MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2022, 18(11), em2184
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/12559

The association between the gender gap in science achievement and students' perceptions of their own attitudes and capabilities

Sulaiman M. Al-Balushi [1*], Nasser Mansour [2], Rashid S. Almehrizi [1], Abdullah K. Ambusaidi [3], Ibrahim S. Al-Harthy [4]

[1] Sultan Qaboos University, Muscat, OMAN
[2] Qatar University, Doha, QATAR
[3] Ministry of Education, Muscat, OMAN
[4] National University for Science and Technology, Muscat, OMAN

Received 31 May 2022 ▪ Accepted 02 October 2022

Abstract
Among the countries that participated in the trends in international mathematics and science study (TIMSS) 2019 for grade 8 science, Oman had the highest gender gap in favor of girls. The current study explores the gender gap in science achievement in Oman and relates it to students' varying perceptions of their own attitudes and capabilities. The sample in the study comprised 467 grade 9 students, 266 female and 201 male. The participants were given a TIMSS-like science test, along with four self-perception surveys; these explored metacognitive awareness, self-regulation (SR), science learning self-efficacy (SLSE), and attitudes to science (AS). The results indicated that student self-perceptions of SR, SLSE, and AS, were significantly related to the gender gap in students with higher-level science achievement. The results were different when looking at the gender gap in scores for lower-level questions; here, there was no relation to any of the four self-perception variables explored in the study.

Keywords: gender gap, TIMSS science achievement, metacognitive awareness, self-regulation, science learning self-efficacy, attitudes to science

INTRODUCTION

A gender gap in the trends in international mathematics and science study (TIMSS) has persisted globally for over 20 years, with boys being the higher-achieving group in both Western and Far Eastern countries (Meinck & Brese, 2019). The situation has been the opposite, however, in Middle Eastern countries such as Bahrain, Egypt, Iran, Jordan, Kuwait, Qatar, Saudi Arabia, and the United Arab Emirates, and with girls achieving higher scores in science TIMSS (Mullis et al., 2020). Of all the countries that participated in the grade 8 TIMSS in 2019, Oman saw the biggest gender gap in performance in both science and mathematics, with girls out-performing boys. This disparity between the achievements of girls and boys has been evident in all the TIMSS cycles in which Oman has participated. In science, in 2007, 2011, 2015, and 2019, girls scored respectively 61, 78, 45, and 54, respectively more points in science than boys (Mullis et al., 2020). While this gender gap phenomenon in Oman has not been explored, there has long been an awareness within the educational system of the challenge of boys' underachievement. The latest report on the quality of education in Oman, produced jointly by the country's Ministry of Education and World Bank (2012), admits that boys perform badly in most subjects, including science and mathematics, and indeed in their overall academic results. The report recommends that a comprehensive empirical study should be undertaken to investigate the gender gap in science achievement in Oman and to attempt to identify the factors underlying boys' lack of achievement. Their weak performance in school has been reflected in their under-representation in tertiary education, where only 42.75% of higher education students in Oman are male (National Center for Statistics and Information, Oman, 2020). This gender gap in both school performance and tertiary

✉ sbalushi@squ.edu.om **(*Correspondence)** ✉ n.mansour@qu.edu.qa ✉ mehrzi@squ.edu.om ✉ ambusaidi40@hotmail.com ✉ ibrahimsa@nu.edu.om

Contribution to the literature

- The key contribution of the current study is that it provides further insight into the factors associated with the gender gap phenomenon in science achievement in countries with a similar gender gap pattern (where girls out-perform boys) and similar socio-cultural contexts (primarily the Middle Eastern countries).
- Unlike most studies on the gender gap in science, the current study explored the variables that might influence the gender gap for different achievement levels (higher- vs. lower-achieving students) and different test-question levels (higher- vs. lower-level test questions).
- An important outcome of the study was that while there was a significant association between the gender gap in scores for higher-level TIMSS-like questions and some self-perceptions variables, the gender gap in scores for lower-level questions did not relate to any of the self-perception variables.

representation runs counter to the international trends of male out-performance in science, mathematics and reading and of female under-representation in prestigious higher education institutions (Gray et al., 2019), and therefore clearly needs to be investigated.

Oman's participation in international science studies such as TIMSS offers a lens through which students' achievement in science can be monitored over time. It is this reason which lies behind the current study and its exploration of the factors behind the gender gap in science achievement in Oman. To the knowledge of the authors, there have been no previous studies that have addressed the TIMSS science gender gap in Oman itself, and only limited studies addressing the issue in countries with similar socio-cultural contexts (e.g., Ababneh & Abdel Samad, 2018; Barry, 2019). It is therefore hoped that the current study can provide further insight into the factors associated with this gender gap phenomenon, not only in Oman but also in other Middle Eastern countries, and that this exploration might help to enhance boys' achievement by identifying the factors that lead to their lagging behind girls. It is hoped that this will ultimately help to improve the overall level of achievement of Oman and other Middle Eastern countries in international studies like TIMSS.

The Role of Self-Perception of Attitudes and Capabilities in the Gender Gap in TIMSS Results

The lack of balance between male and female school students in attitudes towards science and mathematics, and in levels of achievement in those fields, could have a negative impact on the workforce engaged in STEM careers (Berger et al., 2020). A large body of evidence in a variety of studies attributes a gender gap in achievement to student perceptions of their attitudes towards the subject matter, to self-confidence, and to self-concept perceptions (Askin & Oz, 2020; Bofah & Hannula, 2015; Naidoo & Sibanda, 2020; Thomson, 2008; Tweissi et al., 2014). A number of studies show that self-perceptions about attitudes towards science are a predictor of student achievement in TIMSS (Al-Mutawah & Fateel, 2018; Berger et al., 2020; Geesa et al., 2020). Research has also shown a consistent relationship between science and mathematics achievements and students' corresponding self-concepts (Chiu, 2008).

A study by Thomson (2008) investigated the gender differences in results in the grade 8 science TIMSS for 2003 in Australia, in which boys out-performed girls. The findings indicated that self-confidence exerted a considerable and positive influence on science achievement. The study also claimed that attitudes and self-concept are essential to success, especially when the studying of science subjects is optional. Another study, by Askin and Oz (2020), analyzed the grade 8 TIMSS 2015 science data for all participating countries, and examined gender-related variability in science success. They found that there was a significant gender variation in the data of ten countries, including Oman. Their analysis also showed that the gender gap was associated with confidence about doing well in science, with home educational resources, and with future educational goals.

In countries where the gender gap in achievement is minimal, the gender gap in attitudes also seems minimal. In Singapore, for instance, where there was no gender gap in mathematics achievement in TIMSS 2011, there was also no gender difference in attitudes towards learning mathematics (Yoo, 2018). However, research has shown that where there is a gender gap in self-concept and attitudes towards a particular subject matter, this gender gap pattern matches the gender gap pattern in performance in that subject. Research exploring the gender gap where boys out-performed girls found that boys had higher self-esteem than girls, and that girls considered science and mathematics as male fields (Mata et al., 2012), and as less relevant to their futures than other areas of study (Dee, 2006). One example is in the data for the science and mathematics TIMSS in 2015 in Australia, where the gender gap in both achievement and attitudes favored boys (Berger et al., 2020).

Another relevant study is by Mejía-Rodríguez et al. (2020), who conducted a regression analysis to assess the gender gap in the TIMSS 2015 grade 8 mathematics data of 32 countries. They found that in the majority of countries, boys significantly out-performed girls, but that in most countries where the girls had higher self-concepts, such as Oman, girls also out-performed boys in mathematics achievement. The study also found that in these countries, including Oman, parents had greater educational expectations of their daughters than of their sons.

An interesting finding in the study above (Mejía-Rodríguez et al., 2020) was that, internationally, girls had a lower self-concept than boys. Even if the effect of achievement was considered, the girls' perceptions of their capabilities were less favorable than those of the boys. In Finland, for instance, even when girls' achievements were higher than the boys', their self-concept was lower. Berger et al. (2020), investigating the results for the Australian TIMSS 2015, came to similar but less consistent conclusions. They found that although, overall, boys have a better self-concept than girls, girls who were "enthusiastic" about learning science and mathematics had similar self-concept beliefs to that of the boys at the same attitudinal level. However, girls in the "very enthusiastic" category had a lower self-concept about mathematics than boys, in spite of their higher achievement. The current study explores similar issues related to the gender gap in science achievement in Oman and looks at different student self-perceptions of their attitudes and capabilities; these are considered in the light of the specific context of science teaching and learning in Omani schools. It is hoped that this might contribute epistemologically to the studies in this field and will introduce a new perspective on the gender gap in science achievement in schools worldwide.

