Self-concept, motivation and amotivation have the expected (and statistically significant) relationships with performance. However, controlling for all, domain-specific motivation is negatively (and significantly) related to science performance. Furthermore, a significant gender difference in favour of boys of approximately 3 to 4 points is found for both subjects, although gender is not a significant moderator of the relationship between self-concept and performance. Controlling further for the gender composition of the classroom, the coefficient on student gender becomes larger (although turns insignificant for science). This indicates that the gender effect in favour of boys is related to gender compositions in the classroom; specifically, girls taught in classrooms that are either male or female dominated perform significantly lower than their male counterparts, controlling for self-concept and motivation. Where gender compositions are more equal, there is no significant gender difference in performance when comparing students with similar levels of self-concept and motivation.[8] However, it should be pointed out that larger gender performance gaps in gender-skewed classrooms are accompanied by higher average performance for both genders.

Tables 7 to 11 report results for the same model specification estimated for the samples of Q1to4 students and Q5 students separately. It is noteworthy that motivation and amotivation appear to have different relationships with performance (controlling for self-concept) in Q5 and Q1to4 schools. Whilst all three of the self-belief and attitude measures appear to be correlated in Q5 schools (that is, the effect of self-concept captures the effect of motivation), the role of self-concept and motivation on performance in the poorer schools appear to be distinct; a student who has a low or average self-concept in mathematics can still experience a positive performance effect through high motivation and/or interest in the subject. In science, as with the full sample, the relationship is negative. Closer analysis of the components of the motivation construct indicates that external regulation, which may induce feelings of anxiety, is negatively related to performance, whilst integrated regulation is positively related to performance.

One result that is common to both school samples is a stronger negative relationship between performance and amotivation/ anxiety for science than what is estimated for mathematics. Conversely, a significant gender gap is only estimated for Q5 schools, which is only eliminated in mathematics classrooms with more equal distributions of boys and girls. The coefficient on self-concept in mathematics in the sample of Q5 schools is estimated to be significantly larger than the coefficient on science self-concept in the same sample, as well as in both

[8] This result remains even when removing single-sex classrooms.

Table 5: Classroom fixed effects model of mathematics achievement, whole sample of schools

	Dependent variable = mathematics test score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	-4.41**				-3.46**	-3.42**	-8.87**
	(1.55)				(1.48)	(1.48)	(4.25)
Girl * self-concept						-0.79	-0.77
						(1.55)	(1.54)
Self-concept		18.44***			9.59***	9.22***	9.23***
		(0.93)			(1.01)	(1.19)	(1.20)
Motivation			14.25 ***		3.43***	3.40**	3.39***
			(0.82)		(0.93)	(0.92)	(0.92)
Amotivation/ anxiety				-21.14***	-15.92***	-15.91***	-15.86***
				(0.82)	(0.89)	(0.88)	(0.88)
Girl * class 40%-60% boys							8.44*
							(4.62)
Girl * class > 60% boys							0.54
							(5.87)
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Classroom fixed effets	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.037	0.027	0.030	0.090	0.058	0.076	0.053
Observations	11 821	10 293	10 439	10 365	9 401	9 401	9 401

Note: math score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Self-concept, motivation and amotivation/ anxiety are computed using polychoric factor analysis and z-scored. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. Student controls include whether the student has their own books at home, how often the student speaks the language of the test, socio-economic status of the household, mother's education and father's education. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Table 6: Classroom fixed effects model of science achievement, whole sample of schools

	Dependent variable = science test score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	-2.45				-3.76**	-3.84**	-6.28
	(2.02)				(1.76)	(1.76)	(3.85)
Girl * self concept						-2.99*	-2.87
						(1.55)	(1.56)
Self-concept		16.26***			6.23***	7.66***	7.60***
		(1.05)			(1.27)	(1.43)	(1.43)
Motivation			9.83***		-3.71***	-3.71***	-3.73***
			(1.01)		(1.21)	(1.20)	(1.20)
Amotivation/ anxiety				-29.36***	-27.86***	-27.88***	-27.87***
				(1.11)	(1.29)	(1.29)	(1.29)
Girl * class 40%-60% boys							4.09
							(4.42)
Girl * class > 60% boys							-0.76
							(6.04)
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Classroom fixed effets	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.159	0.109	0.104	0.173	0.145	0.143	0.140
Observations	11 821	10 268	10 486	10 524	9 577	9 577	9 577

Note: science score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Self-concept, motivation and amotivation/ anxiety are computed using polychoric factor analysis and z-scored. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. Student controls include whether the student has their own books at home, how often the student speaks the language of the test, socio-economic status of the household, mother's education and father's education. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Table 7: Classroom fixed effects model of mathematics achievement, quintile 5 schools

	Dependent variable = mathematics test score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	-16.03***				-6.28**	-6.10**	-12.55**
	(3.19)				(2.83)	(2.90)	(5.12)
Girl * self concept						0.90	1.01
						(1.98)	(1.97)
Self-concept		24.42***			23.26***	23.65***	23.77***
		(1.01)			(1.58)	(1.75)	(1.71)
Motivation			17.65***		-3.36	-3.37	-3.55
			(1.41)		(2.21)	(2.22)	(2.21)
Amotivation/ anxiety				-21.51***	-5.34***	-5.34***	-5.41***
				(1.36)	(1.86)	(1.86)	(1.84)
Girl * class 40%-60% boys							12.32**
							(6.11)
Girl * class > 60% boys							-0.62
							(8.14)
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.178	0.329	0.189	0.225	0.342	0.341	0.318
Observations	2 130	2 009	2 035	2 050	1 898	1 898	1 898

Note: math score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Self-concept, motivation and amotivation/ anxiety are computed using polychoric factor analysis and z-scored. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. Student controls include whether the student has their own books at home, how often the student speaks the language of the test, socio-economic status of the household, mother's education and father's education. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Table 8: Classroom fixed effects model of science achievement, quintile 5 schools

	Dependent variable = science test score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	-16.13***				-11.94***	-13.09***	-15.17**
	(3.89)				(3.69)	(3.76)	(6.83)
Girl * self concept						-3.69	-3.88
						(2.75)	(2.74)
Self-concept		18.74***			10.67***	8.84***	8.78***
		(1.44)			(1.86)	(2.46)	(2.48)
Motivation			12.16***		-0.94	-0.86	-0.88
			(1.57)		(2.26)	(2.26)	(2.26)
Amotivation				-20.76***	-13.80***	-13.76***	-13.71***
				(1.55)	(2.15)	(2.15)	(2.16)
Girl * class 40%-60% boys							5.23
							(8.18)
Girl * class > 60% boys							-3.45
							(9.47)
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.196	0.210	0.210	0.221	0.231	0.244	0.224
Observations	2 130	2 024	2 064	2 059	1 966	1 966	1 966

Note: science score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Self-concept, motivation and amotivation/ anxiety are computed using polychoric factor analysis and z-scored. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. Student controls include whether the student has their own books at home, how often the student speaks the language of the test, socio-economic status of the household, mother's education and father's education. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Table 9: Classroom fixed effects model of mathematics achievement, quintile 1 to 4 schools

	Dependent variable = mathematics test score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	-3.28*				-3.40**	-3.28**	-7.97
	(1.69)				(1.67)	(1.64)	(5.14)
Girl * self concept						-1.39	-1.34
						(1.88)	(1.88)
Self-concept		16.58***			7.28***	6.62***	6.63***
		(1.09)			(1.10)	(1.35)	(1.35)
Motivation			13.52***		4.71***	4.65***	4.67***
			(0.91)		(0.98)	(0.97)	(0.97)
Amotivation/ anxiety				-20.83***	-16.93***	-16.92***	-16.86***
				(0.94)	(0.94)	(0.94)	(0.94)
Girl * class 40%-60% boys							7.29
							(5.51)
Girl * class > 60% boys							-0.06
							(6.77)
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.035	0.053	0.069	0.135	0.132	0.132	0.134
Observations	9 691	8 284	8 404	8 315	7 503	7 503	7 503

Note: math score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Self-concept, motivation and amotivation/ anxiety are computed using polychoric factor analysis and z-scored. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. Student controls include whether the student has their own books at home, how often the student speaks the language of the test, socio-economic status of the household, mother's education and father's education. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Table 10: Classroom fixed effects model of science achievement, quintile 1 to 4 schools

	Dependent variable = mathematics test score						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Girl	-1.18				-2.76	-3.09	-4.84
	(2.20)				(1.90)	(1.90)	(4.25)
Girl * self concept						4.45**	4.37**
						(1.81)	(1.81)
Self-concept		15.49***			5.90***	8.03***	8.00***
		(1.23)			(1.46)	(1.63)	(1.63)
Motivation			9.00***		-3.63***	-3.61***	-3.63***
			(1.18)		(1.42)	(1.41)	(1.40)
Amotivation				-30.36***	-29.02***	-29.05***	-29.04***
				(1.23)	(1.39)	(1.39)	(1.39)
Girl * class 40%-60% boys							3.11
							(4.85)
Girl * class > 60% boys							-1.25
							(6.64)
Student controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Classroom fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.099	0.083	0.083	0.211	0.203	0.202	0.203
Observations	9 691	8 244	8 422	8 465	7 611	7 611	7 611

Note: science score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Self-concept, motivation and amotivation/ anxiety are computed using polychoric factor analysis and z-scored. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. Student controls include whether the student has their own books at home, how often the student speaks the language of the test, socio-economic status of the household, mother's education and father's education. *** p = 0.01, ** p = 0.05, *, p = 0.10.

academic domains of the Q1to4 school sample.

