Study aimed at identifying factors with the potential to reduce the disparities in mathematics anxiety observed between 15-year-old French-speaking boys and girls in Quebec, based on an analysis of PISA data from 2003 and 2012[1]

Étude visant à identifier des facteurs ayant le potentiel de réduire les écarts d'anxiété mathématique observés entre les garçons et les filles francophones de 15 ans du Québec, à partir d'une analyse des données du PISA de 2003 et 2012

Estudo visando identificar fatores com o potencial de reduzir as disparidades na ansiedade matemática observadas entre meninos e meninas francófonos de 15 anos do Quebec, a partir de uma análise dos dados do PISA de 2003 e 2012

Patricia Vohl
ID ORCID : 0000-0001-8279-9730
Université Laval

Nathalie Loye
ID ORCID : 0000-0002-0809-6911
Université de Montréal

1. The French version was published in issue 46(3) 2023: https://doi.org/10.7202/1113335ar

Keywords: mathematics anxiety, mathematics self-concept, intrinsic value of mathematics, utility value of mathematics, PISA

A previous study (Vohl & Loye, 2023) showed that, on average, 15-year-old French-speaking Quebec girls are more anxious about mathematics than French-speaking Quebec boys across the performance continuum. The results also showed that performances in mathematics and mathematics anxiety are two negatively correlated phenomena. This paper aims to identify factors that could explain the differences in mathematics anxiety observed between girls and boys and that may have the potential to reduce the observed differences. The Pekrun's control-value model for achievement emotions (2006) was used to verify whether differences in self-concept, intrinsic value, and utility value observed between girls and boys completely explain the differences in mathematical anxiety. The results in self-concept explain nearly 70% of the differences in mathematics anxiety observed among francophone students in Quebec.

Mots-clés : anxiété mathématique, concept de soi en mathématiques, valeur intrinsèque des mathématiques, valeur utilitaire des mathématiques, PISA

Dans une étude antérieure (Vohl & Loye, 2023), nous avons montré qu'en moyenne, les filles québécoises francophones de 15 ans se disent davantage anxieuses à l'égard des mathématiques que les garçons et ce, sur l'ensemble du continuum des performances. Nous avons également montré que performances en mathématique et anxiété mathématique sont deux phénomènes négativement corrélés. Dans le présent article, nous souhaitons identifier des facteurs qui pourraient permettre d'expliquer les écarts d'anxiété mathématique observés entre les filles et les garçons. En prenant appui sur le modèle du contrôle et de la valeur de Pekrun (2006), nous vérifions si les écarts de concept de soi, de valeur intrinsèque et de valeur utilitaire observés entre les filles et les garçons expliquent complètement les écarts d'anxiété mathématique. Nos résultats révèlent que les écarts de concept de soi expliquent près de 70% des écarts d'anxiété mathématique relevés chez les élèves francophones du Québec.

Authors' note: Correspondence regarding to this article may be addressed to patricia.vohl@fse.ulaval.ca.

Palavras-chaves: ansiedade matemática, autoconceito em matemática, valor intrínseco da matemática, valor utilitário da matemática, PISA

Num estudo anterior (Vohl & Loye, 2024), mostramos que, em média, as raparigas francófonas do Quebec de 15 anos se sentem mais ansiosas em relação à matemática do que os rapazes e em todo o contínuo das performances. Mostramos também que o desempenho em matemática e a ansiedade matemática são dois fenómenos negativamente correlacionados. No presente artigo, pretendemos identificar fatores que possam explicar as diferenças de ansiedade matemática observadas entre raparigas e rapazes. Baseando-nos no modelo de controlo e valor de Pekrun (2006), procuramos verificar se as diferenças de autoconceito, valor intrínseco e valor utilitário observadas entre raparigas e rapazes explicam completamente as diferenças de ansiedade matemática. Os nossos resultados revelam que as diferenças de autoconceito explicam cerca de 70% das diferenças de ansiedade matemática observadas entre os alunos francófonos do Quebec.

Introduction

Mathematics is one of the foundations of our highly technological society. As a result, performing well in the field is a major asset. Longitudinal studies show that performance in mathematics is positively linked to the probability of obtaining a high school diploma, making a successful transition to higher education, and obtaining a post-secondary diploma (Chiu & Klassen, 2010; Ma, 1999; OECD, 2014a; Parsons & Bynner, 2005; Stokke, 2015). Other studies show that performance in mathematics predicts employability and potential professional income (e.g., Fonseca et al., 2021—for empirical results from Quebec; Joensen & Nielsen, 2009; Ma, 1999; Parsons & Bynner, 2005).

In this context, education systems aim to ensure that every learner performs to the best of his or her ability in mathematics (OECD, 2013). Fortunately, the results of major international education surveys, such as TIMSS[2] and PISA[3], show that the Quebec system meet's this objective with relative success. For example, 15-year-old Quebec students have

2. TIMSS for Trends in International Mathematics and Science Study, one of the international studies established by the IEA (International Association for the Evaluation of Educational Achievement).
3. PISA for Programme for International Student Assessment, launched in 2000 by the OECD (Organisation for Economic Co-operation and Development).

ranked among the best in the world in every cycle of the PISA survey. Furthermore, in every case, they have topped the list of Canadian provinces, with the exception of 2003, when Alberta and British Columbia ranked higher than Quebec (Bussière et al., 2004). However, gender-based analyses reveal disparities that merit particular attention, in view of the issues mentioned above. Indeed, several times since the early 2000s, the results of TIMSS and PISA have shown, statistically significant differences between the boys' and girls' average scores, to the advantage of boys. The results of PISA 2003, 2012, and 2018 also revealed that, all things considered, fewer Quebec girls than boys achieved top rankings[4] (Brochu et al., 2013; Mullis et al., 2016; OECD, 2014a; OECD, 2016; O'Grady et al., 2016, 2019).

Concerned by these gender-based performance gaps, our previous study (Vohl & Loye, 2023) aimed to verify whether mathematics anxiety, a concept defined as "feelings of tension and anxiety that interfere with the manipulation of numbers and the solving of mathematical problems in a wide variety of ordinary life and academic situations" (Richardson & Suinn, 1972, p. 551), could contribute to observed differences between the boys' and girls' performance in mathematics, as stipulated by Stoet et al. (2016). Our findings support this hypothesis.

Indeed, through secondary analysis of data from 15-year-old French-speaking[5] Quebec students who participated in PISA 2003 and 2012, we demonstrated the following two elements in particular: (a) on average, 15-year-old French-speaking Quebec girls across the entire performance continuum say they are more anxious about mathematics than French-speaking Quebec boys, and (b) performance in mathematics and mathematics anxiety are two negatively correlated phenomena. Thus, French-speaking Quebec students who say they are more anxious about mathematics tend to do less well in mathematics, on average, than those who say they are less anxious. These two elements are also consistent

4. For example, in 2003, 25.9% of Quebec boys scored at PISA levels 5 and 6, the highest proficiency levels, while only 19.3% of girls achieved this ($p < 0.05$). In PISA 2012, 25.3% of boys scored at levels 5 and 6, compared to 19.5% of girls ($p < 0.05$). Similar differences were also observed in PISA 2015 and 2018.
5. We have given priority to French-speaking Quebec students because this group constitutes the majority in Quebec. This choice is also justified in Vohl (2023) and Vohl and Loye (2023). Note that, in this paper, as in Vohl and Loye (2023), the term "French-speaking Quebec students" refers to Quebec students from schools in the French-language school system that participated in PISA 2003 and 2012.

with the general trends noted by PISA 2003 and 2012 regarding the links between gender, mathematics anxiety, and performance in mathematics. In 2003 and 2012, in all participating OECD countries, girls said they were more anxious about mathematics than boys, on average (OECD, 2013). Also, analyses conducted by the OECD revealed correlations of -0.37 and -0.39 between mathematics anxiety and performance in mathematics , in OECD countries in 2003 and 2012 respectively (OECD, 2013).

In this paper, we pursue the work begun by Vohl and Loye (2023). More specifically, we aim to identify factors that might explain the observed discrepancies in mathematics anxiety between French-speaking boys and girls in Quebec, with a view to identifying and implementing intervention strategies to reduce these discrepancies. This paper seeks to answer the question: What factors help to explain the gender-based disparity in mathematics anxiety observed among 15-year-old French-speaking Quebec students who participated in PISA 2003 and 2012 and what potential do they hold for reducing the disparities?

Theoretical framework

We present three hypotheses that have been proposed to explain the negative link between mathematics anxiety and performance in mathematics since the early 2000s, as they reflect the factors put forward to explain gender differences in mathematics anxiety. We then present Pekrun's control–value theoretical model for achievement emotions (2006), which integrates the three hypotheses to explain the negative link between mathematics anxiety and performance in mathematics, proposes an explanation for the gender disparity in mathematics anxiety, and then integrates all other factors suggested to date to explain the gender disparity in mathematics anxiety.

Three hypotheses to explain the negative link between mathematics anxiety and performance

The first hypothesis to explain the negative link between mathematics anxiety and performance in mathematics was proposed by Ashcraft and Kirk (2001). This is the disruption account (Ramirez et al., 2018, p. 146). According to Ashcraft and Kirk (2001), mathematics anxiety (the cause) harms performance (the effect). They explain that this is caused by working memory, a cognitive system linked to short-term memory, responsible for managing the information necessary for solving mathematical tasks (inhibition of irrelevant information and storage of relevant information) (Baddeley, 1992; Baddeley & Hitch, 1974; Beilock & Carr, 2005; Geary & Widaman, 1992; Miyake & Shah, 1999; Raghubar et al., 2010). Indeed, mathematics anxiety triggers rumination and negative thoughts that occupy a portion of the limited resources of working memory, which limits mathematics performance.

In addition, the unpleasant physical manifestations of the phenomenon (increased heart rate, sweating, trembling, etc.) cause people who say they are anxious about mathematics to avoid mathematical situations. For them, avoiding mathematics is as natural as trying to avoid pain (OECD, 2013) but this also deprives them of learning opportunities. Therefore, according to Ashcraft and Kirk (2001), avoidance also helps to explain the negative link between mathematics anxiety and performance in mathematics.

The second hypothesis to explain the negative link between mathematics anxiety and performance was proposed by Maloney et al., among others (Ferguson et al., 2015; Maloney, 2016; Maloney et al., 2010; Maloney et al., 2011; Maloney et al., 2012). This is the reduced competency account (Ramirez et al., 2018, p. 146). According to these authors, mathematics anxiety stems from a deficit in three basic mathematical skills: the ability to count objects, the ability to order numerically, and spatial processing ability. As these skills are considered essential to developing high-level mathematical skills, a deficit in these skills would lead to poorer performance in mathematics (the cause) from childhood onwards, triggering mathematics anxiety (the effect).

The third hypothesis to explain the negative link between mathematics anxiety and performance in mathematics was proposed by Ramirez et al. (2018). This is the interpretation account (Ramirez et al., 2018, p. 151). This follows from the appraisal theory for the genesis of emotions (Arnold, 1950; Barrett, 2006; Lazarus, 1991; Schacter & Singer, 1962, cited in Ramirez et al., 2018, p. 151) and the attitude-as-constructions view (Bem, 1972; Chaiken & Yates, 1985; Wilson et al., 2000, cited in Ramirez et al., 2018, p. 151). These two theories define emotions/attitudes as the result of an interpretation of events experienced. Thus, according to the interpretation account, how individuals interpret their past and present learning experiences in mathematics and the results of these learning experiences causes mathematics anxiety more than the experiences and results themselves. Negative, inappropriate, or unrealistic interpretations may lead to mathematics anxiety.

The majority of the empirical results offered by Ramirez et al. (2018) to support the interpretation account point to links between mathematics anxiety, performance in mathematics, and the learner's perceptions, such as the perception of his or her mathematical competency, operationalized in particular by Bandura's sense of self-efficacy (1977, 1997, 2003) and self-concept in mathematics (e.g., Marsh et al., 1988; Marsh & Hau, 2004; Marsh & Yeung, 1998; Martin & Marsh, 2008), as well as the perception of the value of mathematics, operationalized in particular by the intrinsic value and utility value of mathematics (Ryan & Deci, 2009; Viau, 2009; Wigfield & Eccles, 2000). The concepts of sense of self-efficacy, self-concept in mathematics, intrinsic value, and utility value are defined below.

For Bandura, the concept of perceived self-efficacy refers to "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3). The concept is seen as "a powerful incentive to act or to persevere in the face of difficulties" (OECD, 2013, p. 89). According to Bandura (1997), four sources influence a person's self-efficacy beliefs: "enactive mastery experiences that serve as indicators of capability; vicarious experiences that alter efficacy beliefs through transmission of competencies and comparison with the attainments of others; verbal persuasion and allied types of social influences" (Bandura, 1997, p. 79), and lastly, psychological and affective states.

Marsh et al. (1988) and Shavelson et al. (1976) describe self-concept as one's belief in one's own abilities (OECD, 2013). Like the sense of self-efficacy, self-concept is strongly correlated with performance in a given field (Marsh et al., 1985a, 1985c; Marsh & Shavelson, 1985b). The main sources of the phenomenon are evaluations provided by significant others (parents, teachers), causal attributions (how people explain their successes/failures), comparison with peers, and finally, comparison of performance among various fields of study (Marsh & Scalas, 2011).