The Influence of Socio-Cultural Factors on Students' Self-Perceptions

Research has attributed the relationship between student self-perception and confidence on one hand and the gender gap in achievement on the other to a variety of socio-cultural factors (Berger et al., 2020; Chiu, 2008; Mejía-Rodríguez et al., 2020; Naidoo & Sibanda, 2020; Yoo, 2018). For instance, Mejía-Rodríguez et al. (2020) argued that the gender gap in self-concept in Middle Eastern countries like Oman might be attributable to the single-sex school system there. Eisenkopf et al. (2015) argued similarly that a single-gender school environment in Switzerland enhances female students' self-confidence. A more general study by Chiu (2008) analyzed the achievements of 28 countries in the science and mathematics TIMSS 2003, as well as the students' self-concepts. She found that there were high correlations between self-concepts in science and in self-concepts in mathematics. However, there were differences in how students viewed the relationship between the two subjects. In ten countries out of the 28 countries, students' perceptions showed a contrasting process, where they considered science and mathematics as distinctly different subjects, while students' perceptions in four out of the 28 countries showed an integrating process, where they considered science and mathematics as complementary or related subjects. Remarkably, three of the countries with integrating perceptions of science and mathematics were in the Middle East (Egypt, Jordan, and the Palestinian National Authority) and one was in western Africa (Ghana). Chiu (2008) partly attributed these differences in contrasting and integrating perceptions to cultural stereotypes in the different countries. She argued that the overall negative correlation between mathematics achievement and science self-concept might suggest there are specific socio-cultural factors in some countries that act as moderating factors and impact the associations linked to achievement in science. The current paper takes into consideration the socio-cultural contexts of the Omani educational system when exploring and interpreting the gender gap in achievement and attitudes.

In another relevant study, Naidoo and Sibanda (2020) examined the gender gap in the science data for TIMSS 2015 in South Africa, where girls' scores were higher than those of boys. The researchers suspected that the gender difference could be attributed to different socio-cultural environments and learning experiences. The study also related the gender gap to a number of perceptions and attitudes, such as the students' perceptions of their ability to do science and their attitudes towards science and science learning. Berger et al. (2020) also attribute the gender gap in science and mathematics TIMSS scores to the external comparisons that influence girls' self-concept when they compare their achievement with those of their high-achieving peers. Interestingly, Yoo (2018), investigating the situation with Singaporean students, found that teachers and parents had a more significant impact on girls' confidence in mathematics than they had on that of the boys; girls' confidence also had a more significant influence on their TIMSS achievement than was the case with boys.

A THEORETICAL FRAMEWORK TO EXPLAIN THE GENDER GAP IN SELF-PERCEPTION AND ACHIEVEMENT

This section discusses two theoretical models, put forward in previous research, that seek to explain the role of student perceptions of their own attitudes and capabilities in the achievement gender gap. These models are the expectancy-value theory (EVT) (Eccles, 2009) and the internal/external (I/E) frame of reference (Marsh, 1986).

Many researchers have argued that individuals' self-perception of their attitudes and capabilities results from the interactions between a number of personal and social factors (Eccles, 2009; Khishfe & BouJaoude, 2016; Oon et al., 2020). Eccles (2009) proposed a model based on this idea, called the EVT, which explained individuals' performance and choice by linking them to task-related expectancy and value beliefs. EVT has been used to explain individuals' motivation, choices, persistence, and attitudes (Eccles, 2009; Geesa et al., 2020).

Eccles (2009) predicts that individuals' expectations and the values they attach to different options are the most influential psychological factors that affect their decisions as to how to spend their energy and time. She argues that the collective social idea of gender roles influences boys and girls to place different values on different goals and achievements. Eccles (2009) explains, for instance, that if girls and boys receive different information, opportunities, opinions, and feedback on the importance of different school subjects, they are likely to develop different self-perceptions and different patterns of expectations of success; they will also place different values on vocational and educational paths (Eccles, 2009; Geesa et al., 2020). Geesa et al. (2020) used EVT to explain the level of performance of South Korean students in TIMSS 2015 science by pointing out that education is very highly regarded in South Korea and constitutes a collective cultural value, exemplified by the common belief that success in school leads to success in life.

The second theory seeking to explain how students' self-perceptions of attitudes affect their performance is the model posited by Marsh (1986) and named the I/E frame of reference model. Brunner et al. (2008) explain that the I/E model assumes that two comparison processes are involved in the formation of academic self-concepts. First, students engage in a social comparison process within an external frame of reference, comparing their ability in a particular domain with the ability of other students in the same domain. Second, students apply an internal frame of reference to compare their ability in a particular domain with their ability in other domains (Brunner et al., 2008, p. 140).

In another study, Marsh and Hau (2004) found that when students judge their self-concept of a particular school subject, they use both an internal frame of reference, when they compare their performance in one subject with their performance in other subjects, and also an external frame of reference, when they compare their own performance with that of other students. In another study, March et al. (2008) found that students' external comparisons with their peers could lead them to underestimate or overestimate their own capabilities.

Brunner et al. (2008) expanded the I/E model to include potential relationships between domain-general academic self-concept and domain-general cognitive ability. Other research (Abu-Hilal, 2005; Berger et al., 2020; Ogunjuyigbe et al., 2006) has indicated that girls and boys do not have a similar internal/external frame of reference when they assess their self-concept regarding their achievement. Berger et al. (2020), for instance, used the I/E model to explain why girls had lower self-concept perceptions than boys by arguing that external comparisons with high-achieving peers influenced self-perceptions in girls, but did not do so in boys. In the current study, this I/E model is used to interpret possible cultural factors in Oman that might affect the gender gap when students engage with the processes of the formation of their academic self-concept of science.

Perceptions of Self-Regulation, Metacognition, and Learning Self-Efficacy as Sources of the Gender Gap in Science Achievement

A number of previous research studies have focused on student AS as a source of the gender gap in science achievement and TIMSS performance (Al-Mutawah & Fateel, 2018; Askin & Oz, 2020; Berger et al., 2020; Bofah & Hannula, 2015; Chiu, 2008; Geesa et al., 2020; Naidoo & Sibanda, 2020; Thomson, 2008; Tweissi et al., 2014). However, other literature in science education presents other variables that play an essential role in student achievements in science. These variables include self-regulation (SR) skills, metacognitive awareness (MA), and science learning self-efficacy (SLSE), and it seems likely, in the light of previous research, that these variables, along with AS, play a part in shaping the existing gender gap in science TIMSS performance (Gestsdottir et al., 2014; Kurman, 2001; Matthews et al., 2009; Schwery et al., 2016; Vantieghem et al., 2014; Williams et al., 2016).

One crucial variable that influences student science achievement is metacognition (Casselman & Atwood, 2017; Hong et al., 2020; Oyelekan et al., 2019; Wang & Chen, 2014; Wang et al., 2014). Metacognition enhances students' ability to retrieve prior knowledge from long-term memory (Al -Harthy, 2016), a critical test-taking and science-learning process. It has been shown to characterize the thinking skills of high achievers in science (Larkin, 2006) and to be positively associated with science learning achievement (Hong et al., 2020; Oyelekan et al., 2019). Moreover, metacognition has been positively associated with performance in science exams (Casselman & Atwood, 2017) as well as with two learning aspects that are essential for test-taking: comprehension of science texts (Wang & Chen, 2014; Wang et al., 2014) and problem-solving skills (Akben, 2020; Aurah et al., 2014).

Another variable essential to learning in science is SR. Research shows that student SR skills are associated with high learning motivation, task-orientation cohesiveness in science classrooms, and positive perceptions of scientific investigations (Velayutham & Aldridge, 2013). Self-regulated learners have also been found to have a high inherent value that is significantly linked to their science achievement (Kingir et al., 2013); consequently, they gain more from their learning of science than their counterparts who have lower SR skills (Zheng et al., 2020).

A third variable stressed in science education literature is learning self-efficacy, which plays an important role in influencing student science achievement (Tsai et al., 2011; Wan, 2021). Although the link between self-efficacy and performance in science TIMSS has been positively identified (Wan, 2021), self-efficacy has rarely been studied in relation to the gender gap in science TIMSS performance, a research gap which the current study seeks to address.

The Interaction Between Gender, Science Achievement, and Self-Perceptions

Despite substantial work in science education literature on the relationship between cognitive and attitudinal variables and achievement, there has not been a great deal of research examining the link between the gender gap in achievement and self-perception variables, but there have been several interesting findings. For instance, in the study by Berger et al. (2020) of the Australian TIMSS 2015 science and mathematics results, described earlier, a relationship between achievement and gender could be inferred from the mathematics data but not from the science data. Mejía-Rodríguez et al. (2020), looking at the TIMSS 2015 mathematics results in Finland, also found a relationship between gender and student achievement. Interestingly, this study showed that even though the girls' achievement was higher than that of the boys, their self-concept was actually lower. Smith et al. (2014) investigated the relationship between gender, enjoyment of science and confidence in doing the subject among grade 4 and grade 8 students in the USA. They found that student gender affected the relationship between science achievement and enjoyment of science for grade 4 students, and also affected the relationship between science achievement and confidence in science for grade 8 students. In both cases, this relationship was stronger in boys than in girls. Grade 4 boys who liked the subject more and had greater confidence about doing it also achieved better results.

A number of researchers have called for further exploration of the relationship between student attitudes towards science and science achievement (Berger et al., 2020; Thomson, 2008), and how that affects the gender gap. The current study is an attempt to respond to this call, and although a single study relating to a single country will be unable to make sense of the whole gender gap phenomenon, it is hoped that its exploration will help to clarify some of the factors influencing this gap and will contribute to the overall study of the phenomenon.