6.3 Teacher gender and training as potential mechanisms for the test performances of female students relative to male students

One environmental factor that might influence the relationship between the constructs of self-concept and motivation and performance, as well as gender and performance, is the gender of the teacher. As mentioned in the literature review, classroom interactions, which may or may not differ by teacher gender, can allow for gender-differentiated learning expectations and experiences. For example, a teacher's expectations based on stereotypical gender beliefs can become self-fulfilling prophecies (e.g. Cushman, 2010). The teaching styles of male and female teacher's can also conform to normative gender roles, with male teachers preferring compliant students and a teacher-centred instruction style, whilst female teachers may rely on more collaborative classroom environments (Brophy, 1985).

In order to formally determine the effect of female teachers on student achievement, the following regression is estimated for mathematics:[9]

$$\text{(4)} \quad y_{ic} = \alpha + \pi_1 \text{Girl}_{ic} + \pi_2 \text{FemaleTeacher}_{ic} + \pi_3 \text{Girl} * \text{FemaleTeacher}_{ic} + \theta \boldsymbol{X}_{ic} + \vartheta \boldsymbol{TC}_{ic} + \varepsilon_{ic}$$

where the interaction allows for teacher gender to differentially affect female and male students, the vector of student characteristics, $\boldsymbol{X}$, now includes the constructs for self-concept, motivation and amotivation, and $\varepsilon_{ic,t}$ is defined as previously. Note that this model does not include classroom fixed effects, but other teacher/ classroom controls, such as teacher age, teaching experience and job satisfaction,[10] are included in the vector $\boldsymbol{TC}$. Single-sex classrooms are excluded and standard errors are clustered at the classroom level.

The results for the Q5 and Q1toQ4 school samples are presented in table 12. Focusing first on the results for Q5, relative to male students taught by female teachers, female students score worse on the mathematics test, approximately 9% of a sample standard deviation. This difference is, however, insignificant at conventional levels. Similarly, girls taught by male teachers are not estimated to perform significantly better or worse than boys taught by a

[9] This part of the analysis is limited to mathematics given that the negative coefficient on girl was largely confined to this domain.

[10] Teacher age is a categorical variable that takes on 6 values: younger than 25 years; 25-29 years; 30-39 years; 40-49 years; 50-59 years; and 60 years or older. Teaching experience is a categorical variable that takes on 4 values: 0-5 years; 6-10 years; 1 1-15 years; and 16 years or more. Job sentiment is a continuous variable (z-scored) that is generated using polychoric factor analysis of 6 survey items relating to teacher sentiments about their profession e.g. "I had more enthusiasm when I began teaching than I have now." A higher value of this variable implies a more positive sentiment.

male teacher in Q5 schools. The difference-in-differences coefficient (π_3), which gives the relative difference in mathematics test scores between female students with female versus male teachers and male students with female versus male teachers, is large (-11.1, which is approximately 15% of a sample standard deviation) and statistically significant at the 10% level. These results are very similar to those of Antecol et al (2014), and taken together suggest that female grade 9 mathematics teachers relative to male teachers in Q5 schools adversely influence the mathematics outcomes of female students but not male students. With regards to Q1toQ4 schools, the relative difference in the math performance of female and male students taught by a female teacher is very large and significant (-38 points, or 59% of a standard deviation). The relative difference in scores of female and male students taught by a male teacher is similarly large and significant (-30 points), but, as indicated by the difference-in-differences coefficient, is significantly smaller than the relative difference under a female teacher. Therefore, in Q1toQ4 schools, both male and female teachers adversely influence the mathematics outcomes of girls, but more so in the case of the latter.

Table 12: Classroom fixed effects model of math achievement with gender interactions

	(1) Q1toQ4	(2) Q5	(3) Q1toQ4	(4) Q5
Girl (π_1)	-30.09***	4.52	-4.93	-8.79*
	(8.82)	(11.33)	(5.06)	(4.62)
Female Teacher (π_2)	16.37***	6.56		
	(4.67)	(9.20)		
Girl*Female Teacher (π_3)	-7.96**	-11.10*	-5.66	-2.17
	(3.97)	(6.64)	(3.63)	(4.89)
Girl * class 40%-60% boys	-18.85***	29.83***	7.05	10.94**
	(6.20)	(10.11)	(5.22)	(4.66)
Girl * class > 60% boys	-28.16***	8.44	-1.20	-2.07
	(8.21)	(10.67)	(6.54)	(6.71)
$\pi_1 + \pi_3$	-38.06***	-6.58	-10.60**	-10.95***
	(8.69)	(10.51)	(5.24)	(3.93)
Student controls	Yes	Yes	Yes	Yes
Teacher controls	Yes	Yes	No	No
Classroom fixed effects	No	No	Yes	Yes
R-squared	0.239	0.472	0.131	0.335
Observations	6 861	1 422	7 492	1 571

Note: math score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. *** p = 0.01, ** p = 0.05, *, p = 0.10.

The model results in columns (3) and (4) control for classroom fixed effects in order to ensure that the results are not driven by unobservable differences between male and female teachers.[11] The relative difference between female and male students taught by a female teacher in Q5 schools is now larger negative (-10.95 points, or 14.4% of a standard deviation) and statistically significant. Girls taught by male teachers are also estimated to perform 8.79 points (approximately 12% of a standard deviation) worse relative to boys taught by male teachers, implying that, after controlling for unobservable teacher characteristics, the mathematics performance gap in favour of boys under a female teacher does not appear to be relatively larger than that found under a male teacher in wealthier school settings. In poorer school settings, accounting for teacher unobservables dramatically reduces the male-female performance gap under a female teacher to a coefficient that is now also 10.6 points (16% of a standard deviation), as in the case of Q5 schools. Girls taught by male teachers are estimated to perform 4.93 points worse, although this coefficient is not statistically significant. Therefore, the mathematics performance gap in favour of boys under a female teacher is relatively larger than that found under a male teacher in poorer school settings.

It is conceivable that the teaching styles of teachers may differ depending on the strength of their backgrounds in the subject of teaching, as too could the teaching styles between male and female teachers, and this may affect the academic achievements of male and female students differently. Antecol et al (2014) posit three ways in which a weaker test score for girls in, for example, mathematics could arise. First, teachers without strong subject backgrounds may adopt a more mechanical teaching style that is fine for male students but not for female students. Alternatively, female teachers may teach STEM subjects just as well as male teachers, but female students may respond more positively (negatively) to a male (female) teacher at certain ages. Both of these reasons are judged as unlikely by Antecol et al (2014), and rather they find suggestive evidence in support of math anxiety, which may be reduced the stronger the math background of the female teacher and/ or the less held stereotypical beliefs are in the classroom. Recent evidence from the educational psychology literature, in particular that of Beilock, Gunderson, Ramirez and Levine (2010), finds that higher math anxiety in female primary school teachers hurts the math performances of female students but not that of male students. Specifically, the more anxious female teachers are in mathematics classes, the more likely female students are to endorse the stereotype that boys are better at math than girls. This seems to suggest that stereotype threat does not only affect the girl students in a classroom, but female teachers as well. Interestingly, the gender achievement gap is closed when students are taught by female teachers in classrooms with more equal distributions of student gender. This could be

[11] In this case, the regression will take the form $y_{ic} = \alpha + \pi_1 \mathrm{Girl}_{ic} + \pi_3 \mathrm{Girl} * \mathrm{FemaleTeacher}_{ic} + \theta \boldsymbol{X}_{ic} + \vartheta_{c,t} + \varepsilon_{ic}$, where $\vartheta_{c,t}$ are classroom fixed effects.

suggestive of an environment in which previously held stereotypes can be challenged.