Intrinsic value and utility value both refer to the judgment of the value of a field or an activity within a given field (Viau, 2009; Wigfield & Eccles, 2000). Intrinsic value refers to the pleasure a person derives from accomplishing a task (Berger, 2015; Viau, 2009; Wigfield & Eccles, 2000) or to "the drive to perform an activity purely for the joy gained from the activity itself" (OECD, 2013, p. 73) Utility value refers to the benefits a person derives from performing an activity (Viau, 2009) and its function for (Berger, 2015, p. 71) the person's academic, professional and/or personal life (Wigfield & Eccles, 2000). Intrinsic value is similar to the concept of intrinsic motivation in self-determination theory (Ryan & Deci, 2009) and to that of interest (Hidi & Renninger, 2006). Utility value is similar to the concept of extrinsic motivation in Ryan and Deci's self-determination theory (2009).

Pekrun's control-value model and the three hypotheses

Pekrun's control-value model for achievement emotions (2006) groups mathematics anxiety with 14 other emotions typical of the academic context. These 15 emotions, known as achievement emotions, can be divided into three categories: (a) emotions experience during a learning activity underway (enjoyment, anger, frustration, boredom); (b) emotions about prospective outcomes of a learning activity (anxiety, anticipatory joy, hope, hopelessness, anticipatory relief); and, (c) emotions about retrospective outcomes of a learning activity (joy, pride, gratitude, sadness, shame, anger) (Pekrun, 2006).

The basic premise of Pekrun's control-value model for achievement emotions is as follows: achievement emotions arise, first and foremost, from a cognitive appraisal of control and value – the control the person believes he or she can exert over the activity and its outcomes, and the value of the activity and its outcomes. According to Pekrun, self-concept attests to the cognitive appraisal of control, while intrinsic value and utility value attest to the cognitive appraisal of value. With this basic premise, a parallel can be drawn with the interpretation account, while the emotion of interest is mathematics anxiety, albeit with a focus on self-concept alone rather than both the sense of self-efficacy and self-concept, in relation to the perception of competency.

Pekrun's control-value model is composed of four main spheres which interact with each another: (a) the learning environment, (b) the cognitive appraisal of control and value, (c) achievement emotions, and lastly, (d) learning/achievement/performance where learning refers to cognitive resources, strategies employed, and engagement/avoidance, while performance relates to the outcomes of learning assessment. In addition to these four spheres, there are individual affective predispositions (e.g., genes and temperament) and individual cognitive predispositions (e.g., intelligence and competences).

Pekrun proposes a series of links between the various spheres of the model (see Figure 1):

- cognitive appraisal of control and value gives rise to achievement emotions – the basic postulate of the model (link 1) (which allows, with links 2 and 3, to draw a parallel with the interpretation account);
- learning/achievement/performance influence the cognitive appraisal of control and value (link 2);
- the learning environment, including interaction between the learner and the learner's peers, teachers, and parents, influences the cognitive appraisal of control and value (link 3);
- achievement emotions act on cognitive resources (including working memory), on the choice of learning strategies (flexible or rigid), on engagement/avoidance, which has an impact on achievement/performance (link 4) (which allows a parallel to be drawn with the disruption account);
- individual cognitive predispositions (e.g., intelligence and competences) influence learning/achievement/performance (link 5) (which, in combination with link 7, allows for a parallel to be drawn with the reduced competency account);
- individual affective predispositions (e.g., genes and temperament) influence achievement emotions (link 6);
- learning/achievement/performance influence emotions (link 7), and finally;
- learning/achievement/performance influence the environment (link 8).

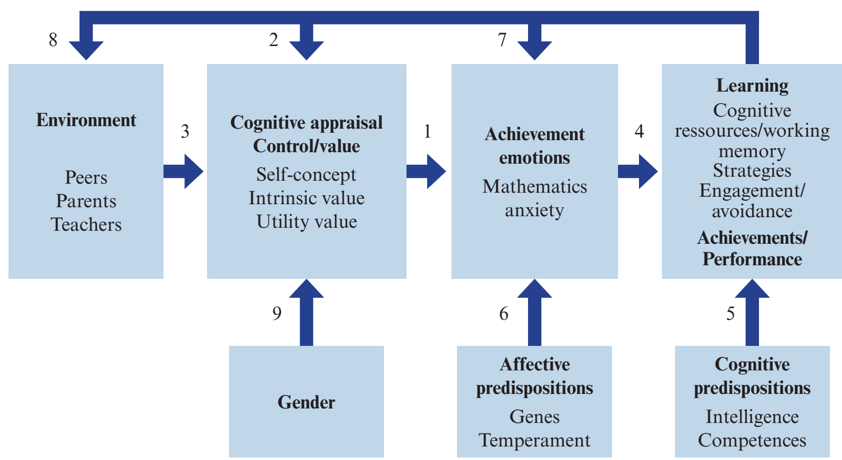
Building on the work of Meece et al. (1990) and of Seegers and Boekaerts (1996), Pekrun assumes that all of the above links are gender- and culture-invariant. In other words, according to Pekrun, there is no expected moderation effect of the proposed links by gender and by culture.

The track potentially explaining the disparity in mathematics anxiety between boys and girls, according to Pekrun's control–value model

According to Pekrun's control–value model for achievement emotions, gender directly influences the cognitive appraisal of control and value (link 9, Figure 1) and, in turn, the cognitive appraisal gives rise to achievement emotions: mathematics anxiety, in this case. Thus, the relationship between gender and mathematics anxiety are mediated by the learner's cognitive appraisal, operationalized by self-concept, intrinsic value, and utility value.

As the links between self-concept, intrinsic value, utility value, and mathematics anxiety are assumed to be gender-invariant, if mediation is complete, the disparity in mathematics anxiety between boys and girls would be explained by the gender disparity in cognitive appraisal: lower mathematics self-concept, intrinsic value of mathematics and utility value attributed to mathematics among girls than among boys (Frenzel et al., 2007).

Figure 1
Pekrun's control–value model for achievement emotions (2006)



Note. Model adapted from Pekrun (2006, p. 328).

In order to empirically verify this explanatory track, it is necessary to proceed in three steps:

- Verify whether there is a gender difference in the mean level of self-concept, intrinsic value, and utility value.
- Verify whether the relationship between gender and mathematics anxiety is completely mediated by self-concept, intrinsic value, and utility value.
- Verify whether the links between self-concept, intrinsic value, utility

value, and mathematics anxiety are gender-invariant. These links are expected to be negative between both self-concept and intrinsic value with respect to mathematics anxiety, but positive between utility value and mathematics anxiety. Thus, a lower self-concept would be associated with a higher level of mathematics anxiety, a lower intrinsic value would be associated with a higher level of mathematics anxiety, and a lower utility value would be associated with a lower level of mathematics anxiety (Pekrun, 2006, cited in Henschel & Roick, 2017).

Once these three steps are verified, it can be asserted that the disparity in mathematics anxiety between boys and girls is explained by a gender-based disparity in self-concept, intrinsic value, and utility value. Below, we report empirical findings that support each of these steps.

Supporting the first step, several research studies (Frenzel et al., 2007; Goetz et al., 2008; Goetz et al., 2013; Kyttälä & Björn, 2010, Meece et al., 1990; Mullis et al., 2000), including meta-analyses (Else-Quest et al., 2010; Hyde et al, 1990), have revealed discrepancies between boys and girls for self-concept, intrinsic value, and/or utility value for mathematics. The meta-analysis of Hyde et al. (1990) found that girls tend to report a lower mathematics self-concept than boys (with an effect size of 0.25). Subsequently, the meta-analysis of Else-Quest et al. (2010) found that girls tend to report a lower self-concept, intrinsic value of mathematics, and utility value attributed to mathematics than boys (with respective effect sizes of 0.33, 0.20, 0.24).

PISA 2003 and 2012 data also identified gender gaps in self-concept, intrinsic value, and utility value. Indeed, on average, 15-year-old girls who participated in PISA 2003 and 2012 reported statistically lower self-concept, intrinsic value, and utility value than boys in OECD countries ($p < 0.05$). The same was true across Canada. The differences in average indices observed in Canada between boys and girls (B–G) for the three concepts were, respectively, 0.36, 0.18, and 0.13 in 2003 and 0.39, 0.23, and 0.13 in 2012 (indices calibrated so that the mean is 0 and the standard deviation is 1, in OECD countries) (OECD, 2005a, 2013).

Controlling for previous performance, the study by Frenzel et al. (2007), conducted among 2,053 German students with an average age of 11, found statistically significant gender differences for all three perceptions (with effect sizes of 0.49, 0.31, and 0.02 for self-concept, intrinsic value, and utility value, respectively). Other studies that specifically document gender differences in self-concept have shown that, for equal performance, on average, girls report a lower mathematics self-concept than boys (Correll, 2001; Goldman & Penner, 2016; Mejía-Rodríguez et al., 2021).

Supporting the second step, the study by Frenzel et al. (2007) shows that self-concept, intrinsic value, and utility value completely mediate the relationship between gender and mathematics anxiety. The authors proceeded as proposed by Kenny et al. (1998, cited in Frenzel et al., 2007) to show that by adding the three perceptions as independent variables, the regression coefficient of the independent gender variable changes from statistically significant to nonsignificant in a multiple regression model where mathematics anxiety is defined as the dependent variable ($b = 0.13$, $p < 0.01$, for the gender variable before the inclusion of the three perceptions and $b = 0$, after inclusion of the three perceptions). Another study, by Kyttälä and Björn (2010), conducted over two years with 116 Finnish students aged 13 and 14, showed a complete mediation of the relationship between gender and mathematics anxiety via self-concept. To do this, the authors studied the direct and indirect effects between the variables of gender and mathematics anxiety by structural link analyses. They obtained a nonsignificant regression coefficient for the direct effect of the gender variable on mathematics anxiety (standardized regression coefficient of -0.07, $p = 0.350$) and a statistically significant regression coefficient for the indirect effect of gender on mathematics anxiety via self-concept (statistically significant standardized coefficient of 0.14, $p < 0.05$). However, this study did not find a significant indirect link between gender and mathematics anxiety via utility value.

Supporting the third step, the study by Frenzel et al. (2007) used multigroup analyses to show that the structural links between mathematics self-concept, intrinsic value, and utility value (independent variables) and mathematics anxiety (dependent variable) can be considered gender-invariant, while controlling for prior performance. To arrive at this conclusion, they compared the chi-square fit index of the unconstrained multigroup model (model in which the regression parameters are free to vary, between the girl group and the boy group) and that of the constrained multigroup model (model in which the parameters are forced to be equal between the two groups). As there was no statistically significant change in the chi-square fit index from the unconstrained model to the constrained model, they concluded that the links can be considered gender-invariant ($\Delta\chi^2 = 4{,}93$, $ddl = 4$, $p = 0{,}29$). The study also reported a negative link between mathematics self-concept and mathematics anxiety, a negative link between intrinsic value and mathematics anxiety, and a positive link between utility value and mathematics anxiety. The study by Kyttälä and Björn (2010) also used multigroup analyses to show that the structural links between self-concept, utility value (independent variables), and mathematics anxiety (dependent variable) can be considered gender-invariant. The link between self-concept and mathematics anxiety link was found to be negative, and the link between utility value and mathematics anxiety was found to be positive.

Finally, as the study by Frenzel et al. (2007) verified all three steps, the authors concluded that the differences in mathematics anxiety observed between the girls and boys in their study are explained by the differences in perception between girls and boys: a lower mathematics self-concept, a lower intrinsic value of mathematics, and a lower utility value attributed to mathematics among girls.

Our study aimed to verify whether the disparity observed between the mathematics anxiety of 15-year-old French-speaking Quebec girls and boys who participated in PISA 2003 and 2012 can be explained by the gender disparity in cognitive appraisal: a lower self-concept, intrinsic value, and utility value among girls than among boys. To this end, we defined three specific research objectives:

SO1: Verify whether there is a gender disparity in self-concept, intrinsic value, and utility value among 15-year-old French-speaking Quebec students who participated in PISA 2003 and 2012, controlling for performance in mathematics.

Research hypothesis for SO1: French-speaking Quebec girls report lower levels of self-concept, intrinsic value, and utility value than French-speaking Quebec boys, controlling for performance in mathematics.

SO2: Verify whether the relationship between gender and mathematics anxiety is mediated completely by self-concept, intrinsic value, and utility value, among 15-year-old French-speaking Quebec students who participated in PISA 2003 and 2012, controlling for performance.

Research hypothesis for SO2: Self-concept, intrinsic value, and utility value completely mediate the relationship between gender and mathematics anxiety, controlling for performance.

SO3: Verify whether the structural links between self-concept, intrinsic value, and utility value (independent variables) and mathematics anxiety (dependent variable) are gender-invariant among French-speaking Quebec students who participated in PISA 2003 and 2012.

Research hypothesis for SO3: The structural links are gender-invariant. The links between self-concept and mathematics anxiety and between intrinsic value and mathematics anxiety are negative, while the link between utility value and mathematics anxiety is positive, controlling for performance.

Taking SO1, SO2 and SO3 as a whole and the resulting research hypothesis: If, for the three perceptions considered, girls have a lower average level than boys, if the link between gender and mathematics anxiety is mediated completely by the three perceptions, and if the links between perceptions and mathematics anxiety are gender-invariant, we conclude that the disparity in mathematics anxiety observed between 15-year-old French-speaking Quebec girls and boys who participated in PISA 2003 and 2012 are explained by lower self-concept, intrinsic value, and utility value among girls than among boys. Controlling for performance will ensure that the observed disparities are not simply attributable to differences in performance between boys and girls. The research hypothesis that arises from taking SO1, SO2, and SO3 as a whole is as follows: The differences in mathematics anxiety between girls and boys can be explained by a lower self-concept, intrinsic value, and utility value among girls than among boys.