This study focuses on the situation in schools in Oman, in the hope that that its findings might be of use not only to the educational system there, but also to those in countries that share a similar gender gap pattern (where girls out-perform boys) and similar socio-cultural contexts (primarily the Middle Eastern countries). The aim is then for these countries to design solutions that tackle this gap in achievement and attitudes, to support boys to achieve better results in science, and to reach a situation where there is a gender balance in students entering future STEM careers.

Purpose of the Study

The purpose of the current study, then, was to explore the gender gap in student science achievement in relation to four types of self-perceptions of attitudes and capabilities: MA, SR, SLSE, and AS. Most studies exploring these interactions have examined Western data in Western contexts, where the gender gap tended to favor boys (e.g., Berger et al., 2020; Smith et al., 2014).

The findings of the current study, in contrast, will illuminate the gender gap phenomenon in Oman and similar countries, where science achievement tends to favor girls (Mullis et al., 2020). Another issue addressed - to our knowledge - for the first time in this study, is the research studying the variables that influence the gender gap in the performance of higher-scoring and lower-scoring students, and in their achievement in higher-level and lower-level test questions. This detailed exploration will hopefully give a deeper insight into the variables affecting the achievement gender gap at each level of test questions.

The main research question for the current study is therefore:

What is the nature of the relationship between the gender gap in science achievement and students' perceptions of their own attitudes and capabilities?

Two further research questions stemmed from this main question:

1. What is the nature of the relationship between the gender gap in lower-level science achievement and students' self-perceptions of their attitudes and capabilities?
2. What is the nature of the interaction between the gender gap in higher-level science achievement and students' self-perceptions of their attitudes and capabilities?

Context

The school system in the Sultanate of Oman comprises three stages: cycle I schools are for grades 1-4; these are mixed-gender and are taught by female teachers only, cycle II (grades 5-10), and cycle III (grades 11-12) are single-gender schools taught by teachers of the same gender as the students. In Omani schools, girls achieve better results than boys at all levels (Ministry of Education & World Bank, 2012). This is especially true in science and mathematics, and is reflected in TIMSS results (Mullis et al., 2020). This gender gap was the main motivation for the current study.

Table 1. Demographics of students, classes, and schools participating in the study

School	Grade levels	NS	NC	Gender	NC-1	NC-2[a]
1	5-9	763	24	Girl	2	63
2	5-9	776	24	Boy	3	96
3	5-10	836	27	Boy	3	96
4	1-9	819	26	Girl	2	64
5	7-10	837	28	Girl	4	130
6	5-12	978	32	Girl	2	59
7	5-12	981	28	Girl	2	68

Note. NS: Number of students; NC: Number of classes; NC-1: Number of classes participating in each grade level; NC-2: Number of students participating in each grade level[a]; & [a]These numbers differ from the actual number of participants included in the current study analysis as we included only the participants who completed all of the five study instruments

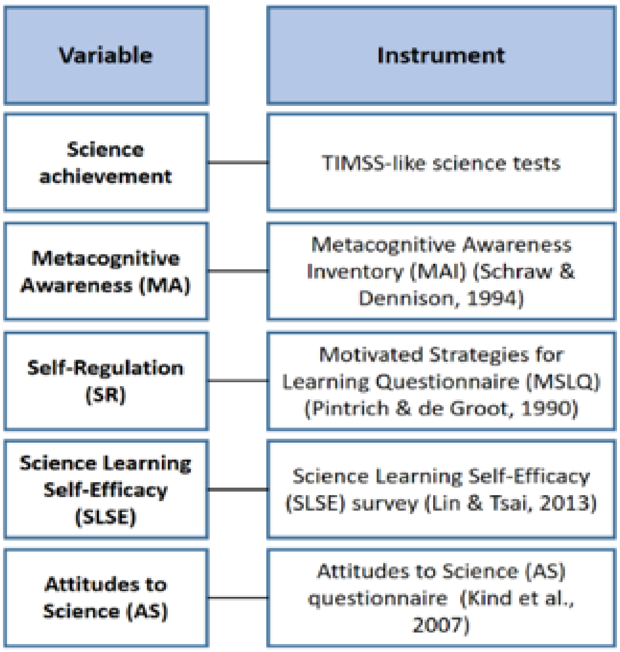

Figure 1. Study variables and the research instruments used to measure them (Source: Authors own elaboration)

METHODS

The current study is exploratory. We used a TIMSS-like test to examine students' science achievement and also surveyed their self-perceptions, namely MA, SR, SLSE, and AS. We explored the interaction between these four variables and the gender gap in lower-level and higher-level science achievement, as measured by student scores in lower-level and higher-level test questions.

Participants

The sample comprised 467 grade 9 Omani students, 266 girls and 201 boys, studying in 17 classes selected randomly from seven different schools located in three neighboring regions (**Table 1**). Although the original TIMSS study is usually done with grade 8 students, the date set for the study was the end of their school year when there was not enough time to test them. Instead, the study was postponed until the start of next school year when the students were just beginning grade 9, on the assumption that the knowledge they might have gained during the vacation was insignificant. The socio-economic status of students in these schools was largely middle-class, with most participants' parents holding secondary school diplomas or the equivalent, and many having higher education qualifications. Before the study was conducted, research ethics approval was obtained from the Ministry of Education in the Sultanate of Oman, and from each of the host schools, which in turn obtained parental consent in accordance with the procedures of each school.

Design and Procedure

Participants completed five different measures. One instrument, the TIMSS-like test, measured their science achievement, and four surveys explored their perceptions of their own attitudes and capabilities. All the instruments were mounted in a mobile application designed for the research project; the app was then installed on mobile tablets which were given to the participants. They were given 40 minutes to complete the TIMSS-like science test and 10 minutes to complete each self-perception survey. Because of lesson timing (40 minutes), they completed the TIMSS-like test and the surveys on different days.

One science teacher was recruited as a research assistant in each of the seven schools; their task was to conduct the study in randomly selected ninth-grade classrooms. The teachers were trained by the research team on how to conduct the study and how to help the participants use the mobile tablets. Two technicians were also present at each school while the study was being administered, to assist in resolving any technical issues.

Instruments Used in the Study

The five instruments used in the study are illustrated in **Figure 1** and described below.

TIMSS-like tests

The current study was motivated by the significant gender gap that exists in Omani student performance in science TIMSS (Mullis et al., 2020).

Item 2: What does the following equation represent?

Water + carbon dioxide ➔ food + oxygen

A. Photosynthesis B. Breathing C. Global warming D. Burning

Item 20: Lila and Rana were discussing gas viscosity.

When Lila said that gases were more viscous than liquids, Rana disagreed.

Who was right, Lila or Rana? ……………… Explain ……………………………………………………………………………

………

Figure 2. Examples of the TIMSS-like science test items (Source: Authors own elaboration)

As a result, our science achievement test was designed to be similar to the actual TIMSS science tests. Fifteen experienced local science teachers were selected to write the tests, along with five graphic designers whose job was to design the illustrations for some of the items. The teachers were trained by an international expert in assessment and measurement in the nature of TIMSS assessment and in how to develop TIMSS-like questions. According to the TIMSS framework, different test booklets are designed, and each participant is randomly assigned one of these versions (Mullis & Martin, 2017). To enable the same process to be followed, the research team and the trained teachers then designed 18 different matching versions of the TIMSS-like tests, each with 27 items. In accordance with the framework, eight items were knowledge (lower-level) questions, and 19 were application and reasoning (higher-level) questions. Since the language of instruction in Omani public schools is Arabic, the tests were written in Arabic. To ensure the matching accuracy of the different versions, they were assembled according to the test blueprint described in the TIMSS framework.

Figure 2 gives an example of a lower-level test question (item 2) and also an example of a higher-level test question (item 20), both translated from Arabic. The total score and the sub-totals for lower-level and higher-level questions were computed as mean scores, to overcome the issue of differences in scores caused by their different number of items. Hence, all scores on TIMSS-like tests ranged between 0 and 1.

Each version had different types of questions: textual, pictorial, multiple-choice, and short open-ended, with graphic designers designing the pictures used in the questions. The test then went through two review stages; first by a panel of local TIMSS assessment experts who had long experience managing TIMSS assessment in Oman; they checked item accuracy and the matching of the versions. The tests were then reviewed by an international TIMSS expert. We then piloted the tests on 575 students, who were randomly selected from different schools located in different regions in Oman, with each version of the test being piloted in two classes. Items with low reliability were then revised. The final Cronbach's alpha reliability coefficients ranged between 0.661 and 0.885, with an average of 0.779. Given that the reliability estimates for the 2015 TIMSS grade 8 science test ranged between 0.74 and 0.88 (Martin et al., 2016), the estimates of test reliability in the current study were judged to be acceptable. More details about the design of the tests, the accuracy of version matching, and the validity and reliability measures taken can be found in Al-Balushi et al. (2022).