The main area of study of the teacher is observed from the TIMSS contextual surveys, with some interesting differences in the distribution of post-secondary study-area across subjects and teacher gender emerging. Table 13 indicates the distribution of teacher study area by school SES quintile, gender and subject. It is evident that the majority of male and female teachers in Q1toQ4 schools tend to be specialists in their subjects, with math teachers showing slightly more generalisation with training in both mathematics and sciences. This may correlate to differences in the quality of teaching across mathematics and science classrooms. Overall, there do not appear to be any important gender differences in the distribution of focal areas across grade 9 science and math teachers in Q1toQ4 schools. Interesting differences, however, are found for Q5 schools. Male mathematics and science teachers in Q5 schools tend to be more generalist than their female counterparts; whilst just more than half of male mathematics teachers and close to 40 percent of male science teachers studied both mathematics and science, just less than a quarter of female mathematics and science teachers have similar focal areas. A fifth of female mathematics teachers in Q5 schools do not denote their post-secondary focal area as either mathematics or science, or even general education.

Following Antecol et al (2014), a teacher's training background is controlled for using the following regression specification for mathematics:[12]

$$(5) \qquad y_{ic} = \alpha + \pi_1 \text{Girl}_{ic} + \pi_3 \text{Girl} * \text{FemaleTeacher}_{ic} + \pi_6 \text{Girl} * \text{Training}_{ic} \\ + \text{Girl} * \text{FemaleTeacher} * \text{Training}_{ic} + \theta \boldsymbol{X}_{ic} + \vartheta_{ic} + \varepsilon_{ic}$$

where all variables are defined as previously, single-sex classrooms are again excluded, and classroom fixed effects are added.[13] The variable 'training' is coded to take three values: diploma or degree in mathematics; diploma or degree in math and science; and a diploma or degree with neither a mathematics nor science focus. The distribution of these post-secondary qualifications across school SES, teacher gender and subject are indicated in the second panel of table 13. As not all teachers studied a degree or diploma with a focus on education, equation (5) is estimated separately for two groups of teachers, those with post-secondary studies that included a focus on education and those that did not. The

[12] The same model is estimated for science and shown in table A3 of the appendix. However, the results are not as easily interpretable, especially for the Q5 sample, given the substantial differences in the distributions of male and female teachers across tertiary training categories (see table 13).

[13] The coefficient subscripts have been intentionally chosen to compare with equation (4).

Table 13: Distribution of teacher main area of post-secondary study over school SES quintile, teacher gender and subject

	Q1toQ4				Q5			
	Male		Female		Male		Female	
Post-secondary main area of study	Math	Science	Math	Science	Math	Science	Math	Science
Mathematics, physics, science and biology	5.6	8.0	1.7	3.7	18.2	32.6	0.9	5.6
Math, no science	30.6	8.3	41.1	7.9	28.1	0.0	42.4	0.0
Math and science	47.0	18.3	39.0	21.2	37.5	4.6	20.6	16.9
Science, no math	4.6	48.0	5.0	42.6	7.1	57.6	6.9	57.0
Math education, no general education	3.1	1.7	2.0	0.1	0.0	0.0	1.8	0
Science education, no general education	0	3.4	0	2.9	0.0	0.0	0	6.8
General education and math education	0	0.1	2.4	0.8	0.0	0.0	3.3	0
General education and science education	0	1.7	0	1.4	0.0	0.0	0	0
General education only	2.3	3.4	2.4	7.4	0.0	0.0	4.1	5.2
Other/ not provided	6.8	7.1	6.4	12.0	9.1	5.2	20.0	8.5
Education diploma/ degree with subject specific (math or science) focus	8.1	22.7	21.5	17.3	15.3	26.6	21.4	33.3
Education diploma/ degree with math and science	28.8	17.9	14.0	9.8	14.6	39.8	18.1	1.6
Education diploma/ degree with other subject focus	12.0	13.9	13.3	20.0	2.3	0.0	20.8	18.3
Diploma/ degree with subject specific (math or science) focus	22.6	25.1	19.6	25.3	5.3	26.1	17.5	20.8
Diploma/ degree with math and science	23.9	8.5	26.7	15.0	45.6	1.8	10.1	18.4
Other diploma/ degree	4.5	11.1	4.8	12.6	16.9	5.7	12.1	7.6

Note: Sample weights are applied.

results are presented in table 14. Three coefficients of interest are included in the table. $(\pi_1 + \pi_3)$ indicates the relative difference in the scores of girls and boys taught by female teachers that studied a diploma or diploma without a math focus; $(\pi_3 + \pi_{7,1})$ indicates the relative difference in the math scores ofgirls taught by female versus male teachers and boys taught by female versus male teachers with mathematics as a main area of post-secondary study; similarly, $(\pi_3 + \pi_{7,2})$ and π_3 indicate the relative performances of girls and boys when taught by a female versus male teacher whose post-secondary studies included mathematics and science and no mathematics, respectively.

Within poorer (Q1toQ4) school contexts, girls relative to boys taught by a female teacher with a teaching qualification with no math focus perform worse (-19.18 or 29.4% of a sample standard deviation), and this coefficient is significant at the 5% level. The gender performance gap for students taught by a similarly qualified male teacher is large negative (-16.81 or 25.8% of a sample standard deviation) and statistically significant. The coefficients $(\pi_3 + \pi_{7,1})$ and $(\pi_3 + \pi_{7,2})$ are not estimated to be significantly different from zero, indicating that, in Q1toQ4 schools, girls taught by female teachers with education qualifications without a focus on math do not have significantly lower math achievement outcomes than girls taught by similarly educated female teachers with math backgrounds. However, it can be noted that the coefficient $\pi_{7,1}$ is large positive (11.29 or 17.3% of a sample standard deviation). Even though the result is not significant at conventional levels, girls see an improvement in their relative performance when taught by a female teacher with a mathematics background compared to being taught by a female teacher with no mathematics background.

When attention is turned to teachers without education qualifications, the relative difference in scores between girls with female versus male teachers and boys with female versus male teachers is large and negative when teachers report having studied both math and science at tertiary level. When taught by a teacher reporting no or only a focus on mathematics, the relative performance difference is not significantly different; the test score of girls remains at approximately 20-24% of a sample standard deviation lower than boys when taught by a female teacher. However, when taught by teachers with a broader study focus (math and science), the relative gender gap is doubled. It is worth noting that the coefficient $\pi_{7,1}$ is large and positive, indicating that there is a benefit for girls being taught by female teachers with tertiary qualifications with a focus on mathematics. The small sample size of the group of female and male teachers in Q1toQ4 schools that studied a diploma or degree that had no mathematics focus may be generating this result, especially if there are significant differences in the subject/s that formed the focus of the qualification across gender.

In the case of wealthier school contexts, a dissimilar result is found. No statistically significant difference in the relative performances of female and male students in Q5 schools is estimated when the gender and relative mathematics background of the teachers are changed. It is interesting to note, however, that the relative performance gap between girls and boys taught by a female versus a male teacher reported to have majored in both math and education is 18.3% of a standard deviation larger (not statistically significant) than the relative performance gap when comparing the performance of girls and boys taught by female versus male teachers with education training but no mathematics focus. Finally, when comparing Q5 teachers that did not report education as a major area of study, the relative performance gap between girls and boys when taught by a female teacher versus a male teacher increases with the math background of the teacher. As with the Q1toQ4 sample, these results may be driven by differences in the distribution of male and female teachers across the difference domains of post-secondary study.

An alternative interpretation for the result that female teachers versus male teachers adversely influence the mathematics performance of girls but not that of boys is that male teachers relative to female teachers positively influence the performance of girls. Re-estimations of model (5) for separate samples of male and female teachers are shown in tables A1 and A2 of the appendix. The adverse effect of female teachers on the mathematics performances of girls persists even when male teachers are excluded from the analysis. No positive influence of male teachers on girls is suggested from the estimates, except in the case of male teachers in Q5 schools that studied an education degree where girls are estimated to perform significantly better than boys. In combination with the findings of table 14, this suggests a higher relative performance of girls in Q5 schools taught by teachers with university qualifications in education, and proposes that differences in the socialisation of girls and boys may interact with the styles and/or modes of teaching that may be utilised by teachers with different training rather than mathematics backgrounds. In the case of Q1toQ4 schools, the relative performance of girls in comparison to boys is improved when taught by a female teacher with qualifications in education, although this positive influence is cancelled out when the teacher has a university degree.

These findings are different to those of Antecol et al (2014) in that they do not provide, at least direct, evidence in support of the math anxiety hypothesis. However, where suggestive evidence may exist is in the comparison of the relative performance of girls assigned to female versus male teachers and boys assigned to female versus male teachers when teachers do not report having an education qualification. When assigned to a female teacher with higher exposure to mathematics at post-secondary levels, the math performances of girls in relation to that of boys worsen when compared with assignment to male teachers with (assumedly) similar backgrounds in mathematics.