Methodology

The secondary data used for this study, taken from PISA 2003 and PISA 2012, have three characteristics to be considered for their analysis. In Vohl and Loye (2024), we present these three characteristics in detail, as well as the analysis techniques required to process them appropriately. We refer the reader to the paper by Vohl and Loye (2024) for further details. However, for clarity in this paper, we provide a brief description of the three particular characteristics inherent in PISA data and the required analysis techniques to be used.

The first methodological consideration inherent in PISA data

PISA samples are generated in each of the participating countries using complex sampling design (e.g., Lohr, 2019; Rutkovski et al., 2010; Skinner & Wakefield, 2017; Stapleton, 2013). PISA's complex sampling design is a two-stage stratified random design. In the first stage schools are selected, and in the second stage students are selected from each of the schools. In this sampling design, the probabilities of selecting students may be unequal. This explains why the PISA databases include a weighting for each student, known as the survey weight. Survey weights are also used to adjust for non-response. They must be incorporated into calculations when estimating parameters, such as the mean of a group.

Furthermore, in a two-stage stratified random design, as students are nested within schools, student data cannot be considered independent. For this reason, PISA provides replicated survey weights for each student, enabling the variance of the estimated parameters to be estimated using a resampling method, without the analyst having to carry out the resampling themselves. As a result, when conducting a secondary analysis of PISA data, survey weights and replicated survey weights must be used, unless multilevel modeling is chosen (see Vohl & Loye, 2024). Survey weights and replicated survey weights are integrated into the analysis plan presented below.

The second methodological consideration inherent in data taken from PISA

In the PISA 2003 and 2012 tests, participants were asked to complete two types of questionnaires: cognitive tests and a contextual questionnaire (OECD, 2005b, 2014b). The 120-minute paper-and-pencil cognitive tests were used to assess mathematical literacy, scientific literacy, and reading comprehension. The 30-minute paper-and-pencil contextual questionnaire was used to document various non-cognitive factors related to performance in the three domains assessed. As mathematical literacy was identified as a major area of assessment in both cycles, mainly non-cognitive factors relating to performance in mathematics were documented. As a result, mathematics anxiety, mathematics self-concept, intrinsic value, and utility value were documented in these cycles.

In 2012, PISA employed a method of item rotation known as three-form design (Graham et al., 1996) when developing the contextual questionnaires: three forms of booklet were produced (forms A, B, and C), and then each non-cognitive factor was included in only two of the three forms. As each student responded to only one of the three forms (form A, B, or C), given at random, the PISA 2012 database is missing at least one-third of data regarding mathematics anxiety, mathematics self-concept, intrinsic value, and utility value. Hence, a suitable method must be incorporated into the analysis plan to adequately deal with these missing data. The state of the art in this context is to use the full information maximum likelihood (FIML) method (Enders, 2010). This method is integrated into the data analysis plan presented below.

The third methodological consideration inherent in data taken from PISA

PISA 2003 and 2012 reported performance with five estimates of students' skill level based on responses to cognitive test items and on the difficulty level of the items, rather than scores. These five estimates are called plausible values. The plausible values approach has been used by PISA since 2000 to account for students' level of ability. For any analysis involving level of ability, the five plausible values must be combined using the Little and Rubin approach (2002). This method is integrated into the data analysis plan presented below.

The samples

The two samples analyzed in this study are: (a) the subsample of Quebec students from schools where French was the main language of instruction during PISA 2003, taken from the Canadian sample from PISA 2003; and (b) the subsample of Quebec students from schools where French was the main language of instruction during PISA 2012, taken from the Canadian sample from PISA 2012.

The 2003 subsample of French-speaking Quebec participants consists of 2,151 15-year-old students, split between 1,102 girls and 1,049 boys (51.2% and 48.8%, respectively, with or without the use of survey weights) from 119 schools. The 2012 subsample contains 2,385 15-year-old students, split between 1,251 girls and 1,134 boys (52.5% and 47.5%, respectively, without survey weights, and 50.9% and 49.1% when the data are weighted using survey weights), from 109 schools.

Ethical aspects

In the PISA 2003 and 2012 surveys, written parental consent was required for students to participate in the tests (in all countries and economies where current regulations required such consent) (OECD, 2005b, 2014b). Furthermore, in order to conduct this research, a request for ethical certification was submitted to the Université de Montréal's Comité d'éthique de la recherche en éducation et en psychologie (CEREP). The request was accepted, and this research is considered low risk for participants.

Variables studied and data analyzed

To meet the specific objectives of our research, six variables were used: gender, performance in mathematics, and the four non-cognitive variables of mathematics self-concept, the intrinsic value of mathematics, the utility value attributed to mathematics, and mathematics anxiety.

Gender

In 2003 and 2012, PISA documented participants' gender using a dichotomous item[6].

Performance in mathematics

Since 2003, PISA has analyzed a mathematics skill known as mathematical literacy. This skill is defined as "an individual's capacity to formulate, employ and interpret mathematics in a variety of contexts. It includes reasoning mathematically and using mathematical concepts, procedures, facts, and tools to describe, explain and predict phenomena. It assists individuals in recognizing the role that mathematics plays in the world and to make the well-founded judgements and decisions needed by constructive, engaged and reflective citizens" (OECD, 2014a, p. 37).

To account for each participant's level of mathematical literacy, PISA provides five plausible values[7]. In 2000, the OECD calibrated each of the plausible values so that the mean in OECD countries was 500 and the standard deviation was 100.

The four non-cognitive variables

On the basis of theoretical and empirical consideration, four measurement instruments were developed for PISA 2003 and reapplied in 2012 to understand mathematics self-concept, the intrinsic value of mathematics, the utility value attributed to mathematics, and mathematics anxiety (OECD, 2005b, 2013, 2014b).

Based on the participants' responses in 2003 and 2012, PISA calculated an index for each of the four instruments. The partial credit model (Masters &t Wright, 1997, cited in OECD, 2014b), a generalization of the one-parameter Rasch model (Rasch, 1960), was used to scale each instrument item, then the index "was scaled using a weighted likelihood estimate (WLE) (Warm, 1989)" (OECD, 2013, p. 193). These indices were calculated and calibrated so that the mean across OECD countries is 0 and the standard deviation is 1.

6. Item ST01Q03 in 2003 and item ST01Q04 in 2012.
7. In the PISA 2003 and 2012 databases, the five values are identified by the variables PV1MATH, PV2MATH, PV3MATH, PV4MATH, and PV5MATH.

For each of the non-cognitive variables studied in this research, Table 1 presents the name and definition of the index proposed by PISA, the wording of each instrument item, the psychometric qualities of the instruments as documented by the OECD for Canada during PISA 2003 and 2012, and finally, the psychometric qualities of the four instruments as observed by this study for the two samples concerned. The fit indices from the confirmatory factor analysis (CFA) conducted, both by the OECD and in this study, indicate an adequate fit between the theoretical model underlying the instruments and the empirical results, except for the RMSEA indices, which, in some cases, show a poor fit as they are greater than 0.1 (Byrne, 2012). As the majority of the indices show a good fit, the high RMSEA values do not compromise the use of the data. In addition, all fidelity indices show good internal consistency.

As our first specific research objective is concerned with gender differences in self-concept, intrinsic value, and utility value in mathematics, we conducted multigroup confirmatory factor analyses (Bollen, 1989; Hayduk, 1987; Jöreskog, 1971; Sörbom, 1974) to verify the metric invariance (equality of factor loading coefficients, between groups, for each item of the measurement instrument) and scalar invariance (equality of intercepts between groups for each item) (Meredith, 1993; Widaman & Reise, 1997) of these three instruments, on the 2003 and 2012 samples.

Indeed, few studies to date have verified the metric and scalar invariance of self-concept, intrinsic value, and utility value measurement instruments according to gender before comparing the observed means between boys and girls, yet it is strongly recommended before drawing conclusions when means are compared between groups. (Meredith, 1993; Steenkamp & Baumgartner, 1998; Steinmetz, 2011; Vandenberg & Lance, 2000). This verification is necessary before concluding that the observed mean deviations do indeed reflect mean deviations of the underlying trait of interest, as they could also reflect metric non-invariance and/or scalar non-invariance across groups (see Steinmetz, 2011, for a detailed explanation).

Table 1

Instruments for measuring mathematics self-concept, the intrinsic value of mathematics, the utility value of mathematics, and mathematics anxiety, as proposed by PISA in 2003 and 2012

Non-cognitive variable considered in this study	**Name and definition of the index as proposed by PISA 2003 and 2012**	**Contextual questionnaire items**	**Psychometric qualities as documented by the OECD in 2003 for Canada**	**Psychometric qualities as documented in this research**
Mathematics self-concept	Mathematics self-concept (SCMAT) “constructed index based on students’ responses about their perceived competence in mathematics” (OECD, 2013, p. 88).	**Assessment instrument comprising 5 Likert-type items on a 4-point scale** Thinking about studying mathematics: to what extent do you agree with the following statements? 1- Strongly agree 2- Agree 3- Disagree 4- Strongly disagree 1. I am just not good at mathematics. 2. I get good grades in mathematics. 3. I learn mathematics quickly. 4. I have always believed that mathematics is one of my best subjects. 5. In my mathematics class, I understand even the most difficult work.	Fit indexes derived from CFA Three-factor model (ANXMAT, SCMAT, MATHEFF) RMSEA = 0.082 RMR = 0.042 CFI = 0.92 TLI = 0.91 Reliability coefficient $\alpha = 0.91$	Fit indexes derived from CFA (1-factor model) 2003: RMSEA = 0.079 SRMR = 0.019 CFI = 0.983 TLI = 0.966 2012: RMSEA = 0.149 SRMR = 0.015 CFI = 0.970 TLI = 0.970 Reliability coefficients 2003: $\omega = 0.801$ 2012: $\omega = 0.770$

Table 1

Instruments for measuring mathematics self-concept, the intrinsic value of mathematics, the utility value of mathematics, and mathematics anxiety, as proposed by PISA in 2003 and 2012

Non-cognitive variable considered in this study	**Name and definition of the index as proposed by PISA 2003 and 2012**	**Contextual questionnaire items**	**Psychometric qualities as documented by the OECD in 2003 for Canada**	**Psychometric qualities as documented in this research**
Intrinsic value of mathematics	Intrinsic motivation to learn mathematics (INTMAT) constructed index "based on students' responses about whether they enjoy mathematics and work hard in mathematics because they enjoy the subject" (OECD, 2013, p. 65)	**Assessment instrument comprising 4 Likert-type items on a 4-point scale** Thinking about your views on mathematics: to what extent do you agree with the following statements? 1- Strongly agree 2- Agree 3- Disagree 4- Strongly disagree 1. I enjoy reading about mathematics. 2. I look forward to my mathematics lessons. 3. I do mathematics because I enjoy it. 4. I am interested in the things I learn in mathematics.	Fit indexes derived from CFA Two-factor model: (INTMAT, INSTMOT) RMSEA = 0.096 RMR = 0.031 CFI = 0.97 TLI = 0.95 Reliability coefficient α = **0.91**	Fit indexes derived from CFA (1-factor model) 2003: RMSEA = 0.101 SRMR = 0.007 CFI = 0.997 TLI = 0.990 2012: RMSEA = 0.042 SRMR = 0.007 CFI = 0.977 TLI = 0.991 Reliability coefficients 2003: ω= 0.880 2012: ω= 0.875

Table 1

Instruments for measuring mathematics self-concept, the intrinsic value of mathematics, the utility value of mathematics, and mathematics anxiety, as proposed by PISA in 2003 and 2012

Non-cognitive variable considered in this study	**Name and definition of the index as proposed by PISA 2003 and 2012**	**Contextual questionnaire items**	**Psychometric qualities as documented by the OECD in 2003 for Canada**	**Psychometric qualities as documented in this research**
Utility value of mathematics	Instrumental motivation to learn mathematics (INSTMOT) constructed index "based on student's responses about […] whether they believe mathematics is important for their future studies and careers" (OECD, 2013, p. 65)	**Assessment instrument comprising 4 Likert-type items on a 4-point scale** Thinking about your views on mathematics: to what extent do you agree with the following statements? 1- Strongly agree 2- Agree 3- Disagree 4- Strongly disagree 1. Making an effort in mathematics is worth it because it will help me in the work that I want to do later on. 2. Learning mathematics is worthwhile for me because it will improve my career prospects and opportunities. 3. Mathematics is an important subject for me because I need it for what I want to study later on. 4. I will learn many things in mathematics that will help me get a job.	Fit indexes derived from CFA Two-factor model (INTMAT, INSTMOT) RMSEA = 0.096 RMR = 0.031 CFI = 0.97 TLI = 0.95 Reliability coefficient α =0.90	Fit indexes derived from CFA (1-factor model) 2003: RMSEA = 0.088 SRMR = 0.016 CFI = 0.987 TLI = 0.961 2012: RMSEA = 0.184 SRMR = 0.013 CFI = 0.992 TLI = 0.977 Reliability coefficients 2003: ω = 0.888 2012: ω = 0.888

Table 1

Instruments for measuring mathematics self-concept, the intrinsic value of mathematics, the utility value of mathematics, and mathematics anxiety, as proposed by PISA in 2003 and 2012

Non-cognitive variable considered in this study	**Name and definition of the index as proposed by PISA 2003 and 2012**	**Contextual questionnaire items**	**Psychometric qualities as documented by the OECD in 2003 for Canada**	**Psychometric qualities as documented in this research**
Mathematics anxiety	Mathematics anxiety (ANXMAT) "constructed index based on students' responses about feelings of stress and helplessness when dealing with mathematics" (OECD, 2013, p. 88).	**Assessment instrument comprising5 Likert-type items on a 4-point scale** Thinking about studying mathematics: to what extent do you agree with the following statements? 1- Strongly agree 2- Agree 3- Disagree 4- Strongly disagree 1. I often worry that it will be difficult for me in mathematics classes. 2. I get very tense when I have to do mathematics homework. 3. I get very nervous doing mathematics problems. 4. I feel helpless when doing a mathematics problem. 5. I worry that I will get poor grades in mathematics.	Fit indexes derived from CFA Three-factor model (ANXMAT, SCMAT, MATHEFF) RMSEA = 0.082 RMR = 0.042 CFI = 0.92 TLI = 0.91 Reliability coefficient α =0.86	Fit indexes derived from CFA (1-factor model) 2003: RMSEA = 0.116 SRMR = 0.026 CFI = 0.982 TLI = 0.963 2012: RMSEA = 0.140 SRMR = 0.032 CFI = 0.973 TLI = 0.946 Reliability coefficients 2003: ω = 0.800 2012: ω = 0.812

Note. CFA = confirmatory factory analysis, RMSEA = root mean square error of approximation, RMR = root mean square residual, SRMR = standardized root mean square residual, CFI = comparative fit index, TLI = Tucker–Lewis index, α = Cronbach's alpha (1951), ω= McDonald's omega (1985, 1999). RMSEA values below 0.05 indicate a good fit, while values above 0.1 are interpreted as an unacceptable fit. RMR or SRMR values below 0.08 indicate a good fit, while values below 0.10 indicate an acceptable fit. CFI, NNFI, and TLI values between 0.9 and 0.95 indicate an acceptable fit, while CFI, NNFI, and TLI values above 0.95 indicate a good fit (Hu & Bentler, 1999). The information in columns 2 and 3 is taken from OECD (2013). The information in column 4 is taken from OECD (2005b, p. 290, 291, 293, 294). The information in column 5 is from personal analyses carried out on the 2003 and 2012 samples in this research.