Self-perceptions of attitudes and capabilities

Four different surveys were used to measure students' self-perceptions of their MA: SR, SLSE, and AS; these were selected because they had been widely used in previous research to measure the same variables. All the surveys used a five-point Likert scale: from strongly agree (5) to strongly disagree (1). In order to reduce the load on the participants, who had the burden of completing all five instruments, we selected only the items most relevant to the intended variable and then translated them into Arabic. Two experts who were fluent in Arabic and English checked the accuracy of the translation, and a panel of five specialists in psychology and science education reviewed each modified instrument, to decide whether they accurately measured the intended variables and whether they were appropriate to the age of the targeted participants. We then piloted each modified instrument on 120 students who represented three different classes which were selected randomly from three different schools. The surveys are described below:

1. **Metacognitive awareness (MA):** We used an Arabic version of the metacognitive awareness inventory (MAI) by Schraw and Dennison (1994); this had four sub-scales. We selected 22 out of the 52 items in the original instrument: planning (five items), information management strategies (six items), debugging strategies (six items), and evaluation (five items). The calculated Cronbach's alpha reliability coefficient was 0.95.

Table 2. Means and standard deviations for different self-perception variables by gender

Variable	Gender	n	Mean	Standard deviation
Metacognitive awareness (MA)	Boy	195	3.62	0.95
	Girl	264	3.88	0.62
Self-regulation (SR)	Boy	195	3.55	0.80
	Girl	266	3.84	0.57
Science learning self-efficacy (SLSE)	Boy	195	3.52	0.88
	Girl	245	3.52	0.61
Attitudes to science (AS)	Boy	201	3.45	0.63
	Girl	263	3.53	0.62

2. **Self-regulation (SR):** We used an Arabic version of the motivated strategies for learning questionnaire (MSLQ) that had been used to assess SR (Pintrich & de Groot, 1990). The original instrument had 55 items divided into four sub-scales, and we selected 20 items: goal orientation (6 items), control of learning beliefs (5 items), rehearsal (6 items), and test anxiety (3 items). The calculated Cronbach alpha reliability coefficient was (0.91).
3. **Science learning self-efficacy (SLSE):** This survey was designed by Lin and Tsai (2013), and had 32 items, in five sub-scales, in its original form. We selected 21 items: conceptual understanding (three items), higher-order cognitive skills (four items), practical work (four items), everyday application (six items), and science communication (four items). The calculated Cronbach's alpha reliability coefficient was 0.93.
4. **Attitudes to science (AS):** We used an Arabic version of a questionnaire by Kind et al. (2007) to assess students' attitudes towards science and science learning; the questionnaire had six sub-scales. We selected 25 of the original 46 items: learning science in school (five items), practical work in science (four items), science outside of school (four items), importance of science (four items), self-concept in science (four items), and future participation in science (four items). The calculated Cronbach's alpha reliability coefficient was 0.90.

Data Collection

Since the TIMSS testing format was changed from paper-based to electronic in the 2019 cycle, we created a parallel format by designing a mobile application called trends in Oman science study (TOSS), which was loaded onto iPads. We piloted the app on a purposeful sample of 120 students, a process which resulted in minor changes. Seven science schoolteachers, who had been hired as research assistants in their schools, then administered the study instruments to the actual study sample (**Table 1**); each teacher was supported by two IT technicians to assist in the administration of TOSS on the iPads. One iPad was issued to each participant.

Data Analysis

Several statistical analyses were used to answer the research questions:

1. Descriptive statistics about students' self-perceptions of their attitudes and capabilities were computed for boys and girls.
2. The total scores on each student's self-perception type were classified into two groups based on their median scores in each of the four self-perception surveys (the low group having total scores<median, and the high group having total scores≥median).
3. Two-way ANOVA analyses were conducted to examine the interaction between the gender gap in both lower and higher-level science achievement (as measured by their scores in the lower and higher-level TIMSS-like test questions respectively), and students' self-perceptions of their attitudes and capabilities.
4. Means were plotted for the dependent variables that showed statistically significant relationships.

RESULTS

Table 2 presents the means and standard deviations of boy and girl participants in each type of self-perception surveyed in the current study.

Table 3 displays the descriptive statistics (means and standard deviations) for the scores in the lower-level and higher-level questions in the TIMSS-like test, divided according to gender and also by the two levels (high and low) of each type of student self-perception about attitudes and capabilities.

These results showed that girls had higher mean scores than boys on the two scores of the TIMSS-like test (lower- and higher-level test questions) for all groups of students, with either a high level or a low level of the four variables of students' self-perceptions about attitudes and capabilities. We then carried out two-way ANOVA analyses for the two scores of the TIMSS-like test in order to examine the interaction of gender gap and two-level groups of the four self-perception variables.

Table 3. Gender gap for performance on TIMSS-like test questions (lower- and higher-level) on different level groupings of self-perceptions about attitudes and capabilities

Variable	PL	Gender	Lower-level questions			Higher-level questions		
			n	Mean	SD	n	Mean	SD
Metacognitive awareness (MA)	Low	Boy	106	0.24	0.16	106	0.27	0.18
		Girl	103	0.39	0.18	103	0.46	0.19
	High	Boy	89	0.32	0.21	89	0.37	0.22
		Girl	161	0.45	0.16	161	0.52	0.16
Self-regulation (SR)	Low	Boy	112	0.23	0.17	112	0.26	0.18
		Girl	96	0.40	0.18	96	0.47	0.18
	High	Boy	83	0.34	0.20	83	0.39	0.20
		Girl	170	0.45	0.16	170	0.49	0.16
Science learning self-efficacy (SLSE)	Low	Boy	97	0.23	0.17	97	0.27	0.18
		Girl	123	0.41	0.16	123	0.49	0.18
	High	Boy	98	0.32	0.20	98	0.36	0.21
		Girl	122	0.46	0.19	122	0.50	0.17
Attitudes in science (AS)	Low	Boy	119	0.24	0.16	119	0.28	0.19
		Girl	139	0.41	0.17	139	0.49	0.19
	High	Boy	82	0.33	0.22	82	0.38	0.21
		Girl	124	0.46	0.17	124	0.50	0.16

Note. PL: Perception level & SD Standard deviation

Table 4. Two-way ANOVA analysis of performance on lower-level TIMSS-like test question, in relation to gender and the different self-perception variables

Variable	Effect	SS	df	MS	F	p-value	Eta-squared
Metacognitive awareness (MA)	MA (high vs. low)	0.57	1	0.57	18.11	0.000*	0.038
	Gender	2.21	1	2.21	70.22	0.000*	0.134
	MA x gender	0.01	1	0.01	0.44	0.508	0.001
	Error	14.32	455	0.03			
Self-regulation (SR)	SR (high vs. low)	0.62	1	0.62	20.19	0.000*	0.042
	Gender	2.07	1	2.07	67.35	0.000*	0.128
	SR x gender	0.10	1	0.1	3.15	0.077	0.007
	Error	14.07	457	0.03			
Science learning self-efficacy (SLSE)	SLSE (high vs. low)	0.58	1	0.58	18.02	0.000*	0.040
	Gender	2.79	1	2.79	86.94	0.000*	0.166
	SLSE x gender	0.04	1	0.04	1.19	0.275	0.003
	Error	13.97	436	0.03			
Attitudes in science (AS)	AS (high vs. low)	0.47	1	0.47	14.77	0.000*	0.031
	Gender	2.53	1	2.53	78.75	0.000*	0.146
	AS x gender	0.04	1	0.04	1.39	0.239	0.003
	Error	14.78	460	0.03			

Note. *p<0.001

Table 4 looks at the results for lower-level TIMSS-like test questions and shows that there were statistically significant differences in the test scores of low- and high-level perception groups for each of the four variables. Participants with a higher level of the four types of self-perceptions of attitudes and capabilities had significantly higher scores for lower-level TIMSS-like questions. The results also showed that there was a statistically significant gender gap in students' performance on lower-level TIMSS-like test questions. However, there was no statistically significant relationship between the gender difference in scores on these questions and any of the variables of students' self-perceptions about their attitudes and capabilities. This finding indicated that the effect of gender did not depend on the effect of the students' self-perceptions about their attitudes and capabilities when looking at their performance on these lower-level questions.

Table 5 presents the results of the analysis of the test scores on the higher-level TIMSS-like questions, showing that there were statistically significant differences between the test scores of the low and high-level perception groups for each of the four variables. Participants with a higher level of the four types of self-perceptions also had significantly higher scores in the higher-level questions. The results also showed that girls achieved statistically significant higher scores than boys on these questions. There was also a statistically significant relationship between the gender gap in higher-level science scores and three of the four self-perception variables, namely SR, SLSE, and attitudes towards science, indicating that these three variables had

Table 5. Two-way ANOVA analysis for performance on higher-level TIMSS-like test questions by gender and different variables

Variable	Effect	SS	df	MS	F	p-value	Eta-squared
Metacognitive awareness (MA)	MA (high vs. low)	0.65	1	0.65	19.78	0.000*	0.042
	Gender	3.15	1	3.15	95.48	0.000*	0.173
	MA x gender	0.06	1	0.06	1.78	0.183	0.004
	Error	15.01	455	0.03			
Self-regulation (SR)	SR (high vs. low)	0.57	1	0.57	17.91	0.000*	0.038
	Gender	2.59	1	2.59	81.96	0.000*	0.152
	SR x gender	0.28	1	0.28	8.97	0.003*	0.019
	Error	14.45	457	0.03			
Science learning self-efficacy (SLSE)	SLSE (high vs. low)	0.27	1	0.27	7.96	0.005*	0.018
	Gender	3.49	1	3.49	103.06	0.000*	0.191
	SLSE x gender	0.14	1	0.14	4.01	0.046*	0.009
	Error	14.76	436	0.03			
Attitudes in science (AS)	AS (high vs. low)	0.42	1	0.42	12.14	0.001*	0.026
	Gender	3.12	1	3.12	90.79	0.000*	0.165
	AS x gender	0.23	1	0.23	6.58	0.011*	0.014
	Error	15.83	460	0.03			