Table 14: Classroom fixed effects model of math achievement controlling for teacher education, by school SES

	(1)	(2)	(3)	(4)
	Education included in main area/ s of study		Education not included in main area/ s of study	
	Q1toQ4	Q5	Q1toQ4	Q5
Girl (π_1)	-16.81**	-11.58	16.99**	-7.63
	(7.39)	(7.31)	(6.77)	(10.29)
Girl * Female Teacher (π_3)	-2.36	4.67	-15.88	-5.03
	(8.80)	(6.53)	(13.69)	(11.39)
Girl * Studied math ($\pi_{6,1}$)	-6.02	4.41	-17.83**	25.61***
	(7.97)	(9.65)	(7.36)	(7.93)
Girl * Female Teacher * Studied math ($\pi_{7,1}$)	11.29	-17.98	20.44	-14.73
	(12.16)	(11.97)	(14.83)	(12.60)
Girl * Studied math & science ($\pi_{6,2}$)	0.44	-1.87	-15.88**	1.35
	(7.05)	(8.12)	(6.95)	(10.49)
Girl*Female Teacher * Studied math & science ($\pi_{7,2}$)	2.98	-7.75	-0.01	-12.89
	(10.81)	(9.86)	(14.44)	(16.46)
Girl * class 40%-60% boys	17.63***	21.87***	-0.05	-12.50*
	(6.33)	(4.87)	(5.68)	(6.86)
Girl * class > 60% boys	18.68**	6.81	-12.29*	3.83
	(7.66)	(8.46)	(6.69)	(8.29)
$\pi_1 + \pi_3$	-19.18**	-6.91	1.11	-12.67*
	(9.72)	(5.33)	(13.13)	(7.12)
$\pi_3 + \pi_{7,1}$	8.92	-13.32	4.56	-19.77***
	(8.60)	(9.23)	(6.03)	(5.38)
$\pi_3 + \pi_{7,2}$	0.62	-3.09	-15.88***	-17.92*
	(5.57)	(8.31)	(5.31)	(9.52)
Student controls	Yes	Yes	Yes	Yes
Teacher controls	No	No	No	No
Classroom fixed effects	Yes	Yes	Yes	Yes
R-squared	0.133	0.325	0.143	0.378
Observations	4 310	963	3 182	608

Note: math score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. *** p = 0.01, ** p = 0.05, *, p = 0.10.

One explanation for this might be found in stereotype threat theory (Steele, 1997), whereby women that identify strongly with a particular domain, in this case mathematics, may experience more anxiety or concern when they have the potential to confirm a negative stereotype about their social group, such as “women are worse at math than men”. This implies that the ‘threat’ may be stronger for highly identified women. The threat associated with domain identification may operate differently across the different school SES contexts. In relatively poorer school contexts where performance is dramatically weaker, a woman who has obtained a university degree (academically identified) may experience greater threat as the expectations of performance, and hence anxiety, are much greater.

The level of (stereotype) threat could, amongst other things, be influenced by the relative distributions of boys and girls in classrooms, as well as the gender distribution of other mathematics and science teachers in their school. From table 14 we see that girls perform relatively better in gender equal environments, when the teacher has an education qualification. Therefore, teaching styles and classroom interactions that generate more equity in the performances of boys and girls are likely to arise from particular education training, as well as classrooms in which representation is more equal. TIMSS unfortunately does not provide information on the gender distribution of math and science teachers in each school, making it difficult to assume that the sample of teachers (or the teacher selected from each school) is necessarily random. For example, in a country such as South Africa where the school-exit (matriculation) exam determines a student's access to and placement in university, some schools may decide to allocate higher quality math teachers, which the person/ s making the hiring-decision might assume to be men, at grades 10 and above.

6.4 Student fixed effect model results

In order to determine whether or not the gender and qualification of teachers plays a role on performance through the self-concept of a student, a within-student between-subject model is estimated. This allows for bias driven by non-random sorting of students into and within schools to be eliminated. As there are two observations for each student (math and science), within-student variation in performance and self-concept and motivation can be exploited to identify δ, γ and ρ after accounting for subject-invariant unobservable student and home background traits. Following Clotfelter and Ladd (2007), the model now takes the following form:

$$(6) \qquad y_{ic,s} - y_{ic,m} = \alpha_s - \alpha_m + \delta(SC_{ic,s} - SC_{it,m}) + \gamma(M_{ic,s} - M_{ic,m}) \\ + \rho\,(AM_{ic,s} - AM_{ic,m}) + \varepsilon_{ic}(s) - \varepsilon_{ic}(m)$$

where φ_i refers to a set of student specific fixed effects. As student prior achievements are not available, the assumption is made that the student fixed effect captures any overall ability or achievement level, and furthermore that ability is independent of subject. Finally, the effects of self-concept, motivation and amotivation are assumed to be subject-invariant, for example, $\beta F_{ic,s} = \beta F_{ic,m}$.

Equation (6) is equivalent to:

$$(7) \qquad \left(y_{ic,t} - \bar{y}_i^*\right) = \delta\left(SC_{ic,t} - \overline{SC}_i^*\right) + \gamma(M_{ic,t} - \bar{M}_i^*) + \rho\,(AM_{ic,s} - \overline{AM}_i^*) + (\varepsilon_{i,t} - \bar{\varepsilon}_i^*) + (\varepsilon_{ic,t} - \varepsilon_{i,t})$$

where the variables with asterisks are student-specific means, $(\varepsilon_{i,t} - \bar{\varepsilon}_i^*)$ refers to a student-specific error term that varies across subject and $(\varepsilon_{ic,t} - \varepsilon_{i,t})$ refers to a subject-specific error term that varies with unobservable student (and classroom) characteristics. A student's achievement in subject t (in classroom c) is therefore measured relative to the average of their achievement based on both tests. Similarly, teacher and classroom characteristics will be measured relative to the average classroom and teacher characteristics of that student.

Model (7) will provide unbiased estimates of δ, γ and ρ if neither $(\varepsilon_{i,t} - \bar{\varepsilon}_i^*)$ nor $(\varepsilon_{ic,t} - \varepsilon_{i,t})$ are correlated with the demeaned variables. Taking $\varepsilon_{i,t}$ to represent the student's ability in subject t, the term $(\varepsilon_{i,t} - \bar{\varepsilon}_i^*)$ would equal zero if student ability does not differ across math and science, and therefore no statistical issues would arise. The second error term, $(\varepsilon_{ic,t} - \varepsilon_{i,t})$, accounts for the effects on student achievement of unobservable student and classroom characteristics, such as global self-esteem and learning strategies. This term will not bias the coefficients of interest if these characteristics are randomly distributed among students and classrooms. It cannot be proven conclusively that the above-mentioned conditions hold, and therefore the analysis cannot be interpreted as completely free from bias.

As gender is fixed within a student and across subjects, equation (7) is estimated separately for boys and girls. Furthermore, interactions between self-concept and the gender and training of the teacher, specifically whether the teacher studied an education qualification, are included as follows:

$$(8) \qquad \left(y_{ic,t} - \bar{y}_i^*\right) = \delta_1\left(SC_{ic,t} - \overline{SC}_i^*\right) + \delta_2\left(SC_{ic,t} - \overline{SC}_i^*\right) * \mathrm{StudiedEducation}_{ic,t} + \delta_3\left(SC_{ic,t} - \overline{SC}_i^*\right) * \mathrm{FemaleTeacher}_{ic,t} * \mathrm{StudiedEducation}_{ic,t} + \theta\left(\boldsymbol{TC}_{ic,t} - \overline{\boldsymbol{TC}}_i^*\right) + \gamma(M_{ic,t} - \bar{M}_i^*) + \rho\,(AM_{ic,t} - \overline{AM}_i^*) + (\varepsilon_{ic,t} - \bar{\varepsilon}_i^*) + (\varepsilon_{ic,t} - \varepsilon_{i,t})$$

where SC, M and AM are as previously defined, and $\boldsymbol{TC}$, a vector of teacher characteristics, includes teacher age, teaching experience, job sentiment and an indicator of whether they studied at university.

The estimation results of model (8) are shown in table 15. Focusing first on the results for the Q5 sample, the coefficients on the subject of focus in the post-secondary studies of teachers agree with those found in table 14 and table A2 of the appendix; controlling for learner subject-invariant characteristics, subject focus is not found to have a significant effect on the performance of students, although a more narrow focus on the subject of teaching as opposed to a broad STEM focus is positively related to performance. The significantly higher performance (at the 10% level) of boys taught by female teachers suggests that the negative influence of female teachers on girls may be driven by a positive influence of female teachers have on boys' performance. It is possible that this coefficient arises from the fact that significantly fewer male math and science teachers in quintile 5 schools have post-secondary education that is not in the subject of teaching (with no indication of the relative quality of these qualifications across male and female teachers). As these teachers are not specialised in the subject that is being taught, the teaching style adopted by, in this case, predominantly women teachers, may be more mechanical in nature, which might be less harmful to boys than girls.