To verify metric and scalar invariance, we proceeded in stages, first verifying configural invariance, followed by metric invariance, then scalar invariance, between the groups (Meredith, 1993; Widaman & Reise, 1997, Wang & Wang, 2020). Moving from one stage to the other, we studied the changes (Δ) in the following fit indices: the Comparative Fit Index (CFI) (Bentler, 1990), the Root Mean Square Error of Approximation (RMSEA) (Steiger, 1990), and the Standardized Root Mean Square Residual (SRMR) (Bentler, 1995; Muthén, 1998–2004). To conclude, we followed the recommendations of Chen (2007, p. 501): a $\Delta|CFI| \geq ,010$ plus an $\Delta RMSEA \geq ,015$ or an $\Delta SRMR \geq ,030$ between the unconstrained model (configural invariance model) and the model where the factor loading coefficients are constrained to equality (metric invariance model), between the groups for each item, indicates rejection of the metric invariance hypothesis; then, to test scalar invariance, a $\Delta|CFI| \geq ,010$ plus an $\Delta RMSEA \geq ,015$ or an $\Delta SRMR \geq ,010$, between the metric invariance model and the scalar invariance model indicates rejection of the scalar invariance hypothesis.

The results of these analyses are presented in Table 2. They support not rejecting the hypotheses of metric invariance and scalar invariance of the measures of the intrinsic value and utility value of mathematics between girls and boys, for 2003 and for 2012. However, for self-concept, the 2003 and 2012 results indicate that the hypothesis of scalar invariance between boys and girls should be rejected.

The study of modification indices led us to relax the equality constraint on the intercept of item 2, in 2003, and then to relax the equality constraint on the intercepts of items 2 and 5, for 2012. By doing so, the variations in fit indices were found to be acceptable (see Table 2). From this, as we were able to verify the scalar invariance of more than two items for the self-concept instrument, we concluded that it is adequate to compare the means between girls and boys, with regard to this perception (Byrne et al., 1989; Steenkamp & Baumgartner, 1998; Steinmetz, 2011).

Table 2

Study of the invariance of the measure of mathematics self-concept, intrinsic value of mathematics, and utility value of mathematics

PISA Cycle	Non-cognitive variable studied	Type of measurement invariance	CFI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR
2003	Self-concept	Configural	0.986	0.074	0.018			
		Metric	0.982	0.072	0.046	-0.004	-0.002	0.028
		Scalar	0.963	0.090	0.064	-0.019	0.018	0.018
		Release of constraint, item 2	0.976	0.075	0.052	-0.006	0.003	0.006
	Intrinsic value	Configural	0.997	0.044	0.009			
		Metric	0.996	0.041	0.030	-0.001	-0.004	0.021
		Scalar	0.993	0.043	0.034	-0.003	0.002	0.004
	Utility value	Configural	0.989	0.077	0.015			
		Metric	0.981	0.078	0.050	-0.008	0.001	0.035
		Scalar	0.970	0.081	0.040	-0.011	0.003	-0.010
2012	Self-concept	Configural	0.995	0.042	0.013			
		Metric	0.990	0.051	0.048	-0.005	0.009	0.035
		Scalar	0.966	0.082	0.070	-0.024	0.031	0.022
		Release of constraints, items 2 and 5	0.986	0.056	0.049	-0.004	0.005	0.001
	Intrinsic value	Configural	0.994	0.064	0.013			
		Metric	0.989	0.063	0.045	-0.005	-0.001	0.032
		Scalar	0.980	0.072	0.063	-0.009	0.009	0.018
	Utility value	Configural	0.995	0.052	0.011			
		Metric	0.995	0.039	0.019	0.000	-0.013	0.008
		Scalar	0.994	0.036	0.018	-0.001	0.003	-0.001

Note. CFI = Comparative Fit Index, RMSEA = Root Mean Square Error of Approximation, SRMR = Standardized Root Mean Square Residual, Δ = change.

Procedures used to obtain data from the samples of French-speaking Quebec participants in PISA 2003 and 2012

The international PISA 2012 database was downloaded from https://www.oecd.org/pisa/data/pisa2012database-downloadabledata.htm[8]. For 2003, Canada decided against publishing the province's information in the international PISA database. We therefore submitted a request to the Council of Ministers of Education, Canada (CMEC) and received the database of the Quebec student sample for PISA 2003[9].

Data analysis

The software

In order to adequately address the three methodological considerations inherent in the PISA data, the data were analyzed using survey weights, replicated survey weights[10], the maximum likelihood method and the approach of Little and Rubin (2002). All analyses were carried out using Mplus Version 8 software, which has the following features: (a) it supports survey weights and replicated survey weights for students; (b) it incorporates the full information maximum likelihood (FIML) method (Arbuckle, 1996; Little & Rubin, 2002) "in conjunction with ML estimator" (Wang & Wang, 2020, p. 47), the maximum likelihood estimator; and finally, (c) it supports plausible values, processes them according to the approach of Little and Rubin (2002). For processing the data, Mplus Version 8 applies the MLR (maximum likelihood robust) estimator[11] (Muthén & Muthén, 2017).

8. In this database, Canadian participants were identified by the variable CNT = CAN, and French-speaking Quebec participants were identified by the variable STRATUM = CAN0545 or CAN0546 or CAN0547.
9. In this database, the *lang_sector* variable was used to identify Quebec participants from French-speaking schools.
10. The student survey weights are identified by the variable W_FSTUWT in these databases, and then the replicated survey weights are identified by W_FSTR1 to W_FSTR80 (80 replicated survey weights).
11. The MLR estimator is "maximum likelihood parameter estimates with standard errors and a chi-square test statistic (when applicable) that are robust to non-normality and non-independence of observations when used with TYPE=COMPLEX" (Muthén & Muthén, 2017, p. 668). TYPE = COMPLEX is used here because the sampling design is a complex sampling design (see Vohl & Loye, 2024).

The data analysis plan

Analysis for SO1

To meet Specific Objective 1, for each of the three perceptions, a linear regression model with two covariates was used. This model is as follows:

$$Perception = intercept + \beta_1 \cdot gender + \beta_2 \cdot performances. \quad (1)$$

In this model, the gender variable is a dichotomous variable. $\hat{\beta}_1$ estimates the mean deviation of this perception by gender, for each of the three perceptions, controlling for performance.

For perceptions where the gender gap is statistically significant, the effect size is estimated using Cohen's *d*. By setting θ_{boys} and θ_{girls}, the mean of the perception among boys and among girls, *d* is obtained by the formula:

$$d = \frac{\hat{\theta}_{girls} - \hat{\theta}_{boys}}{\sqrt{\dfrac{\sigma^2_{\hat{\theta}_{girls}} + \sigma^2_{\hat{\theta}_{boys}}}{2}}} = \frac{\hat{\beta}_1}{\sqrt{\dfrac{\sigma^2_{\hat{\theta}_{girls}} + \sigma^2_{\hat{\theta}_{boys}}}{2}}} \quad (2)$$

According to Feingold (2019), the value of the expression $\dfrac{\sigma^2_{\hat{\theta}_{girls}} + \sigma^2_{\hat{\theta}_{boys}}}{2}$ in the equation (2) is obtained by taking the residual variance of the regression coefficient of the gender variable in the following single-covariate regression model:

$$Perception = intercept + \beta \cdot \text{gender} \quad (3)$$

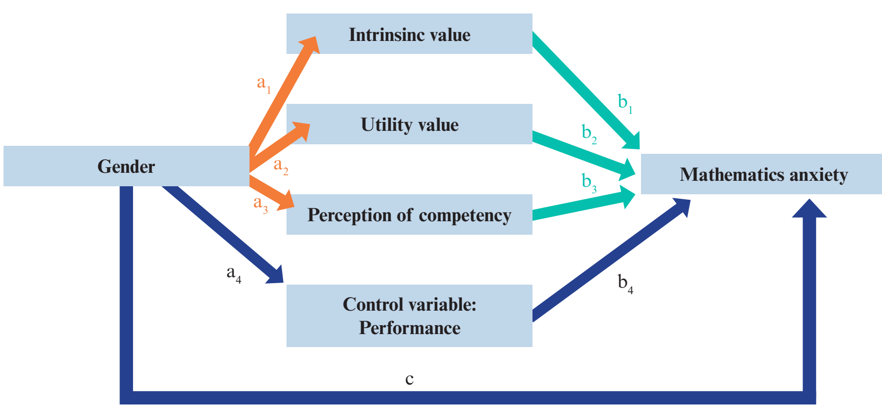
Analysis for SO2

In order to meet Specific Objective 2, we verified whether the relationship between gender and mathematics anxiety is mediated completely by self-concept, intrinsic value, and utility value among 15-year-old French-speaking Quebec students who participated in PISA 2003 and 2012. To do so, we used structural link analysis (e.g., Byrne, 2012; Kline, 2016; Wang & Wang, 2020) to study the direct link between gender and mathematics anxiety and the indirect link between gender and mathematics anxiety via self-concept, intrinsic value, and utility value, controlling for performance. A nonsignificant direct link and a statistically significant indirect link would lead us to conclude that the perceptions completely mediate the link between gender and mathematics anxiety. A significant direct link and a significant indirect link would lead us to conclude that the perceptions partially mediate the link between gender and mathematics anxiety.

To verify whether the indirect link is significant, a 95% confidence interval was provided. A confidence interval was also provided for each of the specific indirect links, namely the link between gender and mathematics anxiety via intrinsic value, the link between gender and mathematics anxiety via utility value, the link between gender and mathematics anxiety via self-concept, and the link between gender and mathematics anxiety via performance. The confidence intervals provided took into account the fact that the distribution of the estimated parameters may be non-normal. Consequently, they were not necessarily symmetric around the estimated parameter (Muthén & Muthén, 2017, p. 613). Figure 2 proposes a representation of the links studied in order to meet Specific Objective 2 and the parallel with Figure 1.

The mathematics anxiety variable was the dependent variable in this analysis; therefore it was necessary to verify the metric invariance of the measure of mathematics anxiety (Wang & Wang, 2020). This procedure was carried out in Vohl and Loye (2023)[12].The results suggested against rejecting the metric invariance hypothesis.

Figure 2
Direct and specific indirect links as studied to meet Specific Objective 2



Note. c represents the direct link between gender and mathematics anxiety. The specific indirect links are given by the following products: $a_1 \bullet b_1$, $a_2 \bullet b_2$, $a_3 \bullet b_3$ et $a_4 \bullet b_4$. The total indirect link between gender and mathematics anxiety is given by the sum $a_1 \bullet b_1 + a_2 \bullet b_2 + a_3 \bullet b_3 + a_4 \bullet b_4$.

12. To compare the underlying means, the scalar invariance of the measurement instrument must be verified. When underlying variables are involved in a linear regression model, their metric invariance must be verified.

Analysis for SO3

In order to study the gender-invariance of the structural links between the three perceptions and mathematics anxiety and controlling for performance, we conducted multigroup structural link analyses (e.g., Byrne, 2012; Wang & Wang, 2020). This involves a three-step approach. First, we studied the goodness-of-fit of the unconstrained model, where the regression coefficients are free to vary by gender. Secondly, we studied the goodness-of-fit of the constrained multigroup model, where the regression coefficients are constrained to equality in both groups. To judge the quality of the overall fit of the data to the proposed model, we used the following indices at each of the two steps: Root Mean Square Error of Approximation (RMSEA) (Steiger, 1990), Comparative Fix Index (CFI) (Bentler, 1990), Tucker–Lewis Index (TLI), (Bentler & Bonett, 1980; Tucker & Lewis, 1973), and Standardized Root Mean Square Residual (SRMR) (Bentler, 1995; Muthén,1998–2004). We applied the interpretive guidelines proposed by Hu and Bentler (1999), which stipulate that an RMSEA value below 0.05, CFI and TLI values above 0.95, and an SRMR value below 0.08 indicate a good fit.