Note. *p<0.001

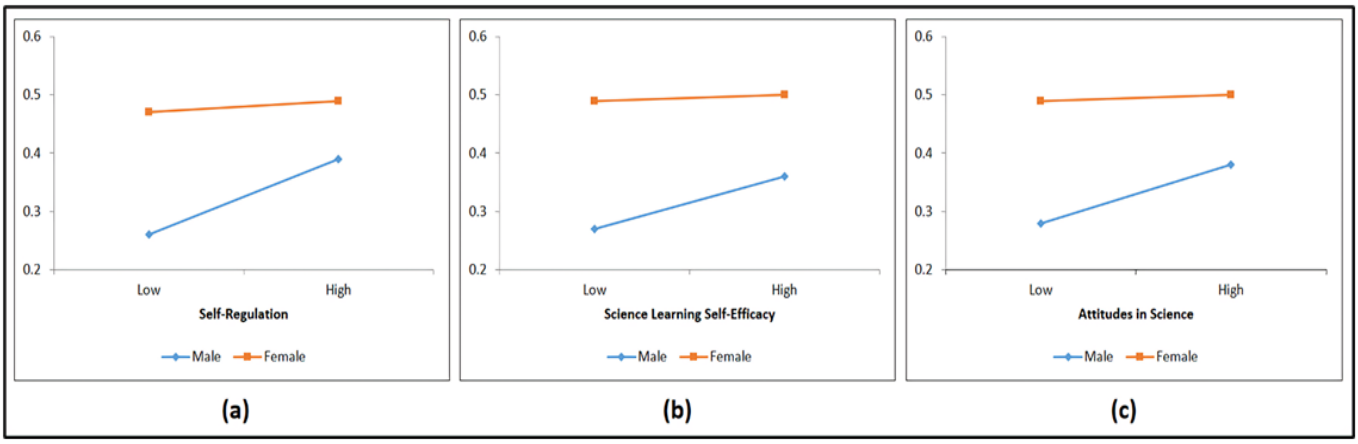

Figure 3. Mean plots of gender gap in higher-level questions of TIMSS-like test on different levels of self-regulation (SR), science learning self-efficacy (SLSE), and attitudes in science (AS) (Source: Authors own elaboration)

a statistically significant influence on the gender gap in scores on these questions.

Figure 3 presents mean plots for these three variables and shows that the gender gap in scores on these questions was greater for the low-level perception group for these three variables than it was for the high-level perception group. Figure 3 also shows that the difference in scores was greater between boys with low and high-level perceptions in these three variables than it was between girls of low- and high-level perceptions in the variables.

The Gender Gap and Student Self-Perceptions and Performance

The current study explored the gender gap in science achievement in TIMSS-like tests and how this was related to student self-perceptions of four areas of their attitudes and capabilities: MA, SR, SLSE, and AS. Achievement was measured separately for scores in high-level and low-level TIMSS-like test questions. The findings indicated that, in the case of performance on higher-level questions, there was a significant relationship between gender and student self-perceptions of SR, SLSE, and AS, but none related to MA. We therefore concluded that these three types of self-perceptions influenced student performance in higher-level questions of TIMSS-like tests, and the gender gap in that performance. The situation was different in the case of performance on lower-level questions, where there was no sign of statistically significant interaction between gender, the gender gap, and any of the four self-perception variables explored in the current study. Additionally, although girls had significantly greater metacognitive perception than boys, this had no impact on the gender gap in science achievement in these test questions.

Given the limited research exploring gender gap in student performance on higher-level science questions, the finding constitutes a significant contribution to studies of gender gap in science achievement.

The international TIMSS 2019 report on science indicated that students struggled more with higher-level science test questions than with lower-level ones (Mullis et al., 2020). Other studies reported similarly, and unsurprisingly, that students found higher-level science test items more complicated than lower-level ones (Kablan & Kaya, 2013; Liou & Bulut, 2020). Higher-level questions in TIMSS include questions in the "applying" and "reasoning" levels. According to the TIMSS science assessment framework (Mullis & Martin, 2017), in the "applying" questions, students are required to "compare, contrast, and classify groups of objects or materials; relating knowledge of a science concept to a specific context; generating explanations; and solving practical problems", while in the "reasoning" questions, they are required to "use evidence and science understanding to analyze, synthesize, and generalize, often in unfamiliar situations and complex contexts" (p. 52).

In the current study, the interaction of student self-perceptions of their attitudes and capabilities with the gender gap in science higher-level achievement could indicate that when male students have greater SR, SLSE, and more positive attitudes towards science, they are more likely to achieve better results. As **Figure 3** shows, the gender gap in higher-level questions of TIMSS-like questions was wider for the lower-level perception group than for the higher-level perception group.

Our findings indicate a clear gender pattern related to students' evaluation of their SR, SLSE, and AS. Female students had better self-perceptions than male in each of these three variables. Interestingly, **Figure 3** shows that there was a bigger difference in higher-level science achievement between boys with high- and low-level perceptions than there was between girls with high- and low-level perceptions; that is, boys with low self-perceptions are likely to do worse in high-level science questions than girls with low self-perceptions. This indicates a strong link between male students' self-perceptions of their attitudes and capabilities and their higher-level science achievement. Indeed, a substantial body of previous research has attributed the gender gap in science and mathematics achievement to a parallel gender gap in student self-confidence and attitudes towards the subject matter, including self-concept, inherent value, and confidence in that particular subject (Berger et al., 2020; Desy et al., 2011; Mata et al., 2012; Smith et al., 2014; Thomson, 2008; Yoo, 2018). Gender has been found to play a significant role in AS and also in TIMSS science achievement in a range of countries (Geesa et al., 2020).

A study by the Alrajhi et al. (2019), looking specifically at Oman, found that female students had a higher self-concept than male students in all three grade levels included in the sample (grades 7, 9, and 11). Interestingly, as students progressed to grade 11, the self-concept perceptions of boys grew weaker, while those of girls tended to grow stronger. Omani girls also exhibited higher academic self-efficacy and task value than Omani boys (Al-Harthy & Aldhafri, 2014).

Single-Gender School Context

Most countries, where girls out-performed boys most strongly in TIMSS science were in the Middle East (Mullis et al., 2020) and had single-gender school systems. The current study has provided new insights into the impact of self-perceptions of attitudes and capabilities on the gender gap in science achievement in a situation where students study in single-gender schools. The findings could be explained partially by the fact that Oman has single-gender schools in which teachers and managers are all the same gender as the students. Some researchers (Eisenkopf et al., 2015; Mejía-Rodríguez et al., 2020) have argued that single-gendered schooling provides greater support to girls' self-concept than it does for that of boys.

In Oman, a key factor in the single-gender school system is the difference between the attitudes to teaching of male and female science teachers. This difference goes some way to explaining the gender gap in science achievement and in the different self-perception variables discussed in this study, as well as in the significant link between them. For instance, research shows that female science teachers in Oman have significantly more positive attitudes towards teaching science than do male science teachers (Ambusaidi & Al-Farei, 2017). A similar study by Abu-Hilal et al. (2016) found that female teachers in Oman, across all subjects, had significantly higher self-efficacy perceptions than their male counterparts about their ability to motivate students and engage them in learning. This is important because engaging students is widely seen as leading to more positive attitudes to learning (Tran et al., 2012). Another Oman-based study by Aldhafri and Alrajhi (2014) demonstrated the idea that teachers have a key impact on their students' self-perceptions; its main finding was that students who perceived their teachers' teaching style as authoritative and psychologically supportive tended to have higher inherent motivation towards learning than those who perceived their teachers' style as authoritarian.

It is clear that, in a single-gender school context, gender differences in science teachers' attitudes and capabilities will lead to a more favorable learning environment for one gender. Alkharusi (2008), for example, found that in Oman, female students in grade 9 science classes had significantly higher mastery goals, which focused on improving competence, than students in male classes; female students also had higher performance goals, which focused on displaying competence. He also reported that highly experienced female science teachers could narrow the gap in science performance between low- and high-efficacy female students in their classes, while male teachers did not exhibit this ability for their students. These findings could partially explain our finding of the interaction between student performance level and gender regarding different self-perception variables. In our female sample, the achievement difference between students with high and low-level self-perceptions was much lower than the difference observed in the male sample. It might well be plausible to infer from Alkharusi's (2008) study that it was the female science teachers who were responsible for narrowing this gap, and for motivating and helping low self-perception girls to achieve better results.