A positive large (approximately 12-17% of a sample standard deviation) and significant relationship is estimated between student performance and teacher training that did not include education. This is accompanied by a large positive and statistically significant influence of university teacher training for girls (13.96 points, or 17% of a sample standard deviation). This result is an important one, as it provides additional evidence that the gender gap in performance (or relative performance of girls) is smaller (higher) when taught by teachers with university training.

When the above is combined with education training as a moderator of self-concept, the result is even more powerful. The coefficients at the bottom of table 15, namely δ_1, $\delta_1 + \delta_2$, $\delta_1 + \delta_{3,1}$ and $\delta_1 + \delta_2 + \delta_{3,2}$ represent the estimated relationship between self-concept and test scores for students taught by male teachers trained in (at a minimum) education, males teachers not trained in education, female teachers trained in (at a minimum) education and female teachers not trained in education, respectively. The effect of self-concept on test scores is positive and statistically significant (at the 5% and 10% level) for girls taught by teachers whose training included a focus on education; for each 1 standard deviation increase in self-concept, test scores of girls improve by 6.5 to 8.5 percent of a sample standard deviation. The self-concept effect for girls when taught by teachers without education training is not statistically significantly different from zero. For boys, the only similar positive self-concept effect is estimated when taught by female teachers with education training (8.8% of a sample standard deviation). A large positive, but statistically insignificant, self-concept effect is also estimated when boys are taught by male teachers not trained in education. It is relevant to point out that the effects of self-concept are substantially smaller than those estimated under a classroom fixed effects model, suggesting a positive relationship between student and home background unobservables and self-concept.

Table 15: Learner fixed effects model of achievement

	(1)		(2)		(3)		(4)	
	Q1toQ4				Q5			
	Girl		Boy		Girl		Boy	
Female teacher	1.12	(3.52)	-5.96*	(3.34)	5.56	(5.47)	14.32*	(7.38)
Training did not include math or science	8.30	(5.16)	6.24	(5.48)	7.74	(6.97)	9.92	(7.69)
Training focused on subject of teaching	17.81***	(4.45)	14.79***	(5.00)	-3.34	(5.42)	-10.85	(8.07)
Teacher did not do education training	-5.46	(4.47)	-1.70	(3.69)	10.85*	(6.17)	15.42*	(8.58)
Self-concept (δ_1)	15.17***	(2.03)	12.68***	(2.19)	7.03**	(3.13)	2.29	(3.89)
Self-concept * Teacher not trained in education (δ_2)	1.83	(2.90)	2.65	(2.90)	-4.23	(3.43)	3.26	(5.40)
Self-concept * Teacher trained in education * Female teacher ($\delta_{3,1}$)	-1.97	(2.82)	0.57	(2.82)	-1.72	(3.62)	5.00	(4.63)
Self-concept * Teacher not trained in education * Female teacher ($\delta_{3,2}$)	-1.36	(2.58)	-2.07	(3.18)	-0.32	(3.35)	-8.49*	(4.31)
Motivation	1.82	(1.81)	4.59***	(1.64)	-0.74	(2.23)	1.78	(3.20)
Amotivation	-5.85***	(1.46)	-6.96***	(1.51)	-8.18***	(2.57)	-9.26***	(2.63)
Teacher job sentiment	4.33**	(4.68)	1.51	(1.78)	-5.13*	(2.70)	-5.02	(3.60)
Teacher has a university degree	4.23	(11.02)	3.24	(3.82)	13.96**	(6.67)	-5.05	(9.73)
δ_1	15.17***	(2.03)	12.68***	(2.19)	7.03**	(3.13)	2.29	(3.89)
$\delta_1 + \delta_2$	17.00***	(2.57)	15.33***	(2.52)	2.80	(2.29)	5.55	(4.10)
$\delta_1 + \delta_{3,1}$	13.21***	(2.61)	13.25***	(2.42)	5.32*	(3.05)	7.29*	(3.72)
$\delta_1 + \delta_2 + \delta_{3,2}$	15.65***	(2.22)	13.26***	(2.46)	2.49	(3.15)	-2.94	(3.97)
Constant	328.97	(13.40)	341.47	(8.36)	463.32	(14.68)	493.89	(15.57)
School fixed effects	Yes		Yes		Yes		Yes	
Learner fixed effects	Yes		Yes		Yes		Yes	
Teacher controls	Yes		Yes		Yes		Yes	
R-squared	0.053		0.064		0.081		0.053	
Observations	6 675		6 894		1 701		1 477	

Note: test scores are measured by the first plausible value (scaled to an international M = 500 and SD = 100). Single-sex classrooms are excluded. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Turning the focus to students in Q1toQ4 schools, both boys and girls do significantly better (19% to 23% of a sample standard deviation) when taught by a teacher that is more broadly trained in both math and science as opposed to neither or math alone. No significant effect of university and education training is estimated. The coefficient on self-concept for girls and boys is very large positive (13-15 points or 16-19% of a standard deviation) and highly significant; this coefficient is not altered much when interacted with the training and gender of the teacher. It is clear that reported self-concept of students in Q1toQ4 schools is negatively correlated with unobserved student and home background characteristics. This might suggest that a relatively small proportion of (most likely) higher ability students in poorer school settings are able to correctly gauge their academic efficacy against that of their peers. The feedback mechanisms, such as frequent and reliable assessments, may be of such a poor quality that students are unable to properly rate their abilities within the classroom. Therefore, the large significant coefficient on self-concept estimated in table 15 indicates that, accounting for student ability, a strong self-concept is related to augmented test performance, albeit it from a low base.

7. Concluding remarks and policy recommendations

Despite recent improvements, the under-representation of women in tertiary enrolments in STEM fields of study, and subsequently professional STEM careers, continues to be an issue faced by the South African economy. To the knowledge of the author, this is the first paper that has attempted to understand the role of self-concept on observed achievement patterns in mathematics and science in South Africa. The analysis of this paper and interpretation of the results borrows considerably from the social psychology and educational psychology literatures, specifically the role of stereotypes in formulating the self-efficacy and value beliefs of children. The TIMSS 2011 dataset was used, as a multitude of questions related to student attitudes towards mathematics and science were integrated into the contextual surveys, allowing for a number of constructs to be developed through factor analysis.

In summary, this paper finds significant and large gaps in domain specific self-concept, motivation and anxiety for girls and boys, as well as in mathematics and science performance, for the wealthiest subset of schools (Q5); no significant differences are found for the poorer school wealth (SES) quintiles. Controlling for student observable characteristics and the domain specific constructs in classroom fixed effects models of mathematics and science achievement, a performance gap of 13-15 points (14-19% of a sample standard deviation) in favour of boys is estimated for the sample of Q5 schools, and only eliminated in classrooms where the distribution of boys and girls is more equal. The finding that, conditional on the same level self-concept and motivation, a substantial gender performance gap remains is concerning. The relationship between the achievement outcomes of girls and boys and teacher gender is examined through the inclusion of interactions between student and teacher gender in the classroom fixed effects model. Controlling for teacher and classroom unobservables, a significant gender gap in math test scores of approximately 11 points is estimated for students taught by female teachers. A sizeable gap is also estimated for male teachers, indicating that teacher gender alone does not explain the performance gap.

Following Antecol et al (2014), teacher training and education in combination with teacher gender is explored. It is conceivable that the teaching styles, and hence effectiveness, of teachers may contribute to differential performances amongst boys and girls, particularly in the manner in which classroom interactions and engagement interplay with different socialisation experiences of girls and boys, as well as the stereotype beliefs of both the students and the teacher. A distinction is made between the subject background of the teacher and the type of tertiary qualification, as South African high-school teachers, particularly in the STEM fields, may have completed their education training either in combination with or after their mathematics and/ or science training, or may have transitioned into teaching after more specialist tertiary education.[14]

The results show that the relative performance gap between female and male students is reduced when taught by teachers whose tertiary training included a focus on education. This is suggestive of differences in teaching methods that may be less harmful to girls adopted by teachers with education backgrounds; specialist, non-educationalist training may lend itself to, for example, more mechanical styles of teaching that boys find less harmful or more beneficial. In the case of poorer schools, some evidence of the math-anxiety hypothesis posited by Antecol et al (2014) is suggested by the results; that is, the negative effect of having a female teacher on the math scores of girls is reduced for students taught by female teachers whose training included a focus on mathematics.[15]

In Q5 schools, however, quite the opposite result is found; the relative difference in mathematics test scores between girls with female versus male teachers and boys with female versus male teachers is larger the more exposure the teacher has had to mathematics. Whilst this seems to be in contradiction to the math-anxiety hypothesis, this paper argues that this result relates well to stereotype threat theory, whereby strong domain identification of an individual

[14] This section of the analysis focused only on mathematics given dissimilar distributions in the training of male and female science teachers in Q5 schools, which complicates the ease of interpretability of the model results.