In step 3 we tested the hypothesis of structural link invariance by comparing goodness-of-fit between the unconstrained model and the constrained model. To do this, we used the difference in the comparative fit index (ΔCFI) and concluded there was no reason to reject the invariance hypothesis when the decrease in the CFI is less than 0.01 between the two nested models[13] (Cheung & Rensvold, 2002; Wang & Wang, 2020).

Results

Appendix A presents the descriptive statistics for the six variables studied in the 2003 and 2012 samples, and the percentages of missing data for each variable.

13. In general, when the maximum robust likelihood (MLR) estimator is used, a scaled chi-square difference test (Satorra & Bentler, 2001) between the unconstrained multigroup model and the constrained multigroup model is performed to verify the invariance of two nested models. However, this test could not be carried out in this research, given the use of plausible values (the correction factors provided by Mplus for chi-square difference testing using the Satorra–Bentler method are not available when imputed data are analyzed). The chi-square difference test using the Satorra–Bentler method would also have been of limited use in this research, given its high sensitivity to sample size (e.g., Kline, 2016).

Results for Specific Objective 1

Table 3 shows the regression parameters for the gender variable in each of the two-covariate regression models (Equation 1) performed, and the associated effect sizes calculated from the one-covariate model (Equation 3). The results show that French-speaking Quebec girls reported statistically lower mean levels of self-concept, intrinsic value, and utility value than French-speaking Quebec boys in 2003 and 2012, controlling for performance ($p < 0.10$). These results confirm the research hypothesis raised by SO1.

Echoing previous research (Else-Quest et al., 2010; Frenzel et al., 2007; OECD, 2005a, 2013), our results also showed that the difference in self-concept is greater between boys and grils than the difference in intrinsis and utility value attributed to mathematics (effect sizes of 0.32 and 0.22 for self-concept and effect sizes between 0.10 and 0.17 for intrinsic and utility value, for 2003 and 2012).

Results for Specific Objective 2

To verify whether the relationship between gender and mathematics anxiety is mediated completely by self-concept, intrinsic value, and utility value among 15-year-old French-speaking Quebec students who participated in PISA 2003 and 2012, controlling for performance, we studied the direct and indirect links between gender and mathematics anxiety for 2003 and 2012. Table 4 presents the results obtained.

The results show that, in both 2003 and 2012, the direct and indirect links between gender and mathematics anxiety were found to be statistically significant, controlling for performance. As a result, we concluded that the mediation of the link between gender and mathematics anxiety by the three perceptions is partial, rather than complete, as stipulated in the research hypothesis raised by SO2.

However, in both cycles studied, the regression coefficient of the indirect link was higher than that of the direct link. Indeed, in 2003, the total indirect link represented 69.2% of the total link, and in 2012, it accounted for 65.3%. However, in our study of the specific indirect links, only the indirect link between gender and mathematics anxiety via self-concept proved significant in both 2003 and 2012. Indeed, while the 95% confidence intervals of the other specific indirect links included the value of 0, the specific indirect link between gender and mathematics anxiety via self-concept did not include the value of 0.

Results for Specific Objective 3

Tables 5 and 6 show the results of the multigroup analyses conducted to verify whether the structural links between self-concept, intrinsic value, utility value, and mathematical anxiety can be considered gender-invariant.

Table 5 shows that the unconstrained multigroup model fits the data perfectly for both 2003 and 2012. Table 6 shows that the constrained model provides a good fit for both 2003 and 2012. As the change in CFI value between the unconstrained model and the constrained model is less than 0.01 for both 2003 and 2012, we concluded that the structural links between self-concept, intrinsic value, utility value, and mathematics anxiety can be considered gender-invariant.

Furthermore, the parameter values obtained in the constrained model reveal that the invariant link between intrinsic value and mathematics anxiety is nonsignificant for both 2003 and 2012. This opposes the research hypothesis raised by Specific Objective 3. However, in line with this research hypothesis, the invariant link between utility value and mathematics anxiety was found to be positive, while the invariant link between self-concept and mathematics anxiety was found to be negative.

Taken as a whole, the results obtained in response to Specific Objectives 1, 2, and 3 lead us to conclude that the relationship between gender and mathematics anxiety is mediated by self-concept and that this mediation is partial in both 2003 and 2012. According to the stated research hypothesis, we expected that mediation would be complete and that all three perceptions would mediate the link between gender and mathematics anxiety.

Thus, in 2003, 69.2% of the differences in mathematics anxiety between 15-year-old French-speaking Quebec girls and boys who participated in PISA 2003 and 2012 were explained by the differences in self-concept between girls and boys. In 2012, this percentage rose to 65.3%. The 2003 model explains 43.2% of the variance in mathematics anxiety, and the 2012 model explains 47.2%.

The invariance of the structural links led us to conclude that, for both boys and girls, a one-unit increase in the self-concept index decreases the mathematics anxiety index by about 2/3 of a standard deviation, or 0.601 in 2003 and 0.660 in 2012, and that a one-unit increase in the utility value index increases the mathematics anxiety index by 0.078 and 0.070, or around 7/100 of a standard deviation, in 2003 and 2012, respectively (since these indices are calibrated so that the standard deviation is 1, in OECD countries).

Table 3

Gender differences in self-concept, intrinsic value, and utility value observed among French-speaking Quebec students who participated in in PISA 2003 and 2012 and associated effect sizes

	2003 Gender differences (boys – girls)				**2012 Gender differences (boys – girls)**			
Variable	$\hat{\beta}_1$	**SE**	***p***	***d***	$\hat{\beta}_1$	**SE**	***p***	***d***
SCMAT	0.31	0.05	< 0.001	0.32	0.37	0.06	< 0.001	0.22
INTMAT	0.14	0.041	< 0.001	0.15	0.16	0.05	< 0.001	0.17
INSTMOT	0.10	0.047	0.028	0.11	0.10	0.06	0.092	0.10

Note. $\hat{\beta}_1$= estimated non-standardized regression coefficient of the gender variable in the regression model given by *Perception* = $\hat{\beta}_1 \cdot$ *gendre* + $\hat{\beta}_2 \cdot$ *performances* with the reference gender = 0 for girls; SE = standard error of the regression coefficient $\hat{\beta}_1$; *p* = p-value p; *d* = effect size; SCMAT = mathematics self-concept; INTMAT = intrinsic value of mathematics; INSTMOT = utility value of mathematics.

Table 4

Results of direct and indirect links for 2003 and 2012, controlling for performance

LINK(S)		2003					2012				
		Coeff.	*SE*	*p*	% T	c.i.	Coeff.	*SE*	*p*	% T	c.i.
TOTALS	GENDER →ANXMAT GENDER → INTMAT → ANXMAT GENDER → INSTMOT → ANXMAT GENDER → SCMAT → ANXMAT GENDER → Performance → ANXMAT	-0.328	0.051	< 0.05	100.0%		-0.418	0.053	< 0.05	100%	
DIRECT	GENRE → ANXMAT (c)	-0.101	0.039	<0.05	30.8%		-0.145	0.041	<0.05	34.7%	
INDIRECT	GENDER → INTMAT → ANXMAT GENDER → INSTMOT → ANXMAT GENDER → SCMAT → ANXMAT GENDER → Performance → ANXMAT	-0.227	0.033	-	69.2%	[-0.292; -0.161]	-0.273	0.036	-	65.3%	[-0.332; -0.190]
INDIRECT SPECIFIC	GENDER → INTMAT → ANXMAT (a_1* b_1)	0.001	0.005	-	0.0%	[-0.010; 0.011]	0.014	0.010	-	0.0%	[-0.006; 0.034]
	GENDER → INSTMOT → ANXMAT (a_2* b_2)	0.010	0.005	-	0.0%	[0.000; 0.020]	0.008	0.006	-	0.0%	[-0.003; 0.020]
	GENDER → SCMAT → ANXMAT (a_3* b_3)	-0.227	0.035	-	69.2%	[-0.295; -0.159]	-0.285	0.040	-	65.3%	[-0.362; -0.207]
	GENDER → Performance → ANXMAT (a_4* b_4)	-0.010	0.007	-	0.0%	[-0.024; 0.004]	-0.011	0.007	-	0.0%	[-0.226;-0.064]

Table 4 (suite)

Results of direct and indirect links for 2003 and 2012, controlling for performance

LINK(S)		2003					2012				
		Coeff.	*SE*	*p*	% T	c.i.	Coeff.	*SE*	*p*	% T	c.i.
COMPOSITION	GENRE → INTMAT (a_1)	0.156	0.043	< 0.05			0.199	0.046	< 0.05		
SPECIFIC	GENRE → INSTMOT (a_2)	0.116	0.047	< 0.05			0.126	0.058	< 0.05		
INDIRECT	GENRE → SCMAT (a_3)	0.340	0.046	< 0.05			0.391	0.048	< 0.05		
LINKS	GENRE → Performances (a_4)	0.076	0.053	0.154			0.134	0.056	< 0.05		
	INTMAT → ANXMAT (b_1)	0.005	0.034	0.876			0.071	0.052	0.170		
	INSTMOT → ANXMAT (b_2)	0.085	0.026	< 0.05			0.066	0.040	0.097		
	SCMAT → ANXMAT (b_3)	-0.668	0.030	< 0.05			-0.728	0.036	< 0.05		
	Performance → ANXMAT (b_4)	-0.135	0.021	< 0.05			-0.079	0.028	< 0.05		

Note. The colors correspond to those used in Figure 2. Gender = 0 for girl, *Coeff.* = standardized regression coefficient with respect to *y*, *SE* = standard error, *p* = p-value,%T = percentage of total effect, c.i. = intervalle de confiance à 95%, confidence interval, SCMAT = mathematics self-concept, INTMAT = intrinsic value of mathematics, INSTMOT = utility value of mathematics, ANXMAT = mathematics anxiety. For 2003, the percentage of variation explained by the model and the data's fit indices to the model: R^2 = 43.2%, RMSEA = 0.000, CFI = 1.000, TLI = 1.000, SRMR = 0.000. For 2012, the percentage of variation explained by the model and the fit indices are R^2 = 47.2%, RMSEA = 0.000, CFI = 1.000, TLI = 1.000, SRMR = 0.000.

Table 5

Results of the unconstrained multigroup model for 2003 and for 2012

	2003						2012					
	Girls			Boys			Girls			Boys		
Independent variables	**Coeff.**	**SE**	***p***	**Coeff.**	**SE**	***p***	**Coeff.**	**SE**	***p***	**Coeff.**	**SE**	***p***
INTMAT	-0.043	0.031	0.162	0.049	0.047	0.292	-0.033	0.054	0.538	0.159	0.060	0.008
INSTMOT	-0.080	0.024	0.001	0.071	0.035	0.043	0.108	0.044	0.013	0.005	0.055	0.926
SCMAT	-0.607	0.029	< 0.001	-0.580	0.047	< 0.001	-0.612	0.042	< 0.001	-0.716	0.045	<0.001
PERFORMANCES	-0.001	0.000	< 0.001	-0.001	0.000	< 0.001	-0.001	0.000	0.016	-0.001	0.000	0.049
R^2	53.3%			37.2%			50.7%			45.3%		
Fit indices	RMSEA = 0.000 CFI = 1.000 TLI= 1.000 SRMR = 0.000						RMSEA = 0.000 CFI = 1.000 TLI= 1.000 SRMR = 0.000					

Note. Coeff. = non-standardized regression coefficient, SE = standard error, p = p-value, R^2 = percentage of variation explained by the model, INTMAT = intrinsic value of mathematics, INSTMOT = utility value of mathematics, SCMAT = mathematics self-concept, RMSEA = root mean square error of approximation, CFI = comparative fit index, TLI = Tucker–Lewis index, SRMR = standardized root mean square residual.

Table 6

Results of the constrained multigroup model for 2003 and for 2012

	2003						2012					
	Girls			Boys			Girls			Boys		
Independent variables	***Coeff.***	***SE***	***p***	***Coeff.***	***SE***	***p***	***Coeff.***	***SE***	***p***	***Coeff.***	***SE***	***p***
INTMAT	-0.002	0.028	0.945	-0.002	0.028	0.945	0.055	0.046	0.243	0.055	0.046	0.243
INSTMOT	0.078	0.020	< 0.001	0.078	0.020	< 0.001	0.070	0.034	0.041	0.070	0.034	0.041
SCMAT	-0.601	0.024	< 0.001	-0.601	0.024	< 0.001	-0.660	0.031	< 0.001	-0.660	0.031	< 0.001
PERFORMANCES	-0.001	0.000	< 0.001	-0.001	0.000	< 0.001	-0.001	0.000	0.003	-0.001	0.000	0.003
R^2	51.1%			40.3%			51.9%			42.3%		
Fit indices	RMSEA = 0.028 CFI = 0.996 TLI= 0.991 SRMR = 0.017 ΔCFI= 1 - 0.996 = 0.004						RMSEA = 0.033 CFI = 0.994 TLI= 0.989 SRMR = 0.015 ΔCFI= 1-0.994 = 0.006					

Note. Coeff. = non-standardized regression coefficient, SE = standard error, p = p-value, R^2 = percentage of variation explained by the model, INTMAT = intrinsic value of mathematics, INSTMOT = utility value of mathematics, SCMAT = mathematics self-concept, RMSEA = root mean square error of approximation, CFI = comparative fit index, TLI = Tucker–Lewis index, SRMR = standardized root mean square residual, ΔCFI = change in the comparative fit index between the constrained model and the unconstrained model.

Discussion

The overall aim of this paper was to identify factors that could help explain the gender-based disparity in mathematics anxiety observed among 15-year-old French-speaking Quebec students who participated in PISA 2003 and 2012 and consequently, potentially reduce these disparities. To this end, we adopted the track proposed by Pekrun's control-value model for achievement emotions (2006). In line with this line of inquiry, we hypothesized that the gender differences in mathematics anxiety could be explained by differences in perceptions of competency, intrinsic value, and utility value between boys and girls. To test this hypothesis, we stated and tested three specific research objectives.

