INTERNATIONAL JOURNAL OF MATHEMATICAL EDUCATION IN SCIENCE AND TECHNOLOGY
2024, VOL. 55, NO. 3, 650–690
https://doi.org/10.1080/0020739X.2021.2022229

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Investigating the importance of girls' mathematical identity within United States STEM programmes: a systematic review

Felix Fernandez, Merle Froschl, Lara Lorenzetti and Maryann Stimmer

FHI 360, Washington, DC, USA

ABSTRACT

We conducted a systematic review to generate a comprehensive understanding of cisgendered girls' mathematical identity within the United States to better understand: (a) how mathematical identity is associated with girls' participation, engagement, and achievement in mathematics; (b) the factors related to girls' mathematical identity; and (c) strategies for developing strong mathematical identities within girls. We identified 76 articles that examined mathematical identity in young girls and women. Our review suggests that positive girls' mathematical identity was consistently linked with increased performance, participation and persistence in mathematics. Factors and strategies affecting girls' mathematical identity included the role of parental attitudes and beliefs; societal and cultural influences; and the intersections of sex and race. Practical implications for researchers and practitioners are also discussed.

ARTICLE HISTORY
Received 2 October 2020

KEYWORDS
Mathematical identity; systematic review; STEM

1. Introduction

Through this systematic review of girls' mathematical identity literature we seek to further explore the relationship between mathematical identity and STEM-related outcomes, provide deeper understanding of the factors that may moderate or mediate mathematical identity, identify strategies and practices shown to be most effective in promoting girls' mathematical identity, and document the overall level of evidence associated with each. Note that this review focuses specifically on cisgender girls' mathematical identity; the articles included in this review do not explore other types of gender identities. We begin by describing our theoretical approach to mathematical identity which attempts to define or operationalize a common measure to direct the focus of our study. The following three research questions guide our review:

(1) What is the relationship between mathematical identity and girls' participation, engagement, and achievement in mathematics?
(2) What factors moderate and mediate the complexities of girls' mathematical identities?
(3) What are effective strategies for developing positive mathematical identities in girls?

CONTACT Felix Fernandez ffernandez@fhi360.org FHI 360, 1825 Connecticut Avenue, NW, Washington, DC 20009, USA

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

1.1. Theoretical approach to mathematical identity

Our approach to mathematic identity encompasses multiple theoretical frameworks and perspectives building from student beliefs, identity, agency, and social perspectives. Mathematical identity refers to a person's beliefs, attitudes, emotions, and dispositions about mathematics and their resulting motivation and approach to learning and using mathematics knowledge (Martin, 2000). It includes students' beliefs on whether or not they can do mathematics and that they belong in mathematics (Boaler & Greeno, 2000). However, self-perception alone is not sufficient for the development of a mathematics identity – recognition (i.e. to be recognized by others as a 'mathematics person') and interest are also paramount (Cribbs et al., 2015). Students' identities as learners and doers of mathematics are developed as they participate in activities inside and outside the classroom, as well as by the cognitive and social roles students can take (Lave & Wenger, 1991; Nasir & Hand, 2008). In this way mathematical identity is a fluid and dynamic concept developed through social practice, by which that learners develop a sense of who they are. More specifically, mathematical identity is essential to students' beliefs about themselves as learners and as potential mathematicians (Solomon, 2007). It involves the ways students think about themselves in relation to mathematics and the extent to which they have developed a commitment to, are engaged in, and see value in mathematics themselves (Cobb et al., 2009).

However, our theoretical approach to mathematical identity, as well as the field in general, still lacks a singular math identity measure. Often informed by multiple theoretical approaches, mathematical identity has been inconsistently defined with multiple, and often competing, measures in use (Darragh, 2016; Gresalfi & Hand, 2019; Langer-Osuna & Esmonde, 2017). Moreover, we do not yet possess a clear understanding of the strategies that show the most promise for building a positive mathematical identity in girls, nor a clear understanding of the quality and strength of this evidence, partly due to the lack of consistent measurement. Therefore, we propose the following two criteria to operationally define a measure of mathematical identity that will direct the focus of our study and potentially inform the field:

(a) a student's belief, attitudes, emotions, and disposition about mathematics (e.g. I like mathematics; I am confident I can succeed in my mathematics course; I am better/worse at mathematics compared to the rest of the class; I am a mathematics person; others see me as a mathematics person); and
(b) a student's motivation, value, and desire to pursue mathematics (e.g. it is important for me to be good at mathematics; it is important that I do well in mathematics; I am interested in a career in mathematics).