[15] The results are less clear with respect to Q1toQ4 teachers without an education background, and are possibly related to systematic differences in the quality of training across male and female teachers.

holding membership with a group facing a negative stereotype is likely to result in greater threat, and hence anxiety.

As a final exploration of the mechanisms for female student math and science test scores, a student fixed effects models that exploits within student variation in performance, self-concept and motivation across the two subjects is estimated. The relationship between self-concept and performance is allowed to vary with teacher gender and education qualification. In poorer school contexts, the relationship between self-concept and performance does not appear to be influenced by these teacher observables. However, given that the coefficient on self-concept is dramatically increased once controlling for student unobservables, this suggests that the process of social comparison on self-efficacy, from which self-concept derives, may be hampered by the larger dysfunctional nature of schooling within the poorest part of the South African education system. If feedback on performance is invalid or unreliable, students will struggle to accurately evaluate their performance and define their self-efficacy. It therefore comes as no surprise that, on average, better performing students with higher ability or better equipped to evaluate their self-efficacy.

Of more interest and relevance to the primary question at hand are the results estimated for the sample of Q5 schools. As would be expected, and unlike what was found for Q1toQ4 schools, student unobservable characteristics (or at least, subject invariant ones) are positively correlated with self-concept and motivation. Furthermore, controlling for these student unobservables and teacher observable characteristics, girls' test scores are positively significantly influenced by teacher education, in terms of both the level and type. What is clearly suggested by the results of table 15 is that teacher characteristics influence the relationship between student self-concept and performance. Specifically, girls experience a stronger self-concept to performance relationship when taught by a teacher with a focus on education training than not. Additionally, the self-concept effect is larger when girls are taught by male teachers with education backgrounds, which might suggest that girls, believing in the stereotype that men are better than women at maths and science, are likely to undervalue their own efficacy when taught by a female teacher. An alternative argument might be that male teachers are able to instil more discipline in their classrooms, especially at an age where posturing and competition amongst boys could occur, which can assist in creating an environment that is beneficial to girls' learning experiences. Whilst this might work to the favour of girls, it appears to hurt boys; similarly, boys taught by women with education training experience a significant self-concept to performance relationship.

This latter finding (although based on a relatively small, albeit better functioning, sample of South African high schools), in combination with the findings from the classroom fixed effect models, provides suggestive evidence that student endorsement of gender stereotypes as well as stereotype threat experienced by female teachers may be contributing to poorer performance, particularly in math, among female students. This finding is in agreement with the recent findings of Antecol et al (2014) in the economics literature, and stresses the importance of future work in this area as well as advocates for the use of path-analysis modelling that can reveal not only the moderating role, but also the mediating role of teacher characteristics on girls' attainment in STEM subjects.

However, it needs to be pointed out that, given the differences in the gender and training-background distribution of teachers across subject domain, it is difficult to argue that the estimated coefficients are free from bias driven by teacher unobservables, both subject-variant and subject-invariant. The teacher-gender dynamics observed here would need to be tested against other data, preferably the TIMSS grade 4 and grade 8 2015 datasets. Modelling similar relationships in the grade 4 data would serve to understand whether or not indications of differential self-concept linked to performance across boys and girls emerge at younger ages. This could explain at which stage of the school experience the academic course of students, particularly of girls, is shaped. Pooling the TIMSS 2011 and 2015 datasets would help to determine whether or not the results presented here are sample driven. A further mediating factor that was not directly explored by this paper is the SES of the home and parental involvement. The stereotype threat observed within the wealthier part of the school system, and hence amongst middle class families, may be due to differences in the manner in which math related "cultural resources" are shared amongst boys and girls (c.f. Muller, 1998).

The findings of this paper and the proposed future work present a number of policy implications. We need to better understand at which stage of schooling gender differences in performance emerge; knowing that there is a gender difference in the performance and self-concepts of grade 9 boys and girls in math and science might be met with a desire to better target and inform the subject choices of girls in high school, but this alone is not likely to solve the problem. Consideration needs to be given to the social factors that contribute to gender differences during earlier stages of schooling, as well as perpetuate and extend them as students move through the education system, need to be developed. As argued by Penner and Paret (2008: 251), "strategies for reducing gaps… need to be [placed]… within larger structures of inequality... it is important to focus on the ways in which processes of academic achievement are embedded in and shaped by the hierarchical structures of difference and inequality".

Classroom practices and methodologies that reflect and reinforce structures of gender inequality need to be identified, and replaced with ones that enhance gender parity. For example, whilst the learning styles of female students tend to emphasise mastery (pursuing work in the hope of understanding the material) over performance (focus on one's marks) in task completion, the reverse has been shown for male students (Kenney-Benson et al, 2006). Which of these are emphasised or encouraged in a classroom setting (or by the school system in general) may play an important role in the relative performance of girls and boys. Similarly, gender differences in socialisation that contribute to differences in activity level and temperament could affect teachers' subjective perceptions of students, and therefore their performances (Bennet et al, 1993). Finally, complimentary to the above mentioned could be policies targeted at creating and raising the gender awareness of teachers' and parents' in order to counteract stereotypes about female math inferiority could also be emphasised.

Appendix

Table A1: Determinants of math achievement for the Q1toQ4 sample, by teacher gender

	(1)	(2)	(3)	(4)
	Female		Male	
Girl (π_1)	-33.92*	-18.86*	-25.64**	2.29
	(18.49)	(9.56)	(10.27)	(6.58)
Studied math	1.54		5.98	
	(8.77)		(6.58)	
Studied math and science	-1.06		2.71	
	(8.68)		(6.30)	
Girl* Studied math ($\pi_{6,1}$)	4.02	1.62	-12.49***	-10.68**
	(6.51)	(6.61)	(4.66)	(4.79)
Girl* Studied math and science	-9.82	-8.62	-4.56	-4.92
($\pi_{7,1}$)	(6.70)	(6.72)	(4.73)	(4.79)
Studied education	4.07		0.25	
	(8.73)		(6.37)	
Girl*Studied education	12.74**	18.95***	-2.13	-3.84
	(5.99)	(5.36)	(5.39)	(5.14)
Studied a degree	-3.47		6.01	
	(10.96)		(8.58)	
Girl * studied a degree	-12.42**	5.98	12.62**	12.86**
	(6.07)	(7.65)	(4.96)	(5.41)
Studied a degree * studied	-6.74		-4.30	
education	(12.32)		(8.81)	
Girl * studied a degree * studied	5.07	-22.75**	-3.47	-8.56
education	(10.69)	(9.05)	(7.69)	(7.07)
Girl * class 40%-60% boys	26.90	16.50**	1.69	-3.97
	(21.49)	(7.37)	(6.90)	(5.11)
Girl * class > 60% boys	23.14	-0.86	21.27***	-8.16
	(14.92)	(9.89)	(7.67)	(6.08)
Student controls	Yes	Yes	Yes	Yes
Teacher controls	Yes	No	Yes	No
Classroom fixed effects	No	Yes	No	Yes
R-squared	0.273	0.120	0.243	0.143
Observations	2 713	2 713	4 148	4 204

Note: math score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Student controls include whether the student has their own books at home, how often the student speaks the language of the test and the socio-economic status of the household. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Table A2: Determinants of math achievement for the Q5 sample, by teacher gender

	(1)	(2)	(3)	(4)
	Female		Male	
Girl (π_1)	-28.15	-27.04**	-22.13	4.08
	(22.38)	(10.99)	(13.42)	(11.28)
Studied math	3.67		43.28***	
	(12.47)		(8.53)	
Studied math and science	23.95		43.73***	
	(16.09)		(9.89)	
Girl* Studied math ($\pi_{6,1}$)	-10.65	-2.43	-.4.59	-4.08
	(9.33)	(7.21)	(10.18)	(10.36)
Girl* Studied math and science	-11.50	-8.53	-3.12	-3.94
($\pi_{7,1}$)	(8.70)	(6.93)	(7.38)	(7.16)
Studied education	-53.27		-26.99**	
	(34.55)		(10.83)	
Girl*Studied education	30.97**	24.02*	20.00	20.01*
	(12.22)	(12.41)	(13.14)	(10.05)
Studied a degree	-13.41		-12.34	
	(39.56)		(15.46)	
Girl * studied a degree	1.67	18.42	7.92	7.11
	(8.74)	(11.60)	(13.37)	(10.84)
Studied a degree * studied	47.49		35.19**	
education	(39.61)		(15.23)	
Girl * studied a degree * studied	30.88	-17.43	18.10	19.08
education	(18.47)	(14.04)	(16.46)	(15.40)
Girl * class 40%-60% boys	-39.34	9.88	70.05***	-1.18
	(28.71)	(5.93)	(10.73)	(4.75)
Girl * class > 60% boys	-22.57	-6.95	-4.55	-9.96
	(17.19)	(8.73)	(12.83)	(9.97)
Student controls	Yes	Yes	Yes	Yes
Teacher controls	Yes	No	Yes	No
Classroom fixed effects	No	Yes	No	Yes
R-squared	0.493	0.360	0.653	0.230
Observations	953	953	469	496

Note: math score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Student controls include whether the student has their own books at home, how often the student speaks the language of the test and the socio-economic status of the household. Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. *** p = 0.01, ** p = 0.05, *, p = 0.10.