The influence of same-gender teaching has also been observed in other countries. Lee et al. (2019) analyzed standardized assessment data of 31,000 grade 6 students in 1,800 primary schools from ten francophone Western and Central African countries; their aim was to explore the relationship between teacher gender, student gender, achievement, and subject appreciation. They found that when girls were taught by female teachers, their mathematics and reading performance were enhanced, as was their subject appreciation. In the case of boys, however, this was not the case, as the gender of their teachers had no impact on their achievement. However, boys taught by male teachers did show significantly higher subject appreciation. There have been other reports of the positive effects of same-gender teachers on student achievement in both developing and developed countries. Dee (2007), for instance, reported that when students were assigned to a same-gender teacher, it enhanced the achievement and subject engagement of both boys and girls. These results, and those of previous Omani studies mentioned above (Abu-Hilal et al., 2016; Aldhafri & Alrajhi, 2014; Alkharusi, 2008; Ambusaidi & Al-Farei, 2017), could indeed serve to explain the findings of the current study. Single-gender schooling may have improved girls' self-perceptions about their attitudes and capabilities but been less effective with boys. However, further research is needed to provide greater insight into the relationship between gender, teaching efficacy, the attitudes of Omani science teachers, and student perceptions about their own attitudes and capabilities.

The Eccles' (2009) model of gender-collective identities might also provide further insight into the influence of single-gender schooling on students' perceptions about their attitudes and capabilities. This theoretical model might well suggest that the gender differences reported in the current study resulted from the formation of collective identities associated with the country's cultural context and segregated school system. It can be argued that, in the Omani single-gender school context, boys and girls receive different information, opportunities, opinions, and feedback. Although this assumption needs to be verified by further investigation, it is supported by a number of studies already carried out. These have shown that female teachers in Omani schools have more positive and effective attitudes towards teaching science than do male teachers (Ambusaidi & Al-Farei, 2017); they have also been shown to have higher self-efficacy perceptions (Abu-Hilal et al., 2016) and a greater ability to narrow the performance gap between high- and low-efficacy female students (Alkharusi, 2008). This might help to explain the different gender-related patterns in higher-level achievement found for students with low self-perceptions of their attitudes and capabilities (**Figure 3**). Abu-Hilal et al. (2016) also attributed the way that female teachers in Omani schools out-perform male teachers, to the fact that Omani girls enter teacher education programs in Oman with significantly higher school grades than their male counterparts. Indeed, it is now extremely competitive for Omani female high school graduates to gain a place on a teacher training program. This is exacerbated by Oman's cultural context, in which teaching is one of the most acceptable professions for women. Male high school graduates, in contrast, tend to opt first for specializations other than teaching, and, as observed by Abu-Hilal et al. (2016), most of them enter the profession only when other options, such as engineering and medicine, are not available.

Other Socio-Cultural Factors

It is widely agreed that socio-cultural factors have a major impact on student self-perceptions and confidence (Berger et al., 2020; Chiu, 2008; Mejía-Rodríguez et al. 2020; Naidoo & Sibanda, 2020; Yoo, 2018). Students' schools and cultures shape their perceptions of their own attitudes and capabilities, and their responses to measures of their self-perceptions are influenced by cultural factors (Bofah & Hannula, 2015). The possible effects of the single-school context have already been discussed, but another culturally related factor may also have contributed to the gender gap reported in the current study. This is the impact of people such as parents and peers in shaping gender differences in the students' internal/external frame of reference. Marsh (1986) describes this theoretical factor and argues that the development of both male and female self-concepts is influenced by different frames of reference. Other researchers have described similar findings. Berger et al. (2020) reported that girls' self-concepts were influenced by external comparisons with high-achieving peers, while research by Ogunjuyigbe et al. (2006) also indicated that girls were more influenced by peer and parental pressures than were boys.

An interesting study by Abu-Hilal (2005), conducted in the United Arab Emirates, whose Middle Eastern cultural context is similar to that of Oman, found that girls with a high self-concept attributed their performance to internal causes such as effort, while girls with a low self-concept attributed their performance to external causes such as luck or their teachers being biased against them. In contrast, boys, whatever their performance level or self-concept, did not attribute their achievement, or lack of it, to any external cause. This difference may help to explain the pattern observed earlier in **Figure 3**, where, for female learners, the difference between the achievements of those with high and low self-perceived attitudes and capabilities was smaller than the difference between the achievements of male learners with high and low self-perception levels. Further research is needed to explore the association, in the Omani context, between students' science achievement, their self-perceptions, and different frames of reference for attributing their self-perceptions.

Overall, Omani girls have greater advantages than boys, in that they have both parents and teachers who boost their self-perceptions of their attitudes and capabilities. Omani girls live with parents who expect their daughters to achieve more than their sons academically (Mejía-Rodríguez et al., 2020); they also study with female teachers who exhibit more positive attitudes, higher self-efficacy, and greater learning support abilities than male teachers (Abu-Hilal et al., 2016; Alkharusi, 2008; Ambusaidi & Al-Farei, 2017). The external influence of parents and teachers could go a long way to explain the fact that girls consistently out-perform boys in TIMSS science tests. Another research study, looking at students in Singapore, supports this interpretation; it argues that self-concept and confidence have a greater influence on TIMSS achievement in girls than in boys (Yoo, 2018). In addition, it has been argued that girls' self-perceptions are more significantly influenced by external factors than are the self-perceptions of boys (Berger et al., 2020), and that, specifically, the expectations of teachers and parents have a greater impact on girls than on boys (Yoo, 2018). However, so far there has been only very limited research on the impact of general socio-cultural factors on student self-perceptions and confidence in the context of Oman, and on the possible role played by parents and teachers in influencing students' science achievement; further investigation is needed.

RECOMMENDATIONS & LIMITATIONS

Our findings indicate that the gender gap in achievement in higher-order science skills is related to students' perceptions of their own SR skills, SLSE, and AS. This means that the gender gap in science achievement in Oman could be narrowed, especially in relation to TIMSS higher-level questions, if male students' self-perceptions of these variables were enhanced. Research has also shown that the gender gap in science achievement could be reduced by the use of culturally responsive pedagogy that engages students in learning, empowers their agency, acknowledges the knowledge, experiences, and ideas they bring to the classroom, and supports their ownership of their learning (Riegle-Crumb et al., 2019). For instance, a number of countries where boys out-perform girls have launched a variety of projects to encourage girls' participation in science and mathematics, and some of these have been effective in reducing the gender gap. One example is the campaigns carried out in Australia in the 1970s and 1980s, described by Thomson (2008); these aimed to reduce the gender gap by encouraging girls to participate more in mathematics and the sciences. The success of these campaigns was shown by the fact that there was a significant absence of gender differences in the results of TIMSS 1995 and 1999. Countries such as Oman, where girls out-perform boys in science, could initiate similar national projects to encourage boys to engage more with science learning and to improve their self-perceptions of their attitudes and capabilities.

The (male) science teachers in boys' schools will play a key role in any such campaigns and need to make a conscious effort to make science more engaging and exciting for their students. Indeed, all science teachers in the country should aim to create a learning environment that engages both boys and girls. Research has shown that teacher efficacy has an important impact on student attitudes towards the subject (Yoo, 2018) and, according to Desy et al. (2011), that teachers are more influential than parents in instilling positive attitudes towards science in students. These researchers call for more engaging, fun, exciting, and curiosity-capturing hands-on experiences, and curricula for all students as one possible way to reduce the gender gap. They also stress the need for more professional development opportunities for science teachers at all grade levels. Researchers investigating different national contexts agree that when students are engaged, there will be greater learning motivation and better achievement (Abu-Hilal et al., 2016; Dee, 2007; Pavešić, 2008; Tran et al., 2012). Perry et al. (2012) argued that students across the board need to be consistently encouraged to engage with science, should be expected to succeed in it, and must be provided with opportunities to carry out hands-on science-related activities. Other science educators have argued that the science curriculum needs to be reconstructed, with a greater emphasis on contextual learning; this, they say, could lead to more positive student attitudes towards science disciplines (Khishfe & BouJaoude, 2016; Oon et al., 2020).

Another suggestion for ways to improve boys' engagement in science is to increase the time devoted to experimentation and inquiry during science lessons. Pavešić (2008), for example, found that an increase of experimentation time in Slovenian science lessons was linked with higher achievement in the TIMSS 2003 science test, as well as with greater student enjoyment of science. This study also found that the engagement of girls in experimentation was also linked with enhanced achievement, and inquiry-based learning in general was associated with positive student attitudes towards science. More specifically, a higher frequency of inquiry-based instruction was significantly associated with improved self-efficacy and greater interest in learning science and mathematics, especially for boys. Other researchers also see inquiry-based learning as a primary "gender inclusionary strategy" because it responds to student interests and experiences, reveals the social relevance of science, promotes peer collaboration, and stresses the real-life context (Riegle-Crumb et al., 2019). It is clear, therefore, that Omani classrooms need to provide more opportunities for students of both genders to engage in inquiry-based instruction, while being aware that boys and girls might respond differently to these initiatives. Further research is still needed to explore the best instructional methods to help narrow the gender gap in the Omani context, to encourage boys to engage more enthusiastically in science learning, and to improve their self-perceptions of their own attitudes and capabilities and thus their level of achievement.