By combining the results obtained in response to these three objectives, it appears that the gender-based disparity in mathematics anxiety observed among 15-year-old French-speaking Quebec students who participated in PISA 2003 and 2012 can be explained by a weaker mathematics self-concept among girls than among boys, and by the presence of a direct negative link between gender and mathematics anxiety. Disparities in self-concept explain most of the disparities in mathematics anxiety: 69.2% in 2003 and 65.3% in 2012.

These findings have important scientific and practical implications. In fact, firstly, the presence of the direct link prompts us to look at the factors which—in addition to self-concept, intrinsic value, and utility value—could explain the disparities in mathematics anxiety observed between girls and boys. In fact, the literature proposes tracks other than Pekrun's control–value model for achievement emotions (2006) to explain gender differences in mathematics anxiety. Indeed, factors relating to the components of affective and cognitive predispositions in Pekrun's control–value model for achievement emotions (2006) are implicated in explaining the disparities in mathematics anxiety between girls and boys.

With regard to individual affective predispositions, some authors have proposed that gender disparity in mathematics anxiety could be explained by differences in the average level of anxiety experienced by boys and girls (e.g., Dowker et al., 2016; Frenzel et al., 2007; Hill et al., 2016). In fact, studies indicate that girls have a higher average level of general anxiety than boys (e.g., Chapman et al, 2007; Costa et al., 2001; Feingold, 1994) and a higher prevalence of anxiety disorders than boys (e.g., McLean et al., 2011). In this regard, Szczygiel (2020) showed that, among first- and second-year primary school students, general anxiety mediates the relationship between gender and mathematics anxiety. The author concluded that girls may report being more anxious about mathematics because their average level of general anxiety is higher than that of boys.

With regard to cognitive predispositions, and in connection with the reduced competency account, research conducted among adult populations by Delage et al. (2022), Ferguson et al. (2015), Gibeau et al. (2023), Maloney et al. (2012), and Sokolowski et al. (2019) reveal that a disparity in spatial processing ability between men and women could explain, at least in part, the gender differences observed in mathematics anxiety. Indeed, these authors have shown that spatial processing ability mediates the relationship between gender and mathematics anxiety.

These various factors relating to the cognitive and affective predispositions of Pekrun's control-value model for achievement emotions were not documented in the PISA assessments and, consequently, were not studied in the context of this research. It would therefore be relevant for future research to explore or at least consider them as covariates when aiming to explain the disparity in mathematics anxiety between boys and girls via mathematics self-concept, intrinsic value, and utility value. At the same time, given that the disparities in mathematics anxiety observed between French-speaking Quebec girls and boys can largely be explained by an average lower mathematics self-concept among girls than among boys, practical solutions for reducing the disparities would benefit from leveraging self-concept.

Accordingly, we address the main factors underlying the differences in self-concept between girls and boys. As a result, we are able to make recommendations for reducing the gaps in self-concept observed between boys and girls, which would in turn reduce the disparities in mathematics anxiety.

Sources of gender differences in self-concept and recommendations

Gender stereotypes in mathematics play a major role in explaining gender differences in self-concept (e.g., Ertl et al., 2017; Makarova et al., 2019; Rossi et al., 2022). Indeed, studies have shown that, given equal ability, girls who adhere to gender stereotypes in mathematics report a lower self-concept than boys (Ertl et al., 2017; Guimond & Roussel, 2001). In contrast, boys who adhere to gender stereotypes in mathematics disclose a higher self-concept than their female peers (Rossi et al., 2022).

Significant others (parents and teachers) play a key role in the implicit and explicit transmission of gender stereotypes (e.g., Ertl et al., 2017). Indeed, in addition explicitly transmitting their beliefs (Gunderson et al., 2012; Herbert & Stipek, 2005), significant others who adhere to stereotypes tend to have lower expectations of success in mathematics for girls than for boys (Dresel et al., 2007, cited in Ertl et al., 2017).

Also, significant others who adhere to gender stereotypes in mathematics tend to attribute learners' successes and failures in gender-differentiated ways. For example, they tend to attribute boys' successes to their mathematical skill and girls' successes to their diligence (Kessels, 2015), while attributing boys' failures to a lack of effort and girls' failures to a lack of skill (Dresel et al., 2007, cited in Ertl et al., 2017). It has been shown that the development of mathematics self-concept is promoted by attributing success to effort and failure to a lack of effort (Ertl et al., 2017).

Still, with regard to gender stereotypes in mathematics, it appears that parents and teachers who adhere to gender stereotypes in mathematics also tend to adopt differentiated behaviors toward learners, depending on gender. For example, they offer more numeracy-oriented games to boys and more literacy-oriented games to girls (Gunderson et al., 2012). They present boys with greater mathematical challenges than girls (Gunderson et al., 2012), then they tend to interfere more with girls' math-related work. These behaviors are detrimental to girls' mathematics self-concept. Indeed, parents and teachers leave fewer opportunities for girls to develop their familiarity and skills in the field (Ertl et al., 2017; Mejía-Rodríguez et al., 2021; Skaalvik & Skaalvik, 2004). In this context, combating gender stereotypes in mathematics is an essential lever for reducing the disparities in self-concept between girls and boys (Encinas-Martín & Cherian, 2023).

Other beliefs about mathematics can also be addressed to mitigate this issue. For example, the stereotype that mathematics is innate (e.g., "in mathematics, some people have it, others don't!") could also be detrimental to learners' self-concept, especially among girls (Ertl et al., 2017). Combating this stereotype is also extremely important, like valuing mistakes as a way of learning (e.g., Ertl et al., 2017; Ramirez et al., 2018) and promoting effort and a job well done in mathematics. Gourdeau, cited in Kinnard, 2019, argued that success in mathematics requires discipline and hard work rather than intelligence and that all the great mathematicians attributed their success to hard work.

According to Stoet et al. (2016), it is also important to promote the intrinsic value of mathematics among young girls to reduce the gaps in mathematics self-concept A secondary analysis of the PISA 2012 data by Stoet et al. (2016) showed that the countries with the highest equality indices (including Canada) also report the highest gender disparities in mathematics anxiety, self-concept, and intrinsic value. According to the authors, in more egalitarian and more developed countries, girls are less subject to economic considerations when making academic choices. As a result, they are freer to choose according to their own interests. In this context, they choose mathematics more rarely than boys, given them fewer opportunities to develop their knowledge of the discipline and their math skills. As a result, they report a lower mathematics self-concept compared to boys.

One key to successfully combatting mathematics stereotypes and fostering the intrinsic value of mathematics among young girls is encouraging parents, teachers, and significant to resolve their own mathematics anxiety (e.g., Beilock et al., 2010; Casad et al., 2015; Maloney et al., 2015; Vukovic et al., 2013). Indeed, not only are people who report mathematics anxiety likely to adhere to mathematics gender stereotypes in (Beilock et al., 2010; Casad et al., 2015; Ertl et al., 2017; Maloney et al., 2015; Vukovic et al., 2013) and to transmit them (Beilock et al., 2010; Bieg et al., 2015; Casad et al., 2015; Goetz et al., 2013), but they also report and transmit negative perceptions of mathematics. However, studies have shown that adults can overcome mathematics anxiety, improve their performance in mathematics (e.g., Jamieson et al., 2016; Park et al., 2014; Passolunghi et al., 2020), and boost their mathematics self-concept (e.g., Gresham, 2007, cited in Casad et al., 2015).

In summary, reducing the disparities in self-concept between girls and boys requires: (a) combatting stereotypes in mathematics, (b) promoting effort and perseverance in mathematics, (c) fostering the intrinsic value of mathematics among young girls, and (d) encouraging significant others to resolve their own mathematics anxiety. Furthermore, according to Goetz et al. (2013), it is essential to lift the veil on the discrepancy between performance in mathematics and mathematics self-concept for girls. According to these authors, it is vital that girls realize that they report average lower mathematics self-concept than boys, even with equal skill. The simple fact that they wrongly perceive themselves as less competent leads them to report more anxiety about mathematics, to perform less well in mathematics, and to turn away from STEM. This is a negative cascade that harms women and society in general (OECD, 2017). For Devine et al. (2012), this cascade is all the more deplorable given that, despite the significant gaps in mathematics anxiety and self-concept observed between girls and boys, girls manage to achieve results barely below the results of boys, on average, when gaps in performance in mathematics are actually observed. In this case, Devine et al. (2012) conclude that girls may have extremely high potential in mathematics.

Limitations

The cross-sectional nature of the data analyzed in our research does not allow us to conclude that there is a causal relationship between self-concept, intrinsic value, utility value, and mathematics anxiety.

Thus, "to minimize the potential for response bias, data quality standards in PISA require minimum participation rates for schools and students. At the Canada-wide level, a minimum response rate of 85 per cent was required for schools initially selected" (O'Grady et al., 2016, p. 49). PISA "also requires a minimum student participate rate of 80 per cent within all participating schools combined (original sample and replacements) at the national level" (O'Grady et al., 2016, p. 49). However, in 2012, in Quebec, the rate of school participation in PISA was above the minimum required threshold, but the student response rate was 75.6%, below the 80% threshold set by the OECD (Brochu et al., 2013). An analysis of the characteristics of respondents and non-respondents alike revealed that the phenomenon "[…] may marginally impact the results for Quebec […]" (Brochu et al., 2013, p. 52).

Furthermore, although – unlike the vast majority of studies – we documented scalar invariance or partial scalar invariance for the four non-cognitive concepts studied in our research before proceeding with comparisons of the means between boys and girls, we cannot rule out the possibility that the observed gender differences can be explained by methodological considerations (e.g., Devine et al., 2012; Frenzel et al., 2007; Hill et al., 2016). In fact, self-reported measurement instruments, such as those used to document mathematics self-concept, intrinsic value, utility value, and mathematics anxiety, are known to be associated with various types of biases, such as gender-linked response bias (Devine et al., 2012, p. 6), social-desirability bias (Paulhus, 2017), and even response bias attributable to inadequate calibration of students' perceptions and judgments of learning (Winne & Jamieson-Noel, 2002), etc.

Conclusion

In this research, we set out to identify factors that might help explain the disparities in mathematics anxiety observed between 15-year-old French-speaking Quebec boys and girls, documented by PISA data from 2003 and 2012. Building on the basic premise of Pekrun's control-value model for achievement emotions (2006), we attempted to verify whether the differences in self-concept, intrinsic value, and utility value between boys and girls could explain the disparities in mathematics anxiety. Our findings show that differences in self-concept and a direct link between gender and mathematics anxiety explain the differences in mathematics anxiety observed among French-speaking students in Quebec. From this, we propose avenues to explore in future research to explain the direct link between gender and mathematics anxiety, and we put forward a series of recommendations that could foster the proper development of girls' self-concept and, ultimately, enable more of them to reach their potential in mathematics (OECD, 2014a).

Proofreading: Caroline Lefour

Formatting: Emmanuel Gagnon

Portugue abstract: Eusebio Andre Machado

Original version (in French)

Received: February 24, 2024

Final version: April 10, 2024

Accepted: June 11, 2024

LISTE DES RÉFÉRENCES

Arbuckle, J. L. (1996). Full information estimation in the presence of incomplete data. Dans Marcoulides, G.A. & Schumacker, R.E. (dir.), *Advanced structural equation modeling: Issues and techniques* (p. 243 – 277). Lawrence Erlbaum Associates, Inc.

Ashcraft, M. H. & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*(2), 224-237. https://doi.org/10.1037/0096-3445.130.2.224

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559. https://doi.org/10.1126/science.1736359

Baddeley, A. D. & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, *8*, 47-89.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.

Beilock, S. L. & Carr, T. H. (2005). When high-powered people fail: Working memory and "choking under pressure" in math. *Psychological Science*, *16*(2), 101-105. https://doi.org/10.1111/j.0956-7976.2005.00789.x

Beilock, S. L., Gunderson, E. A., Ramirez, G. & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(5), 1860-1863. https://doi.org/10.1073/pnas.0910967107

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238-246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M. (1995). *EQS structural equations program manual* (Vol. 6). Multivariate software.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588-606. https://doi.org/10.1037/0033-2909.88.3.588

Berger, J.-L. (2015). *Apprendre: la rencontre entre motivation et métacognition. Autorégulation dans l'apprentissage des mathématiques en formation professionnelle*. Peter Lang.

Bieg, M., Goetz, T., Wolter, I. & Hall, N. C. (2015). Gender stereotype endorsement differentially predicts girls' and boys' trait-state discrepancy in math anxiety. *Frontiers in psychology*, *6*, Article 1404, 1-8. https://doi.org/10.3389/fpsyg.2015.01404

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.

Brochu, P., Deussing, M.-A., Houme, K. & Chuy, M. (2013). *Measuring up: Canadian Results of the OECD PISA Study - The Performance of Canada's Youth in Mathematics, Reading, and Science - 2012 First Results for Canadians Aged 15*. Council of Ministers of Education Canada. https://cmec.ca/publications/lists/publications/attachments/318/pisa2012_canadianreport_en_web.pdf

Bussière, P., Knighton, T. & Cartwright, F. (2004). *À la hauteur : Résultats canadiens de l'études PISA de l'OCDE : la performance des jeunes du Canada en mathématiques, en lecture, en sciences et en résolution de problèmes : premiers résultats de 2003 pour les Canadiens de 15 ans*. Statistique Canada. https://www.cmec.ca/docs/pisa2003.fr.pdf

Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge.