1.2. Rationale for this review

A comprehensive systematic review of mathematical identity is needed, not only to provide a synthesis of the related research on girls' and women's mathematical identities but to characterize the overall strength and quality of the evidence therein. There remains a need for a validated and consistent definition of mathematical identity. It is our hope that by providing an operationalize definition of a measure of mathematical identity the results of this review will help establish the relationship between mathematical identity

and STEM-related outcomes, provide a deeper understanding of the factors that may moderate or mediate those effects, and identify strategies and practices shown to be most effective in promoting positive mathematical identity. The findings will provide insight into designing STEM programmes and interventions that address inequity and promote positive mathematical identities.

2. Method

This systematic review was designed by a multidisciplinary team composed of education experts and review methodologists. We applied standard methodology for conducting systematic reviews as defined in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, developed by a working group of review methodologists to ensure transparency and adequate reporting within systematic reviews (Liberati et al., 2009). This review was also structured based on guidance from the Campbell Collaboration (The Campbell Collaboration, 2019), the premier organization in developing systematic reviews – specifically within the social sciences – that are comprehensive and methodologically sound. The use of these standard guidelines is meant to reduce bias throughout each stage of the systematic review process and improve the validity of the findings relative to girls' mathematical identity presented in this review.

2.1. Definition of terms

We used two criteria to determine if a study included a measure of student mathematical identity (a) a student's belief, attitudes, emotions, and disposition about mathematics; and (b) a student's motivation, value, and desire to pursue mathematics. A measure of mathematical identity may include a composite measure (i.e. separate measures on student attitudes and values towards mathematics), or a single measure of mathematical identity that addresses both student attitudes and values (e.g. I see myself as a mathematics person; others see me as a mathematics person). Additionally, for the purpose of our review, measures could be quantitative or qualitative in nature (e.g. from survey or interview data). By limiting our review to studies that met these criteria we hoped to provide a consistent and systematic response to our proposed research questions as well as provide validation for our proposed operational definition of a measure of mathematical identity.

To address our research questions, we examined studies that included the following outcomes of interest: academic achievement (e.g. national, state, local assessments, and other mathematics test); non-cognitive measures (e.g. student self-concept, self-efficacy, perseverance, motivation, and dispositions towards math); retention in or pursuit of the STEM career pathway (i.e. pursuit or completion of STEM-related degrees and careers); and community and collaboration outcomes (e.g. teacher and parental engagement, attitudes, beliefs, peer roles, and community influences).

2.2. Information sources & search terms

From December 2017 through February 2018, we conducted a systematic search of the peer-reviewed literature in eight databases: Academic Search Premier, Education

Full Text, Education Resource Information Center, Family and Society Studies, GenderWatch, PsycINFO, Social Work Abstracts, and Women's Studies International. Databases were selected for their focus on the fields of education, psychology, sociology, and gender/women studies and as to minimize duplication of entries. We also included literature from an annotated bibliography that was developed under a National Science Foundation funded project (#1348524), A Capacity Building Project to Advance Research on Girls' Mathematical identity: Improving STEM Learning and Broadening Participation (2014–2017) to (a) explore knowledge and gaps in the research about how girls develop a mathematical identity, and (b) identify studies on girls' mathematical identity.

The electronic search strategy was developed, tested, and refined under the guidance of a reference librarian. In each database, we used Boolean operators to combine three key concepts and restricted search terms to the title, abstract, keywords, and subject terms. We developed search terms based on the key concepts from our research questions, described in detail in the Background above. The key concept areas that comprised our search terms were:

(1) **Math*:** Captures all forms of math (mathematics, mathematical, etc.), including phrases like 'science, technology, engineering and math'.
(2) **'STEM':** Captures STEM intervention programmes that may not have directly indicated mathematics.
(3) **Identity terms:** Captures relevant math identity terms, including, identity, identities, perception, perceptions, perspective, perspectives, perceive, perceives, perceiving, attitude, attitudes, disposition, dispositions, belief, beliefs, and self-concept.
(4) **Outcome terms:** Captures relevant achievement, engagement, participation, and persistence terms, such as, achieve, achieves, achieving, achievement, achievements, engage, engages, engaging, engagement, participate, participates, participating, participation, and persistence.