This current study has a number of limitations. One limitation lies in the fact that Oman has a single-gender school system; this might mean that the findings cannot easily be generalized to mixed-gender school systems. The claims made in this study therefore need to be bolstered and extended by further research that will explore what happens in mixed-gender school contexts. This limitation, however, also has a positive side, in that it contributes to our knowledge and pedagogical understanding of the impact of this unique learning environment on students' engagement in the classroom and on their performance. Another limitation of the study is that it looks only at grade 9 students (relating to the grade 8 TIMSS test); an exploration of the gender gap in science achievement in grade 4 students might produce different results. The findings of the study are also limited by the fact that it only investigates specific components of student self-perceptions of their own attitudes and capabilities in relation to the gender gap; these are: MA, SR, SLSE, and AS. Repeating the study using different cognitive and psychological factors might provide further understanding of the issues underlying the gender gap in science achievement. An exploration of the gender gap in mathematics might also shed further light on the topic and enhance our understanding of the gender gap in science achievement in Oman.

Author contributions: All authors have sufficiently contributed to the study and agreed with the results and conclusions.

Funding: This study is part of a research project (RC/EDU/CUTM/15/01) funded by the Research Council of the Ministry of Higher Education, Research and Innovation, Oman.

Acknowledgements: The authors would like to thank the Sas Center for Entrepreneurship (under the Ministry of Technology and Telecommunications) in designing the TOSS application.

Ethical statement: The authors stated that the Ethical Committee of Sultan Qaboos University approved the study proposal on March 26, 2015.

Declaration of interest: No conflict of interest is declared by authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Ababneh, E. G., & Abdel Samad, M. M. (2018). Gender gap in science achievement for Jordanian students in PISA 2015. *European Journal of Educational Research, 7*(4), 963-972. https://doi.org/10.12973/eu-jer.7.4.963

Abu-Hilal, M. (2005). Frames of reference model of self-concept and locus of control among a sample of Emirati school boys and girls. *Journal of Humanities and Social Sciences, 21*(2), 57-80.

Abu-Hilal, M., Aldhafri, S., Al-Kilani, H., Kazem, A., Al-Qaryouti, I., & Alkharusi, H. (2016). Factorial validity and reliability of teachers' self-efficacy among Omani teachers: Invariance across gender. *Journal of Educational & Psychological Sciences, 17*(1), 610-624. https://doi.org/10.12785/JEPS/170120

Akben, N. (2020). Effects of the problem-posing approach on students' problem solving skills and metacognitive awareness in science education. *Research in Science Education, 50*(3), 1143-1165. https://doi.org/10.1007/s11165-018-9726-7

Al -Harthy, I. (2016). Prediction accuracy: The role of feedback in 6th graders' recall predictions. *International Education Studies, 9*(3), 212-216. https://doi.org/10.5539/ies.v9n3p212

Al-Balushi, S., Al-Harthy, I. & Almehrizi, R. (2022). Attention drifting away while test-taking: Mind-wandering in students with low and high-performance levels in TIMSS-like science tests. *International Journal of Science and Mathematics Education*. https://doi.org/10.1007/s10763-022-10258-6.

Aldhafri, S., & Alrajhi, M. (2014). The predictive role of teaching styles on Omani students' mathematics motivation. *International Education Studies, 7*(6), 135-144. https://doi.org/10.5539/ies.v7n6p135

Al-Harthy, I., & Aldhafri, S. (2014). The relationship among task-value, self-efficacy and academic achievement in Omani students at Sultan Qaboos University. *International Review of Social Sciences and Humanities, 7*(2), 15-22.

Alkharusi, H. (2008). Effects of classroom assessment practices on students' achievement goals. *Educational Assessment, 13*(4), 243-266. https://doi.org/10.1080/10627190802602509

Al-Mutawah, M., & Fateel, M. (2018). Students' achievement in math and science: How grit and attitudes influence? *International Education Studies, 11*(2), 97-105. https://doi.org/10.5539/ies.v11n2p97

Alrajhi, M., Aldhafri, S., Alkharusi, H., Alharthy, I., Albarashdi, H., & Alhadabi, A. (2019). Grade and gender effects on self-concept development. *The*

Open Psychology Journal, 12, 66-75. https://doi.org/10.2174/1874350101912010066

Ambusaidi, A. & Al-Farei, K. (2017). Investigating Omani science teachers' attitudes towards teaching science: The role of gender and teaching experiences. *International Journal of Science and Mathematics Education, 15*, 71-88. https://doi.org/10.1007/s10763-015-9684-8

Askin, O., & Oz, E. (2020). Cross-national comparisons of students' science success based on gender variability: Evidence from TIMSS. *Journal of Baltic Science Education, 19*(2), 186-200. https://doi.org/10.33225/jbse/20.19.186

Aurah, C. M., Cassady, J. C., & McConnell, T. J. (2014). Genetics problem solving in high school testing in Kenya: Effects of metacognitive prompting during testing. *Electronic Journal of Science Education, 18*(8), 1-26.

Barry, A. (2019). Gender differences in academic achievement in Saudi Arabia: A wake-up call to educational leaders. *International Journal of Education Policy & Leadership, 15*(15), 1-17. https://doi.org/10.22230/ijepl.2019v15n15a890

Berger, N., Mackenzie, E., & Holmes, K. (2020). Positive attitudes towards mathematics and science are mutually beneficial for student achievement: a latent profile analysis of TIMSS 2015. *The Australian Educational Researcher, 47*, 409-444. https://doi.org/10.1007/s13384-020-00379-8

Bofah, E., & Hannula, M. (2015). TIMSS data in an African comparative perspective: Investigating the factors influencing achievement in mathematics and their psychometric properties. *Large-scale Assessments in Education, 3*, 4. https://doi.org/10.1186/s40536-015-0014-y

Brunner, M., Lüdtke, O., & Trautwein, U. (2008). The internal/external frame of reference model revisited: Incorporating general cognitive ability and general academic self-concept. *Multivariate Behavioral Research, 43*(1), 137-172. https://doi.org/10.1080/00273170701836737

Casselman, B. L., & Atwood, C. H. (2017). Improving general chemistry course performance through online homework-based metacognitive training. *Journal of Chemical Education, 94*(12), 1811-1821. https://doi.org/10.1021/acs.jchemed.7b00298

Chiu, M.-S. (2008) Achievements and self-concepts in a comparison of math and science: Exploring the internal/external frame of reference model across 28 countries. *Educational Research and Evaluation, 14*(3), 235-254. https://doi.org/10.1080/13803610802048858

Dee, T. S. (2006). The why chromosome: How a teacher's gender affects boys and girls. *Education Next, 6*(4), 68-75.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources, 42*(3), 528-554. https://doi.org/10.3368/jhr.XLII.3.528

Desy, E., Peterson, S., & Brockman, V. (2011). Gender differences in science-related attitudes and interests among middle school and high school students. *Science Educators, 20*(2), 23-30. https://doi.org/10.1037/e617182010-001

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist, 44*(2), 78-89.https://doi.org/10.1080/00461520902832368

Eisenkopf, G., Hessami, Z., Fischbacher, U., & Ursprung, H.W. (2015). Academic performance and single-sex schooling: Evidence from a natural experiment in Switzerland. *Journal of Economic Behavior & Organization, 115*, 123-143. https://doi.org/10.1016/j.jebo.2014.08.004

Geesa, R., Izci, B., Chen, S., & Song, H. (2020). The role of gender and attitudes toward science in fourth and eighth graders' science achievement in South Korea, Turkey, and the United States. *Journal of Research in Education, 29*(2), 54-87.

Gestsdottir, S., von Suchodoletz, A., Wanless, S. B., Hubert, B., Guimard, P., Birgisdottir, F., Gunzenhauser, C., & McClelland, M. (2014). Early behavioral self-regulation, academic achievement, and gender: Longitudinal findings from France, Germany, and Iceland. *Applied Developmental Science, 18*(2), 90-109. https://doi.org/10.1080/10888691.2014.894870

Gray, H., Lyth, A., McKenna, C., Stothard, S., Tymms, P. & Copping, L. (2019). Sex differences in variability across nations in reading, mathematics, and science: A meta-analytic extension of Baye and Monseur (2016). *Large-Scale Assessment Education, 7*(2), 1-29. https://doi.org/10.1186/s40536-019-0070-9

Hong, W., Bernacki, M. L., & Perera, H. N. (2020). A Latent profile analysis of undergraduates' achievement motivations and metacognitive behaviors, and their relations to achievement in science. *Journal of Educational Psychology, 112*(7), 1409-1430. https://doi.org/10.1037/edu0000445

Kablan, Z., & Kaya, S. (2013). Science achievement in TIMSS cognitive domains based on learning styles. *Eurasian Journal of Educational Research, 53*, 97-114. ttps://doi.org/10.14689/ejer.2013.53.6

Khishfe, R., & BouJaoude, S. (2016). Lebanese students' conceptions of and attitudes towards science and related careers based on their gender and religious affiliations. *International Journal of Science and Mathematics Education, 14*, 145-167. https://doi.org/10.1007/s10763-014-9587-0

Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. *International Journal of Science Education, 29*(7), 871-893. https://doi.org/10.1080/09500690600909091

Kingir, S., Tas, Y., Gok, G., & Vural, S. S. (2013). Relationships among constructivist learning environment perceptions, motivational beliefs, self-regulation and science achievement. *Research in Science & Technological Education, 31*(3), 205-226. https://doi.org/10.1080/02635143.2013.825594