Casad, B. J., Hale, P. & Wachs, F. L. (2015). Parent-child math anxiety and math-gender stereotypes predict adolescents' math education outcomes. *Frontiers in psychology*, *6*, Article 1597,1-21. https://doi.org/10.3389/fpsyg.2015.01597

Chapman, O. (2007). Facilitating preservice teachers' development of mathematics knowledge for teaching arithmetic operations. *Journal of Mathematics Teacher Education*, *10*, 341-349. https://doi.org/10.1007/s10857-007-9046-8

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural equation modeling: a multidisciplinary journal*, *14*(3), 464-504. https://doi.org/10.1080/10705510701301834

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, *9*(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5

Chiu, M. M. & Klassen, R. M. (2010). Relations of mathematics self-concept and its calibration with mathematics achievement: cultural differences among fifteen-year-olds in 34 countries. *Learning Instruction*, *20*(1), 2-17. https://doi.org/10.1016/j.learninstruc.2008.11.002

Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American journal of Sociology*, *106*(6), 1691-1730. https://psycnet.apa.org/doi/10.1086/321299

Costa Jr, P. T., Terracciano, A. & McCrae, R. R. (2001). Gender differences in personality traits across cultures: robust and surprising findings. *Journal of personality and social psychology*, *81*(2), 322-331. https://doi.org/10.1037/0022-3514.81.2.322

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

Delage, V., Trudel, G., Retanal, F. & Maloney, E. A. (2022). Spatial anxiety and spatial ability: Mediators of gender differences in math anxiety. *Journal of Experimental Psychology: General*, *151*(4), 921-933. https://psycnet.apa.org/doi/10.1037/xge0000884

Devine, A., Fawcett, K., Szűcs, D. & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and brain functions*, *8*, 1-9. https://doi.org/10.1186/1744-9081-8-33

Dowker, A., Sarkar, A. & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years? *Frontiers in psychology*, *7*, Article 508, 1-16. https://doi.org/10.3389/fpsyg.2016.00508

Else-Quest, N. M., Hyde, J. S. & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin*, *136*(1), 103. http://dx.doi.org/10.1037/a0018053

Encinas-Martín, M. & M. Cherian (2023). *Gender, Education and Skills: The Persistence of Gender Gaps in Education and Skills*, OECD Skills Studies. OECD Publishing. https://doi.org/10.1787/34680dd5-en

Enders, C. K. (2010). *Applied missing data analysis*. Guilford press.

Ertl, B., Luttenberger, S. & Paechter, M. (2017). The impact of gender stereotypes on the self-concept of female students in STEM subjects with an under-representation of females. *Frontiers in psychology*, *8*, Article 703, 1-11. https://doi.org/10.3389/fpsyg.2017.00703

Feingold, A. (1994). Gender differences in personality: a meta-analysis. *Psychological Bulletin*, *116*(3), 429. https://doi.org/10.1037/0033-2909.116.3.429

Feingold, A. (2019). New approaches for estimation of effect sizes and their confidence intervals for treatment effects from randomized controlled trials. *The quantitative methods for psychology*, *15*(2), 96-111. https://doi.org/10.20982%2Ftqmp.15.2.p096

Ferguson, A. M., Maloney, E. A., Fugelsang, J. & Risko, E. F. (2015). On the relation between math and spatial ability: The case of math anxiety. *Learning Individual Differences*, *39*, 1-12. https://doi.org/10.1016/j.lindif.2015.02.007

Fonseca, R., Fontaine, M. M. & Haeck, C. (2021). *Le lien entre les compétences en numératie et les rendements sur le marché du travail au Québec* (2021RP-11, CIRANO). https://www.cirano.qc.ca/fr/sommaires/2021RP-11

Frenzel, A. C., Pekrun, R. & Goetz, T. (2007). Girls and mathematics—A “hopeless” issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, *22*(4), 497-514. https://doi.org/10.1007/BF03173468

Geary, D. C. & Widaman, K. F. (1992). Numerical cognition: On the convergence of componential and psychometric models. *Intelligence*, *16*(1), 47-80. https://doi.org/10.1016/0160-2896(92)90025-M

Gibeau, R. M., Maloney, E. A., Béland, S., Lalande, D., Cantinotti, M., Williot, A., Chanquoy, L., Simon, J., Boislard-Pépin, M.-A. & Cousineau, D. (2023). The Correlates of Statistics Anxiety: Relationships with Spatial Anxiety, Mathematics Anxiety and Gender. *Journal of Numerical Cognition*, *9*(1), 16-43. https://doi.org/10.5964/jnc.8199

Goetz, T., Bieg, M., Lüdtke, O., Pekrun, R. & Hall, N. C. (2013). Do Girls Really Experience More Anxiety in Mathematics? *Psychological Science*, *24*(10), 2079-2087. https://doi.org/10.1177%2F0956797613486989

Goetz, T., Frenzel, A. C., Hall, N. C. & Pekrun, R. (2008). Antecedents of academic emotions: Testing the internal/external frame of reference model for academic enjoyment. *Contemporary Educational Psychology*, *33*(1), 9-33. https://doi.org/10.1016/j.cedpsych.2006.12.002

Goldman, A. D. & Penner, A. M. (2016). Exploring international gender differences in mathematics self-concept. *International Journal of Adolescence and Youth*, *21*(4), 403-418. https://doi.org/10.1080/02673843.2013.847850

Graham, J. W., Hofer, S. M. & MacKinnon, D. P. (1996). Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*, *31*(2), 197-218. https://doi.org/10.1207/s15327906mbr3102_3

Guimond, S. & Roussel, L. (2001). Bragging about one’s school grades: Gender stereotyping and students’ perception of their abilities in science, mathematics, and language. *Social psychology of education*, *4*, 275-293. https://doi.org/10.1023/A:1011332704215

Gunderson, E. A., Ramirez, G., Levine, S. C. & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex roles*, *66*, 153-166.

Hayduk, L. A. (1987). *Structural equation modeling with LISREL: Essentials and advances*. Jhu Press.

Henschel, S. & Roick, T. (2017). Relationships of mathematics performance, control and value beliefs with cognitive and affective math anxiety. *Learning and Individual Differences*, *55*, 97-107. https://doi.org/10.1016/j.lindif.2017.03.009

Herbert, J. & Stipek, D. (2005). The emergence of gender differences in children's perceptions of their academic competence. *Journal of applied developmental Psychology*, *26*(3), 276-295. https://doi.org/10.1016/j.appdev.2005.02.007

Hidi, S. & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111-127. https://doi.org/10.1207/s15326985ep4102_4

Hill, F., Mammarella, I. C., Devine, A., Caviola, S., Passolunghi, M. C. & Szűcs, D. (2016). Maths anxiety in primary and secondary school students: Gender differences, developmental changes and anxiety specificity. *Learning and individual differences*, *48*, 45-53. https://doi.org/10.1016/j.lindif.2016.02.006

Hu, L.-t. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: A Multidisciplinary Journal*, *6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hyde, J. S., Fennema, E. & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*(2), 139-155. https://doi.org/10.1037/0033-2909.107.2.139

Jamieson, J. P., Peters, B. J., Greenwood, E. J. & Altose, A. J. (2016). Reappraising stress arousal improves performance and reduces evaluation anxiety in classroom exam situations. *Social Psychological and Personality Science*, *7*(6), 579-587. https://psycnet.apa.org/doi/10.1177/1948550616644656Joensen, J. S. & Nielsen, H. S. (2009). Is there a causal effect of high school math on labor market outcomes? *Journal of Human Resources*, 44(1), 171-198.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, *36*(4), 409-426.

Kessels, U. (2015). Bridging the gap by enhancing the fit: How stereotypes about STEM clash with stereotypes about girls. *International Journal of Gender, Science and Technology*, *7*(2), 280-296. https://genderandset.open.ac.uk/index.php/genderandset/article/view/392

Kinnard, N. (2019, hiver). Qui a peur des mathématiques ? *Contact*, 26-29. https://contact.ulaval.ca/article_magazine/qui-a-peur-des-mathematiques/index.html

Kline, R. B. (2016). *Principles and practice of Structural equation modeling* (4e éd.). Guilford Press.

Kyttälä, M. & Björn, P. M. (2010). Prior mathematics achievement, cognitive appraisals and anxiety as predictors of Finnish students' later mathematics performance and career orientation. *Educational Psychology Review*, *30*(4), 431-448. https://doi.org/10.1080/01443411003724491

Little, R. J. & Rubin, D. B. (2002). *Statistical analysis with missing data* (2e éd.). John Wiley & Sons.

Lohr, S.L. (2019). *Sampling: Design and Analysis* (2e éd.). Chapman et Hall/CRC. https://doi.org/10.1201/9780429296284

Ma, X. (1999). A meta-analysis of the relationship between anxiety toward mathematics and achievement in mathematics. *Journal for Research in Mathematics Education*, *30*(5), 520-540. https://doi.org/10.2307/749772

Makarova, E., Aeschlimann, B. & Herzog, W. (2019). The gender gap in STEM fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. *Frontiers in Education*, 4, Article 60, 1-11.

Maloney, E. A., Ansari, D. & Fugelsang, J. A. (2011). The effect of mathematics anxiety on the processing of numerical magnitude. *The Quarterly Journal of Experimental Psychology*, *64*(1), 10-16. https://doi.org/10.1080/17470218.2010.533278

Maloney, E. A., Ramirez, G., Gunderson, E. A., Levine, S. C. & Beilock, S. L. (2015). Intergenerational effects of parents' math anxiety on children's math achievement and anxiety. *Psychological Science*, *26*(9), 1480-1488. https://doi.org/10.1177/0956797615592630

Maloney, E. A., Risko, E. F., Ansari, D. & Fugelsang, J. (2010). Mathematics anxiety affects counting but not subitizing during visual enumeration. *Cognition*, *114*(2), 293-297. https://doi.org/10.1016/j.cognition.2009.09.013

Maloney, E. A., Waechter, S., Risko, E. F. & Fugelsang, J. (2012). Reducing the sex difference in math anxiety: The role of spatial processing ability. *Learning Individual Differences*, *22*(3), 380-384. https://doi.org/10.1016/j.lindif.2012.01.001

Maloney, E. A. (2016). Math anxiety: Causes, consequences, and remediation. Dans K. R. Wentzel et D. B. Miele (dir.), *Handbook of motivation at school* (2e éd., p. 408-423). Routledge.

Marsh, H. W., Barnes, J. & Hocevar, D. (1985a). Self–other agreement on multidimensional self-concept ratings: Factor analysis and multitrait–multimethod analysis. *Journal of personality and social psychology*, *49*(5), 1360-1377. https://doi.org/10.1037/0022-3514.49.5.1360

Marsh, H. W., Byrne, B. M. & Shavelson, R. J. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. *Journal of Educational Psychology*, *80*(3), 366-380. https://doi.org/10.1037/0022-0663.80.3.366

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*(1), 56-67. https://doi.org/10.1037/0022-0663.96.1.56

Marsh, H. W. & Scalas, L. F. (2011). Self-concept in learning: Reciprocal effects model between academic self-concept and academic achievement. Dans S. Jarvela (dir.), *Social and emotional aspects of learning* (p. 191-198). Elsevier.

Marsh, H. W. & Shavelson, R. (1985b). Self-concept: Its multifaceted, hierarchical structure. *Educational Psychologist*, *20*(3), 107-123. https://doi.org/10.1207/s15326985ep2003_1

Marsh, H. W., Smith, I. D. & Barnes, J. (1985c). Multidimensional self-concepts: Relations with sex and academic achievement. *Journal of Educational Psychology*, *77*(5), 581-596. https://doi.org/10.1037/0022-0663.77.5.581

Marsh, H. W. & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal*, *35*(4), 705-738. https://doi.org/10.3102/00028312035004705

Martin, A. J. & Marsh, H. W. (2008). Workplace and academic buoyancy: Psychometric assessment and construct validity amongst school personnel and students. *Journal of Psychoeducational Assessment*, *26*(2), 168-184. https://doi.org/10.1177/0734282907313767

McDonald, R. P. (1985). *Factor analysis and related methods*. Lawrence Erlbaum.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Lawrence Erlbaum.

McLean, C. P., Asnaani, A., Litz, B. T. & Hofmann, S. G. (2011). Gender differences in anxiety disorders: prevalence, course of illness, comorbidity and burden of illness. *Journal of psychiatric research*, *45*(8), 1027-1035. https://doi.org/10.1016/j.jpsychires.2011.03.006

Meece, J. L., Wigfield, A. & Eccles, J. S. (1990). Predictors of math anxiety and its influence on young adolescents' course enrollment intentions and performance in mathematics. *Journal of Educational Psychology*, *82*(1), 60-70. https://doi.org/10.1037/0022-0663.82.1.60

Mejía-Rodríguez, A. M., Luyten, H. & Meelissen, M. R. (2021). 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

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525-543. https://doi.org/10.1007/BF02294825

Miyake, A. & Shah, P. (dir.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. Cambridge University Press.

Mullis, I. V., Martin, M. O., Foy, P. & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Boston College, TIMSS & PIRLS International Study Center website.

Mullis, I.V., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Smith, T. A, Chrostowski, S. J. & O'Connor, K. M. (2000). *TIMSS 1999 International Mathematics Report Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. https://timssandpirls.bc.edu/timss1999i/pdf/T99i_Math_All.pdf

Muthén, B. O. (1998-2004). *Mplus technical appendices*. Muthén & Muthén.

Muthén, L. K. & Muthén, B. O. (2017). *Mplus User's guide* (8e éd.). Muthén & Muthén.