The standard search combined the three key concepts using Boolean operators and restricted search terms to the title, abstract, keywords, and subject terms as follows:

(Math$*$ or "STEM") + (Identity Terms) + (Outcome Terms)

2.3. Eligibility criteria

We used nine inclusion criteria against all literature identified through our database search and in the annotated bibliography (see Table 1). Given differences in culture, gender norms, and education systems across countries, we sought to limit the scope of this review by focusing on studies conducted within the United States to ensure findings would be relevant within the U.S. education system.

We also included studies that examined mathematical identity among girls or provided sex disaggregated findings. If a study did not disaggregate its findings between girls and boys, it was excluded from the study. The review team considered this a necessary distinction in order to facilitate our understanding of mathematical identity specifically for *girls*. For studies that examined girls exclusively, we looked for and captured certain within-group differences as relevant to research question 3. As described in Table 1, studies were

Table 1. Inclusion criteria and rationale.

Inclusion criteria	Rationale
1. Study setting in the United States	To avoid cultural and population differences between countries and ensure findings are directly relevant to programme and policy-making in the United States
2. Focus on girls or substantive analysis of difference between girls and boys	To ensure findings were directly related to girls' performance and not the role of mathematical identity in general
3. Students in grades pre-kindergarten through post-secondary school	To evaluate mathematical identity in girls and women across the age/grade spectrum
4. Study includes a measure of mathematical identity	Articles required a measure of mathematical identity aligned with the review's definition to appropriately address study research questions.
5. Study examines relevant outcomes of interest: academic achievement, post-secondary outcomes, engagement, persistence, family or teacher engagement	Articles required at least one of the key outcomes of interest. Depending on the research question, this may include only a measure of mathematical identity.
6. Source is an empirical study of STEM education, (e.g. policy, practice, or intervention programme)	Articles must report on empirical research, with either primary or secondary data analysis.
7. Peer-reviewed publication or high-quality grey literature	To reduce publication bias, grey literature was included if of comparable quality to peer-reviewed literature (i.e. includes methods, appropriate interpretations).
8. Source is in English	Articles limited to reviewers' first language
9. All publication dates	A time bound was not applied to allow for a more comprehensive review.

also excluded if they did not focus on early childhood through post-secondary school populations, did not include an appropriate measure of mathematical identity, did not include an outcome of interest, were not an empirical study, or were not provided in English.

2.4. Study selection & consistency checks

The screening process was completed in two stages. First, articles were assessed for eligibility based on title and abstract, at which point they were deemed relevant or irrelevant. In the second stage, reviewers read the full texts of articles to determine inclusion or exclusion from the review. Reviewers also noted the specific research question(s) to which each article mapped.

Prior to the title/abstract and full-text phases, the research team conducted pre-testing on a sub-set of 20 articles to train reviewers on eligibility criteria and appropriate approaches for sorting articles. Reviewers were provided with a decision-tree composed of the various eligibility criteria to assist with study selection. After an individual assessment of each article, the team discussed the group's decisions. Lead investigators used these pilots as an opportunity to clarify concepts, procedures, and decision tools to ensure reproducibility.

At both stages, each article was assessed by two reviewers, final determinations on all conflicts were made by the review team's education experts, particularly as most conflicts concerned the concept of identity as related to mathematics.

2.5. Analytic approach

For all articles included in the study, full-texts were exported to NVivo for coding. We extracted information on the following: location, grade level, intervention, study design,

outcomes of interest, effect sizes, and major limitations. Given the variety in outcomes as well as measurement approaches for specific outcomes, we did not pursue a quantitative analysis of effects across articles. Instead, for each study question, we provide a narrative summary of findings, which are moderated by the risk of bias assessment described below. We also provide these summaries in tabular format to more easily compare findings and interpretations across included studies.