Table A3: Classroom fixed effects model of science achievement controlling for teacher education, by school SES

	(1)	(2)	(3)	(4)
	Education included in main area/ s of study		Education not included in main area/ s of study	
	Q1toQ4	Q5	Q1toQ4	Q5
Girl (π_1)	-1.39	-0.83	-14.61	62.45***
	(9.44)	(20.55)	(10.37)	(13.04)
Girl * Female Teacher (π_3)	4.35	-29.51	15.47	-44.46***
	(8.89)	(24.51)	(11.08)	(13.93)
Girl * Studied science ($\pi_{6,1}$)	3.43	-21.37	8.12	-56.10***
	(7.97)	(24.07)	(9.33)	(13.66)
Girl * Female Teacher * Studied	0.15	37.34	-15.51	29.13*
science ($\pi_{7,1}$)	(11.61)	(25.49)	(13.28)	(16.96)
Girl * Studied math & science ($\pi_{6,2}$)	17.29*	-19.48	14.07	-36.90***
	(9.27)	(22.34)	(12.01)	(10.29)
Girl*Female Teacher * Studied	-3.88	40.52*	-0.01	25.99
math & science ($\pi_{7,2}$)	(14.53)	(23.49)	(14.44)	(21.72)
Girl * class 40%-60% boys	-5.08	9.68	2.98	-26.49***
	(7.73)	(10.76)	(8.21)	(7.67)
Girl * class > 60% boys	-12.05	-4.48	-8.69	-27.42*
	(8.29)	(10.24)	(9.71)	(14.70)
$\pi_1 + \pi_3$	2.96	-30.34**	0.86	17.99*
	(8.78)	(11.51)	(10.18)	(9.36)
$\pi_3 + \pi_{7,1}$	4.50	7.83	-0.04	-15.33
	(7.98)	(9.73)	(8.20)	(9.85)
$\pi_3 + \pi_{7,2}$	0.47	11.01	1.91	-18.48
	(11.33)	(9.61)	(10.78)	(19.43)
Student controls	Yes	Yes	Yes	Yes
Teacher controls	No	No	No	No
Classroom fixed effects	Yes	Yes	Yes	Yes
R-squared	0.145	0.110	0.178	0.142
Observations	3 574	946	3 428	700

Note: science score is measured by the first plausible value (scaled to an international M = 500 and SD = 100). Cluster (classroom) robust standard errors are shown in parentheses. Sample weighting is taken into account. *** p = 0.01, ** p = 0.05, *, p = 0.10.

References

Antecol, H., Eren, O. and Ozbeklik, S., 2014. The effect of teacher gender on student achievement in primary school. *Journal of Labor Economics*, 33(1): 63-89.

Aronson, J. and Steele, C.M. 2005. Stereotypes and the Fragility of Academic Competence, Motivation and Self-Concept. In A.J. Eliot & C.S. Dweck (Eds.), *Handbook of Competence and Motivation* (pp. 436 – 456). New York: Guilford Press.

Atkinson, R., and Mayo, M. 2010. *Refueling the US Innovation Economy: Fresh Approaches to Science, Technology, Engineering and Mathematics (STEM) Education*. Executive Summary. Information Technology and Innovation Foundation.

Bandura, A. 1977. Self-efficacy: Toward a Unifying Theory of Behavioural Change. *Psychological Review*, 84(2): 191.

Bandura, A. 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall.

Barres, B. A. 2006. Does Gender Matter? *Nature*, 442(7099): 133-136.

Beilock, S.L., Gunderson, E.A., Ramirez, G. and Levine, S.C.. 2010. *Female Teachers' Math Anxiety Affects Girls' Math Achievement,* Proceedings of the National Academy of Sciences, USA, 107(5): 1060-1063.

Bennett, R.E., Gottesman, R.L., Rock, D.A. and Cerullo, F. 1993. Influence of behaviour perceptions and gender on teachers' judgements of students' academic skill. *Journal of Educational Psychology*, 85: 347-356.

Bleeker, M.M. and Jacobs, J.E. 2004. Achievement in Math and Science: Do Mothers' Beliefs Matter 12 Years Later? *Journal of Educational Psychology*, 96(1): 97.

Bong, M. and Skaalvik, E.M. 2003. Academic Self-Concept and Self-Efficacy: How Different are They Really? *Educational Psychology Review*, 15: 1 -40.

Brophy, J.E. and Good, T.L. 1974. *Teacher-student relationships: Causes and consequences*. Holt, Rinehart & Winston.

Brophy, J.E. (1985). *Teacher socialization as a mechanism for developing student motivation to learn*. (Research series).

Byrne, B.M. 1984. The general/academic self-concept nomological network: A review of construct validation research. *Review of educational research*, 54(3): 427-456.

Byrne, B.M. 2006. *Structural Equation Modelling with EQS: Basic Concepts, Applications, and Programming* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Carver, C.S. and Scheier, M.F. 1981. *Attention and Self-regulation: A Control Theory Approach to Human Behaviour*. Berlin: Springer-Verlag.

Ceci, S. J., Williams, W. M., and Barnett, S. M. 2009. Women's Underrepresentation in Science: Sociocultural and Biological Considerations. *Psychological Bulletin*, 135(2): 218.

Clotfelter, C.T., Ladd, H.F. and Vigdor, J.L. 2007. Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26(6): 673-682.

Crocker, J., Major, B. and Steele, C. M. 1998. Social stigma. In D. T. Gilbert, S. T. Fiske, and G. Lindzey (Eds.), *Handbook of Social Psychology* (4th ed., pp. 504-553). Boston: McGraw-Hill.

Cushman, P. 2010. Male primary school teachers: Helping or hindering a move to gender equality? *Teaching and Teacher Education*, 26: 1211–1218.

Deci, E. and Ryan, R., 1985. *Intrinsic Motivation and Self-Determination in Human Behaviour.* New York: Plenum.

Dee, T.S. 2004. Teachers, race and student achievement in a randomized experiment. *Review of Economics and Statistics,* 86(1): 195– 210.

Dee, T.S. 2005. A teacher like me: Does race, ethnicity, or gender matter? *American Economic Review* 95(2): 158–65.

Dee, T.S. 2007. Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3):528–54.

Dweck, C.S. and Leggett, E.L.1988. A social-cognitive approach to motivation and personality. *Psychological review*, 95(2): 256.

Eccles, J.S., Midgely, C. and Adler, T. 1984. Grade-Related Changes in the School Environment: Effects on Achievement Motivation. In J. Nicholls (Ed), *Advances in Motivation and Achievement: The Development of Achievement Motivation* (Vol. 3, pp.283 – 331). Greenwich, CT: JAI Press.

Eccles, J., Adler, T.F., Futterman, R., Goff, S.B., Kaczala, C.M., Meece, J.L. and Midgley, C. 1985. Self-perceptions, task perceptions, socializing influences, and the decision to enroll in mathematics. *Women and mathematics: Balancing the equation*: 95-121.

Ehrenberg, R.G., Goldhaber, D.D. and Brewer, D.J. 1995. Do teachers' race, gender and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *Industrial and Labor Relations Review* 48(3):547–61

Farkas, G. 2003. Racial disparities and discrimination in education. *Teachers College Record*, 105: 1119-1146.

Festinger, L. 1954. A theory of social comparison processes. *Human relations*, 7(2): 117-140.

Fin, L.S. and Ishak, Z. 2012. A Priori Model of Students' Academic Achievement: The Effect of Gender as Moderator. *Procedia-Social and Behavioral Sciences*, 65: 1092-1100.

Fiske, S.T. 1992. Thinking is for Doing: Portraits of Social Cognition from Daguerreo-Type to Laserphoto. *Journal of Personality and Social Psychology*, 63: 877 – 889.

Frome, P.M. and Eccles, J.S. 1998. Parents' Influence on Children's Achievement-Related Perceptions. *Journal of personality and social psychology,* 74(2): 435.

Goni, U. and Bello, S. 2016. Parental socio-economic status, self-concept and gender differences on students' academic performance in Borno State colleges of education: Implications for counselling. *Journal of Education and Practice*, 7(14): 21-27.

Grolnick, W.S. and Ryan, R.M.,1987. Autonomy in children's learning: An experimental and individual difference investigation. *Journal of personality and social psychology*, 52(5): 890.

Guay, F., Larose, S. and Boivin, M. 2004. Academic self-concept and educational attainment level: A ten-year longitudinal study. *Self and identity,* 3(1): 53-68.