Kurman, J. (2001). Self-regulation strategies in achievement settings: Culture and gender differences. *Journal of Cross-Cultural Psychology, 32*(4), 491-503. https://doi.org/10.1177/0022022101032004008

Larkin, S. (2006). Collaborative group work and individual development of metacognition in the early years. *Research in Science Education, 36*(1-2), 7-27. https://doi.org/10.1007/s11165-006-8147-1

Lee, J., Rhee, D.-E., & Rudolf, R. (2019). Teacher gender, student gender, and primary school achievement: Evidence from ten francophone African countries. *The Journal of Development Studies, 55*(4), 661-679. https://doi.org/10.1080/00220388.2018.1453604

Lin, T., & Tsai, C. (2013). A multi-dimensional instrument for evaluating Taiwanese high school students' science learning self-efficacy in relation to their approaches to learning science. *International Journal of Science and Mathematics Education, 11,* 1275-1301. https://doi.org/10.1007/s10763-012-9376-6

Liou, P.-Y., & Bulut, O. (2020). The effects of item format and cognitive domain on students' science performance in TIMSS 2011. *Research in Science Education, 50*(1), 99-121. https://doi.org/10.1007/s11165-017-9682-7

Marsh, H. (1986). Verbal and math self-concepts: An internal/external frame of reference model. *American Educational Research Journal, 23*, 129-149. https://doi.org/10.3102/00028312023001129

Marsh, H. W., & Hau, K.-T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology, 96*, 56-67. https://doi.org/10.1037/0022-0663.96.1.56

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2016). *Methods and procedures in TIMSS 2015.* TIMSS & PIRLS International Study Center.

Mata, M., Monteiro, V., & Peixoto, F. (2012). Attitudes towards mathematics: Effects of individual, motivational, and social support factors. *Child Development Research, 2012,* 1-10. https://doi.org/10.1155/2012/876028

Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology, 101*(3), 689-704. https://doi.org/10.1037/a0014240

Meinck, S., & Brese, F. (2019). Trends in gender gaps: Using 20 years of evidence from TIMSS. *Large-Scale Assessments in Education, 7*, 8. https://doi.org/10.1186/s40536-019-0076-3

Mejía-Rodríguez, A., Luyten, H., & Meelissen, M. (2020). Gender differences in mathematics self-concept across the world: An exploration of student and parent data of TIMSS 2015. *International Journal of Science and Mathematics Education, 19*, 1229-1250. https://doi.org/10.1007/s10763-020-10100-x

Ministry of Education and World Bank. (2012). Education in Oman: The drive for quality. *The Ministry of Education, Oman*. https://unesdoc.unesco.org/ark:/48223/pf0000227402

Mullis, I., & Martin, M. (2017). *TIMSS 2019 assessment framework*. The International Association for the Evaluation of Educational Achievement.

Mullis, I., Martin, M., Foy, D., Kelly, D., & Fishbein, B. (2020). *TIMSS 2019 International Results in Mathematics and Science.* The International Association for the Evaluation of Educational Achievement.

Naidoo, J., & Sibanda, D. (2020). Examining science performance of South African grade 9 learners in TIMSS 2015 through a gender lens. *South African Journal of Education, 40,* S1-10. https://doi.org/10.15700/saje.v40ns2a1717

National Center for Statistics and Information, Oman. (2020). *Statistical yearbook 2020.* https://www.ncsi.gov.om/Pages/NCSI.aspx

Ogunjuyigbe, P. O., Ojofeitimi, E. O., & Akinlo, A. (2006). Science education in Nigeria: An examination of peoples' perceptions about female participation in science, mathematics, and technology. *Journal of Science Education and Technology, 15*(3), 277-284. https://doi.org/10.1007/s10956-006-9014-6

Oon, P-T., Cheng, M. M. W., & Wong, A. S. L. (2020). Gender differences in attitude towards science: Methodology for prioritizing contributing factors. *International Journal of Science Education, 42*(1), 89-112. https://doi.org/10.1080/09500693.2019.1701217

Oyelekan, O. S., Jolayemi, S. S., & Upahi, J. E. (2019). Relationships among senior school students' self-efficacy, metacognition, and their achievement in chemistry. *Cypriot Journal of Educational Sciences, 14*(2), 208-221. https://doi.org/10.18844/cjes.v14i2.2564

Pavešić, B. (2008). Science achievement, gender differences, and experimental work in classes in Slovenia as evident in TIMSS studies. *Studies in Educational Evaluation, 34*(2), 94-104. https://doi.org/10.1016/j.stueduc.2008.04.005

Perry, B. L., Link, T., Boelter, C., & Leukefeld, C. (2012). Blinded to science: Gender differences in the effects of race, ethnicity, and socioeconomic status on academic and science attitudes among sixth graders. *Gender and Education, 24*(7), 725-743. https://doi.org/10.1080/09540253.2012.685702

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40. https://doi.org/10.1037/0022-0663.82.1.33

Riegle-Crumb, C., Morton, K., Nguyen, U., & Dasgupta, N. (2019). Inquiry-based instruction in science and mathematics in middle school classrooms: Examining its association with students' attitudes by gender and race/ethnicity. *American Educational Research Association Open, 5*(3), 1-17. https://doi.org/10.1177/2332858419867653

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*(4), 460-475. https://doi.org/10.1006/ceps.1994.1033

Schwery, D., Hulac, D., & Schweinle, A. (2016). Understanding the gender gap in mathematics achievement: The role of self-efficacy and stereotype threat. *School Psychology Forum, 10*(4), 386-396.

Smith, T. J., Pasero, S. L., & McKenna, C. M. (2014). Gender effects on student attitude toward science. *Bulletin of Science, Technology & Society, 34*(1-2), 7-12. https://doi.org/10.1177/0270467614542806

Thomson, S. (2008). Examining the evidence from TIMSS: Gender differences in year 8 science achievement in Australia. *Studies in Educational Evaluation, 34*(2), 73-81. https://doi.org/10.1016/j.stueduc.2008.04.003

Tran, N., Schneider, S., Duran, L., Conley, A., Richland, L., Burchinal, M., Rutherford, T., Kibrick, M., Osborne, K., Coulson, A., Antenore, F., Daniels, A., & Martinez, M. (2012). The effects of mathematics instruction using spatial temporal cognition on teacher efficacy and instructional practices. *Computers in Human Behavior, 28*(2), 340-349. https://doi.org/10.1016/j.chb.2011.10.003

Tsai, C.-C., Ho, H. N., Liang, J.-C., & Lin, H.-M. (2011). Scientific epistemic beliefs, conceptions of learning science and self-efficacy of learning science among high school students. *Learning and Instruction, 21*(6), 757-769. https://doi.org/10.1016/j.learninstruc.2011.05.002

Tweissi, A., Ababneh, I., & Abu Lebdih, K., (2014). Gender gap in student achievement in Jordan study report monitoring & evaluation partnership (MEP) project. *National Center for Human Resources Development and World Education.* https://www.researchgate.net/publication/281280418_Gender_Gap_in_Student_Achievement_in_Jordan_Study_Report

Vantieghem, W., Vermeersch, H., & Van Houtte, M. (2014). Transcending the gender dichotomy in educational gender gap research: The association between gender identity and academic self-efficacy. *Contemporary Educational Psychology, 39*(4), 369-378. https://doi.org/10.1016/j.cedpsych.2014.10.001

Velayutham, S., & Aldridge, J. M. (2013). Influence of psychosocial classroom environment on students' motivation and self-regulation in science learning: A structural equation modeling approach. *Research in Science Education, 43*(2), 507-527. https://doi.org/10.1007/s11165-011-9273-y

Wan, Z. H. (2021). Exploring the effects of intrinsic motive, utilitarian motive, and self-efficacy on students' science learning in the classroom using the expectancy-value theory. *Research in Science Education, 51*(3), 647-659. https://doi-org.uoelibrary.idm.oclc.org/10.1007/s11165-018-9811-y

Wang, J. R., & Chen, S. F. (2014). Exploring mediating effect of metacognitive awareness on comprehension of science texts through structural equation modeling analysis. *Journal of Research in Science Teaching, 51*(2), 175-191. https://doi.org/10.1002/tea.21131

Wang, J. R., Chen, S. F., Fang, I., & Chou, C-T. (2014). Comparison of Taiwanese and Canadian students' metacognitive awareness of science reading, text, and strategies. *International Journal of Science Education, 36*(4), 693-713. https://doi.org/10.1080/09500693.2013.826841

Williams, K. E., White, S. L., & MacDonald, A. (2016). Early mathematics achievement of boys and girls: Do differences in early self-regulation pathways explain later achievement? *Learning and Individual Differences, 51*, 199-209. https://doi.org/10.1016/j.lindif.2016.09.006

Yoo, Y. (2018). Modelling of factors influencing gender difference in mathematics achievement using TIMSS 2011 data for Singaporean eighth grade students. *Asia Pacific Journal of Education, 38*(1), 1-14. https://doi.org/10.1080/02188791.2017.1334626

Zheng, J., Xing, W., Zhu, G., Chen, G., Zhao, H., & Xie, C. (2020). Profiling self-regulation behaviors in STEM learning of engineering design. *Computers & Education, 143*, 103699. https://doi.org/10.1016/j.compedu.2019.103669