2.6. Risk of bias and quality of evidence assessment

We conducted a risk of bias assessment to evaluate the internal validity of individual studies and understand the implications of this bias on the overall quality of evidence for outcomes. The risk of bias for each article informs the overall quality of the evidence of a given outcome, or in this case, study question. Risk of bias were determined separately for quantitative and qualitative studies. Specifically, we used the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) approach (Schunemann et al., 2013) for assessing bias in randomized and observational studies (Schunemann et al., 2013), the former of which is harmonized with the Cochrane Collaboration's tool for assessing a studies risk of bias (Higgins et al., 2019). For all studies using quantitative methods, including experimental and observational (or non-experimental) studies, reviewers assessed the design and prevailing limitations, ultimately providing a rating for each article. We utilized two assessments: one for randomized controlled trials and a second for quasi-experimental or non-experimental studies. Per the guidance, articles received either a low, unclear, or high risk of bias. A rating of 'unclear' did not necessarily mean the limitations could not be assessed; rather, there was evidence of some bias that raised doubt about the study results. Low risk of bias meant that the identified study bias was unlikely to alter results, whereas high risk of bias weakened confidence in results.

Similarly, for qualitative studies, we assessed the methodological limitations for each study and determined a risk of bias. We used the Critical Appraisal Skills Programme checklist for assessing qualitative study limitations (Critical Appraisal Skills Programme, 2018; Munthe-Kaas et al., 2018). This 10-item assessment examines appropriateness of qualitative methods, including descriptions of recruitment, data analysis, researcher role, and ethical procedures. GRADE' Confidence in the Evidence from Reviews of Qualitative Research (Lewin et al., 2018) guidance suggests providing each article with a rating of very minor concerns, minor concerns, moderate concerns, or serious concerns. To harmonize these ratings with the quantitative risk of bias scores, we equated very minor or minor concerns to low risk of bias, moderate concerns to moderate risk of bias, and serious concerns to high risk of bias.

Finally, once individual articles received a risk of bias rating, they were grouped by outcome to analyze quality of studies. After all articles were assessed, we examined the body of evidence against GRADE's four-tiered system (i.e. high, moderate, low, or very low quality) to rate the quality of evidence for main study outcomes separately. The GRADE approach does not suggest that you 'average' the individual risk of bias scores but rather qualitatively weigh the various limitations to assign a quality score. The research team discussed study limitations and determined quality of evidence for each outcome.

3. Results

We identified 7,142 unique articles from eight databases and an annotated bibliography of girls' mathematical identity (Figure 1). Two reviewers screened the title/abstract for each article, at which point 4,261 were considered irrelevant. Two reviewers then read full texts of remaining articles, ultimately excluding 2,803 with reasons. One-third were excluded because they did not focus on girls or disaggregate findings by sex, while nearly one-quarter did not include an appropriate measure of mathematical identity as we have operationally defined. Seventy-six peer-reviewed articles were included in this review.