O'Grady, K., Deussing, M.-A., Scerbina, T., Fung, K. & Muhe, N. (2016). *Measuring up: Canadian Results of the OECD PISA Study - The Performance of Canada's Youth in Science, Reading, and Mathematics – 2015 First Results for Canadians Aged 15*. Council of Ministers of Education Canada. https://www.cmec.ca/Publications/Lists/Publications/Attachments/365/PISA2015-CdnReport-EN.pdf

O'Grady, K., Deussing, M.-A., Scerbina, T., Tao, Y., Fung, K., Elez, V. & Monk, J. (2019). *À la hauteur: Résultats canadiens de l'étude PISA 2018 de l'OCDE – Le rendement des jeunes de 15 ans du Canada en lecture, en mathématiques et en sciences*. Conseil des ministres de l'Éducation (Canada). https://www.cmec.ca/Publications/Lists/Publications/Attachments/396/PISA2018_PublicReport_FR.pdf

OECD (2005a), *Apprendre aujourd'hui, réussir demain: Premiers résultats de PISA 2003*. OECD Publishing. https://doi.org/10.1787/9789264007260-fr

OECD (2005b). *PISA 2003 Technical Report*. OECD Publishing. https://doi.org/10.1787/9789264010543-en

OECD (2014a). PISA 2012 *Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume I, Revised edition, February 2014)*. OECD Publishing. https://doi.org/10.1787/9789264208780-en

OECD (2013). *PISA 2012 Results: Ready to Learn: Students' Engagement, Drive and Self-Beliefs (Volume III)*. OECD Publishing. https://doi.org/10.1787/9789264201170-en

OECD (2014b). *PISA 2012 Technical Background*. OECD Publishing. https://doi.org/10.1787/6341a959-en

OECD (2016). *Résultats du PISA 2015: L'excellence et l'équité dans l'éducation* (vol. 1). OECD Publishing. https://doi.org/10.1787/9789264267534-fr

OECD (2017). *Recommandation de 2013 du Conseil de l'OCDE sur l'égalité hommes-femmes en matière d'éducation, d'emploi et d'entrepreneuriat*. OECD Publishing. https://doi.org/10.1787/9789264279407-fr

Park, D., Ramirez, G. & Beilock, S. L. (2014). The role of expressive writing in math anxiety. *Journal of Experimental Psychology: Applied*, *20*(2), 103-111. https://doi.org/10.1037/xap0000013

Parsons, S. & Bynner, J. (2005). *Does numeracy matter more?* National Research and Development Centre for Adult Literacy and Numeracy.

Passolunghi, M. C., De Vita, C. & Pellizzoni, S. (2020). Math anxiety and math achievement: The effects of emotional and math strategy training. *Developmental science*, *23*(6), e12964 https://doi.org/10.1111/desc.12964

Paulhus, D. L. (2017). Socially desirable responding on self-reports. *Encyclopedia of personality and individual differences*, 1-5.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*, 315-341. https://doi.org/10.1007/s10648-006-9029-9

Raghubar, K. P., Barnes, M. A. & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning Individual Differences*, *20*(2), 110-122. https://doi.org/10.1016/j.lindif.2009.10.005

Ramirez, G., Shaw, S. T. & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist*, *53*(3), 145-164. https://doi.org/10.1080/00461520.2018.1447384

Rasch, G. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests*. Nielsen & Lydiche.

Richardson, F. C. & Suinn, R. M. (1972). The mathematics anxiety rating scale: psychometric data. *Journal of Counseling Psychology*, *19*(6), 551-554. https://doi.org/10.1037/h0033456

Rossi, S., Xenidou-Dervou, I., Simsek, E., Artemenko, C., Daroczy, G., Nuerk, H. C. & Cipora, K. (2022). Mathematics–gender stereotype endorsement influences mathematics anxiety, self-concept, and performance differently in men and women. *Annals of the New York Academy of Sciences*, *1513*(1), 121-139. https://doi.org/10.1111/nyas.14779

Rutkowski, L., Gonzalez, E., Joncas, M. & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, *39*(2), 142-151. https://doi.org/10.3102/0013189X10363170

Ryan, R. M. & Deci, E. L. (2009). Promoting self-determined school engagement. Dans K. R. Wentzel & A. Wigfield (dir.), *Handbook of motivation at school* (p. 171-195). Routledge.

Satorra, A. & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, *66*(4), 507-514. https://doi.org/10.1007/BF02296192

Seegers, G. & Boekaerts, M. (1996). Gender-related differences in self-referenced cognitions in relation to mathematics. J*ournal for Research in Mathematics Education*, *27*(2), 215-240. https://doi.org/10.5951/jresematheduc.27.2.0215

Skaalvik, S. & Skaalvik, E. M. (2004). Gender differences in math and verbal self-concept, performance expectations, and motivation. *Sex roles*, *50*, 241-252. https://doi.org/10.1023/B:SERS.0000015555.40976.e6

Shavelson, R. J., Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. *Review of educational research*, *46*(3), 407-441. https://doi.org/10.3102/00346543046003407

Skinner, C. & Wakefield, J. (2017). Introduction to the design and analysis of complex survey data. *Statistical Science*, *32*(2), 165-175. https://doi.org/10.1214/17-STS614

Sokolowski, H. M., Hawes, Z. & Lyons, I. M. (2019). What explains sex differences in math anxiety? A closer look at the role of spatial processing. *Cognition*, *182*, 193-212. https://doi.org/10.1016/j.cognition.2018.10.005

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology*, *27*(2), 229-239. https://doi.org/10.1111/j.2044-8317.1974.tb00543.x

Stapleton, L. M. (2013). Multilevel *Structural equation modeling* with complex sample data. Dans G. R. Hancock & R. O. Mueller (dir.), *Quantitative methods in education and the behavioral sciences: Issues, research, and teaching. Structural equation modeling: A second course* (p. 521–562). IAP Information Age Publishing.

Steenkamp, J. B. E. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of consumer research*, *25*(1), 78-90. https://doi.org/10.1086/209528

Steinmetz, H. (2011). Estimation and comparison of latent means across cultures. Dans Davidov, P. Schmidt & J. Billiet (dir.), *Cross-cultural analysis: Methods and applications* (p. 85–116). Routledge/Taylor & Francis Group.

Stoet, G., Bailey, D. H., Moore, A. M. & Geary, D. C. (2016). Countries with higher levels of gender equality show larger national sex differences in mathematics anxiety and relatively lower parental mathematics valuation for girls. *PloS One*, *11*(4), Article e0153857. https://doi.org/10.1371/journal.pone.0153857

Stokke, A. (2015). What to do about Canada's declining math scores? *C.D. Howe Institute Commentary 427*. https://doi.org/10.2139/ssrn.2613146

Szczygiel, M. (2020). Gender, general anxiety, math anxiety and math achievement in early school-age children. *Issues in Educational Research*, *30*(3), 1126-1142. https://search.informit.org/doi/10.3316/informit.465488906598804

Tucker, L. R. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1-10. https://doi.org/10.1007/BF02291170

Vandenberg, R. J. & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, *3*(1), 4-70.

Viau, R. (2009). *La motivation en contexte scolaire* (2e éd.). De Boeck.

Vohl, P. (2023). *Étude des écarts d›anxiété mathématique selon le genre et des facteurs ayant le potentiel de les réduire, chez les élèves québécois francophones de 15 ans ayant participé au PISA de 2003 et de 2012* [thèse de doctorat, Université de Montréal]. Papyrus. https://papyrus.bib.umontreal.ca/xmlui/handle/1866/32456

Vohl, P. & Loye, N. (2023). Portrait des écarts d'anxiété mathématique selon le genre et du lien anxiété mathématique/performances en mathématiques, chez les élèves québécois francophones de 15 ans ayant participé au PISA de 2003 et de 2012. *Mesure et évaluation en éducation*, *46*(2), 128-170. https://doi.org/10.7202/1111101ar

Vohl, P. & Loye, N. (2024). Les défis liés à l'analyse secondaire de données issues des évaluations à grande échelle en éducation. Dans N. Loye & N. Duroisin (dir.), *Évaluation, apprentissage et numériques* (p. 355-398). Peter Lang.

Vukovic, R. K., Kieffer, M. J., Bailey, S. P. & Harari, R. R. (2013). Mathematics anxiety in young children: Concurrent and longitudinal associations with mathematical performance. *Contemporary Educational Psychology*, *38*(1), 1-10. https://doi.org/10.1016/j.cedpsych.2012.09.001

Wang, J. & Wang, X. (2020). *Structural equation modeling: Applications using Mplus*. Wiley.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, *54*(3), 427-450. https://doi.org/10.1007/BF02294627

Widaman, K. F. & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. Dans K. J. Bryant, M. Windle & S. G. West (dir.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (p. 281–324). American Psychological Association. https://doi.org/10.1037/10222-009

Wigfield, A. & Eccles, J. S. (2000). Expectancy–value theory of achievement motivation. *Contemporary Educational Psychology*, *25*(1), 68-81. https://doi.org/10.1006/ceps.1999.1015

Winne, P. H. & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology*, *27*(4), 551-572. https://doi.org/10.1016/S0361-476X(02)00006-1

Appendix A

Table A1

Main descriptive statistics for the variables studied in the sample of French-speaking Quebec participants - 2003

Sample	Variable	n	Mean	Variance	Min	Max	Q1	Median	Q3	Skewness	Kurtosis
	GENDRE = male	2151	0.49	0.25	0	1	-	-	-	-	-
	ANXMAT	2097	-0.10	0.83	-2.37	2.55	-0.47	0.06	0.56	-0.30	0.56
	INTMAT	2069	-0.02	0.87	-1.78	2.29	-0.66	0.00	0.58	0.05	-0.31
	INSTMOT	2055	0.29	0.96	-2.30	1.59	-0.39	0.05	1.11	-0.33	-0.59
	SCMAT	2099	0.25	0.97	-2.18	2.26	-0.29	0.18	0.88	-0.16	-0.05
2003	PVMATH1	2151	536.50	8 557.58	200.53	784.50	472.92	544.98	602.46	-0.31	-0.13
	PVMATH2	2151	535.36	8 741.12	189.63	790.96	474.17	541.47	601.99	-0.27	-0.04
	PVMATH3	2151	535.58	8 634.77	195.70	856.47	473.16	539.99	602.70	-0.23	-0.03
	PVMATH4	2151	535.12	8 836.34	197.41	790.81	475.73	538.74	601.06	-0.27	0.06
	PVMATH5	2151	536.63	8 596.11	214.55	789.64	473.94	542.72	602.62	-0.29	-0.13
	PVMATH	2151	535.84	8 673.20	225.92	766.19	473.98	541.58	602.17	-0.27	-0.05

Note. ANXMAT = mathematics anxiety, INTMAT = intrinsic value of mathematics, INSTMOT = utility value of mathematics, SCMAT = mathematics self-concept, PV1 to PV5 = plausible values 1 to 5, PVMATH = five plausible values combined, Skewness = coefficient of skewness, Kurtosis = coefficient of kurtosis. Results from personal analysis.

Table A2

Main descriptive statistics for the variables studied in the sample of French-speaking Quebec participants - 2012

Sample	Variable	n	Mean	Variance	Min	Max	Q1	Median	Q3	Skewness	Kurtosis
2012	GENDRE = male	2385	0.49	0.25	0	1	-	-	-	-	-
	ANXMAT	1526	0.02	0.95	-2.37	2.55	-0.47	0.06	0.60	-0.27	0.39
	INTMAT	1574	-0.02	0.94	-1.78	2.29	-0.66	0.00	0.58	0.04	-0.52
	INSTMOT	1573	0.32	1.03	-2.30	1.59	-0.39	0.50	1.11	-0.41	-0.67
	SCMAT	1524	0.31	1.10	-2.18	2.26	-0.29	0.41	1.12	-0.11	-0.38
	PVMATH1	2385	536.44	8 230.45	200.51	828.41	474.61	541.60	600.49	-0.18	-0.12
	PVMATH2	2385	536.78	8 033.33	225.43	839.16	476.01	541.06	599.95	-0.16	-0.18
	PVMATH3	2385	537.09	8 170.37	191.94	860.97	476.56	541.60	599.24	-0.20	-0.09
	PVMATH4	2385	537.05	8 278.41	187.26	777.62	477.73	541.60	602.28	-0.23	-0.17
	PVMATH5	2385	537.67	8 125.60	197.31	825.92	477.81	542.30	600.18	-0.19	-0.06
	PVMATH	2385	537.01	8 167.64	225.74	805.35	476.55	541.63	600.42	-0.19	-0.12

Note. ANXMAT = mathematics anxiety, INTMAT = intrinsic value of mathematics, INSTMOT = utility value of mathematics, SCMAT = mathematics self-concept, PV1 to PV5 = plausible values 1 to 5, PVMATH = five plausible values combined, Skewness = coefficient of skewness, Kurtosis = coefficient of kurtosis. Results from personal analysis.

Table A3

Percentage of missing data in the sample of French-speaking Quebec students who participated in PISA 2003 and in the sample of French-speaking Quebec students who participated in PISA 2012

Variable	2003	2012
GENDRE = male	0.00%	0.00%
ANXMAT	2.49%	36.02%
INTMAT	3.93%	34.00%
INSTMOT	4.56%	34.05%
SCMAT	2.42%	36.10%
PV1	0.00%	0.00%
PV2	0.00%	0.00%
PV3	0.00%	0.00%
PV4	0.00%	0.00%
PV5	0.00%	0.00%

Note. ANXMAT = mathematics anxiety, INTMAT = intrinsic value of mathematics, INSTMOT = utility value of mathematics, SCMAT = mathematics self-concept, PV1 to PV5 = plausible values 1 to 5. Results from personal analysis.