Hanushek, E.A. and Rivkin, S.G. 2010. Generalizations about using value-added measures of teacher quality. *The American Economic Review*, 100(2): 267-271.

Helwig, R., Anderson, L. and Tindal, G. 2001. Influence of elementary student gender on teachers' perceptions of mathematics achievement. *Journal of Educational Research*, 95: 93-102.

Hill, C., Corbett, C., and St Rose, A. 2010. Why so few? Women in Science, Technology, Engineering, and Mathematics. American Association of University Women, Washington, DC.

Holling, H. and Preckel, F. 2005. Self-estimates of intelligence––methodological approaches and gender differences. *Personality and Individual Differences*, 38(3): 503-517.

Jacobs, J.E., Lanza, S., Osgood, D.W., Eccles, J.S. and Wigfield, A. 2002. Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child development*, 73(2): 509-527.

Jacobs, J.E. and Eccles, J.S. 1992. The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of personality and social psychology*, 63(6): 932.

Jacobs, J.E., Chhin, C.S. and Bleeker, M.M. 2006. Enduring Links: Parents' Expectations and Their Young Adult Children's Gender-Typed Occupational Choices. *Educational Research and Evaluation*, 12(4): 395-407

Jenks, C., and Phillips, M. 1998. *The black-white test score gap*. Brookings Institution Press: Washington, DC.

Jinks, J. and Morgan, V.,1999. Children's perceived academic self-efficacy: An inventory scale. *The Clearing House*, 72(4): 224-230.

Jones, S.M. and Dindia, K. 2004. A meta-analytic perspective on sex equity in the classroom. *Review of educational research*, 74(4): 443-471.

Kahle, J.B., Meece, J. and Scantlebury, K. 2000. Urban African-American middle school science students: Does standards-based teaching make a difference?. *Journal of Research in Science Teaching*, 37(9): 1019-1041.

Kane, T.J., Taylor, E.S., Tyler, J.H. and Wooten, A.L. 2011. Identifying effective classroom practices using student achievement data. *Journal of human Resources*, 46(3): 587-613.

Kenney-Benson, G.A., Pomerantz, E.M., Ryan, A.M. and Patrick, H. 2006. Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology*, 42: 11-26.

Levine, S.C., Vasilyeva, M., Lourenco, S.F., Newcombe, N.S. and Huttenlocher, J. 2005. Socioeconomic status modifies the sex difference in spatial skill. *Psychological Science*, 16: 841-845.

Lindberg, S.M., Hyde, J.S., Petersen, J.L. and Linn, M.C. 2010. New trends in gender and mathematics performance: A meta-analysis. *Psychol Bull*, 136(6): 1123-1135.

Marsh, H.W. 1993. The multidimensional structure of academic self-concept: Invariance over gender and age. *American Educational Research Journal*, 30(4): 841-860.

Marsh, H. W., Hau, K. T., Artelt, C., Baumert, J., and Peschar, J. L. 2006. OECD's Brief Self-report Measure of Educational Psychology's Most Useful Affective Constructs: Cross-cultural, Psychometric Comparisons Across 25 Countries. *International Journal of Testing*, 6(4): 311-360.

Marsh, H.W. and Craven, R.G. 2006. Reciprocal Effects of Self-Concept and Performance From a Multidimensional Perspective: Beyond Seductive Pleasure and Unidimensional Perspectives. *Perspective Psychology*, 1(2): 133-163.

Marsh, H. W., and O'Mara, A. 2008. Reciprocal Effects Between Academic Self-concept, Self-esteem, Achievement, and Attainment Over Seven Adolescent Years: Unidimensional and Multidimensional Perspectives of Self-concept. *Personality and Social Psychology Bulletin*, 34(4): 542-552.

Marsh, H.W. and Yeung, A.S. 1997. Academic Self-Concept and Achievement. *American Educational Research Journal*, 34(4): 691-720.

Marsh, H.W. 1996. Positive and Negative Global Self-Esteem: A Substantively Meaningful Distinction or Artifactors? *Journal of Personality and Social Psychology,* 70: 810 – 819.

Marsh, H.W. and Craven, R.G. 2006. Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on psychological science*, 1(2): 133-163.

Meece, J.L. and Kurtz-Costes, B.E. 2001. Introduction: The Schooling of Ethnic Minority Children and Youth. *Educational Psychologist*, 36(1): 1-7.

Miserandino, M. 1996. Children who do well in school: Individual differences in perceived competence and autonomy in above-average children. *Journal of educational psychology*, 88(2): 203.

Muller, C. 1998. Gender differences in parental involvement and adolescents' mathematics achievement. *Sociology of Education*, 71: 336-356.

Oakes, J. 1990. Chapter 3: Opportunities, achievement, and choice: women and minority students in science and mathematics. *Review of Research in Education*, 16(1): 153-222.

Okeke, N.A. 2009. Race stereotypes, academic self-concept and racial centrality in African American adolescents (Doctoral dissertation, The University of North Carolina at Chapel Hill).

Pajares, F., and Miller, M. D. 1994. Role of Self-efficacy and Self-concept Beliefs in Mathematical Problem Solving: A Path Analysis. *Journal of Educational Psychology*, 86(2): 193.

Pajares, F. and Schunk, D.H. 2001. Self-Beliefs and School Success: Self-Efficacy, Self-Concept, and School Achievement. Chapter in R. Riding & S. Rayner (Eds), *Perception* (pp.239 – 266). London: Ablex Publishing.

Pajares, F. and Urdan, T. 2006. *Adolescence and Education: Vol. 5. Self-Efficacy Beliefs of Adolescents*. Greenwich, CT: Information Age.

Parsons, J.E., Adler, T.F. and Kaczala, C.M. 1982. Socialization of achievement attitudes and beliefs: Parental influences. *Child development*: 310-321.

Penner, A.M. & Paret, M. 2008. Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research,* 37: 239–253,

Reyes, L.H. 1984. Affective variables and mathematics education. *The Elementary School Journal*, 84(5): 558-581.

Schunk, D.H. and Mullen, C.A. 2012. Self-efficacy as an engaged learner. In *Handbook of research on student engagement* (pp. 219-235). Springer US.

Sheldon, K.M. and Kasser, T. 1995. Coherence and congruence: two aspects of personality integration. *Journal of personality and social psychology*, 68(3): 531.

Skaalvik, E.M. and Hagtvet, K.A. 1990. Academic achievement and self-concept: An analysis of causal predominance in a developmental perspective. *Journal of Personality and Social Psychology*, 58(2): 292.

Skinner, B.F. 1953. *Science and human behavior*. Simon and Schuster.

Slee, R., Weiner, G. and Tomlinson, S. 1998. *School Effectiveness for Whom? Challenges to the School Effectiveness and the School Improvement Movements.* Londer, Falmer Press.

Spencer, S.J., Fein, S., Wolfe, C.T., Fong, C. and Duinn, M.A. 1998. Automatic Activation of Stereotypes: The Role of Self-Image Threat. *Personality and Social Psychology Bulletin*, 24(11): 1139-1152.

Steele, C.M. 1992. Race and the Schooling of Black Americans. *The Atlantic Monthly*, 269(4): 68-78.

Steele, C.M. 1999. Thin Ice: "Stereotype Threat" and Black College Students. *Atlantic Monthly*, 284(2): 44 – 54.

Steele, C. M., and Aronson, J. 1995. Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology*, 69(5): 797.

Steele, C. 2003. *Stereotype threat and African-American student achievement* (pp. 109-130). na.

Thrupp, M. and Lupton, R. 2006. Taking School Contexts More Seriously: The Social Justice Challenge. British Journal of Educational Studies, 54(3): 308-328.

Thrupp, M., Lauder, H and Robinson, T. 2002. School Composition and Peer Effects. International Journal of Educational Research, 37(5): 483-504.

Usher, E.L. and Pajares, F. 2008. Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions. *Review of Educational Research*, 78(4): 751-796.

Valentine, J.C., DuBois, D.L. and Cooper, H. 2004. The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational psychologist*, 39(2): 111-133.

Vallerand, R.J. and Blssonnette, R. 1992. Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of personality*, 60(3): 599-620.

Vyas, S. and Choudhary, G. 2016. Study habits of sr. sec. school adolescent students in relation to their socio-economic status. *IJAR*, 2(6): 134-139.

Webb-Williams, J. 2014. Gender differences in school children's self-efficacy beliefs: Students' and teachers' perspectives. *Educational Research and Reviews,* 9(3): 75.

Winters, M.A., Haight, R.C., Swaim, T.T. and Pickering, K. 2013. The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades: Evidence from panel data. *Economics of Education Review*, 34: 69–75.

Wood, J.V. 1996. What is social comparison and how should we study it?. *Personality and Social Psychology Bulletin*, 22(5): 520-537.