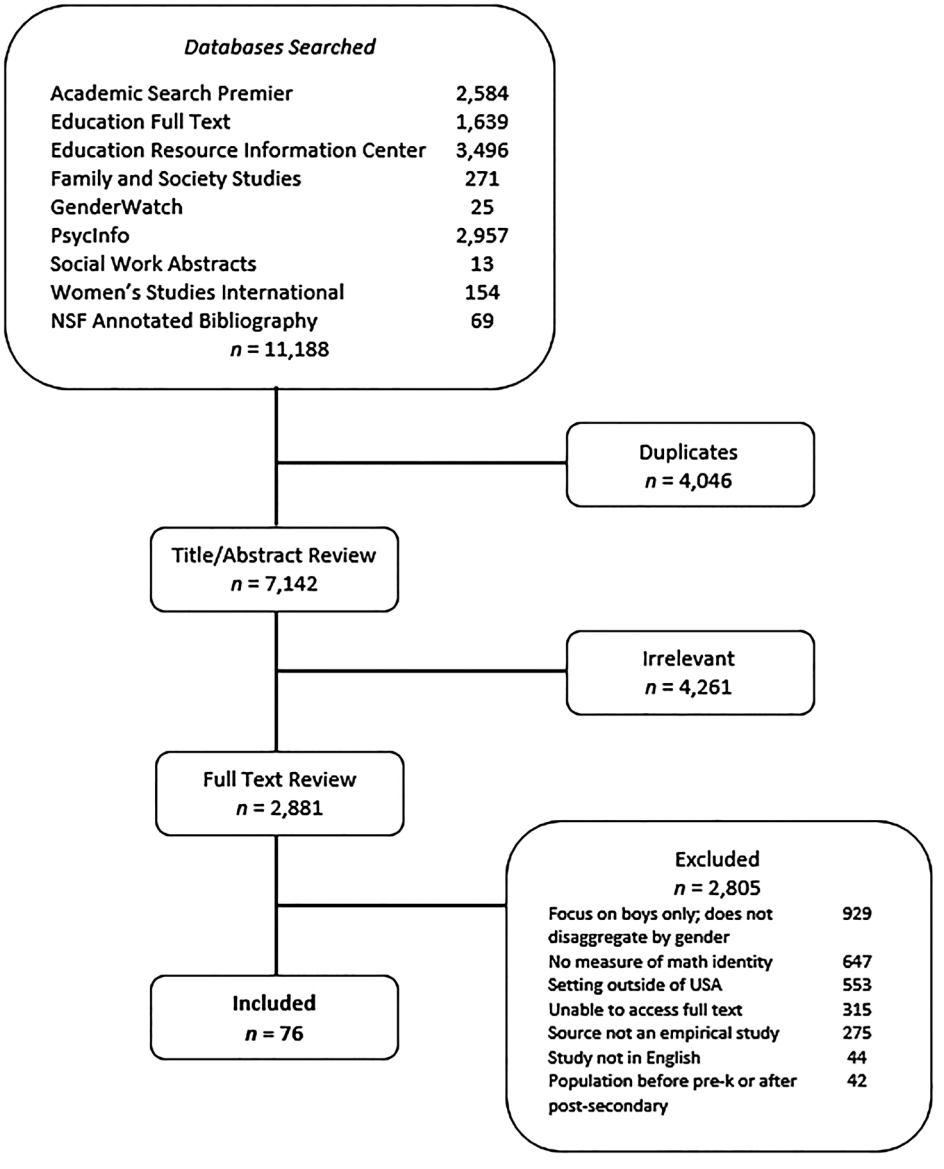


Figure 1. PRISMA diagram.

3.1. Study characteristics

We collected evidence spanning 48 years, with 22% of the literature emerging in the past five years. Most articles (79%) examined mathematical identity in both girls and boys, providing a disaggregated analysis of findings by sex (see Table 2). Roughly half of articles included middle school students or high school students as the study population. No studies examined mathematical identity in early education. Most studies were conducted in the mid-western United States; data from the Northeast or national surveys comprised the second and third largest geographic areas. Non-cognitive measures, including self-concept and engagement, were the most common study outcomes. In terms of design, most studies applied a non-experimental approach (67%), which included any quantitative designs that did not incorporate a control group. Very few studies used experimental or quasi-experimental approaches. We also incorporated mixed-methods (9%) and qualitative studies (8%). For the summary table of studies, please see Appendix A.

Table 2. Study characteristics.

Study characteristics	Count	%
Period		
1998 or earlier	30	39.5
1999–2003	2	2.6
2004–2008	10	13.2
2009–2013	17	22.4
2014–2018	17	22.4
Sex		
Female and male	60	78.9
Female only	16	21.1
Region (USA)		
Midwest	33	43.4
Northeast	12	15.8
Southern	11	14.5
West	11	14.5
National	12	15.8
Age group		
Early education (Pre-K)	0	0.0
Elementary (K–5)	7	9.2
Middle school (6–8)	35	46.1
High school (9–12)	34	44.7
College	20	26.3
Study design		
Non-experimental	51	67.1
Quasi-experimental	5	6.6
Randomized controlled trial	7	9.2
Mixed methods	7	9.2
Qualitative	6	7.9
Outcome measures		
Academic achievement	33	43.4
Non-cognitive measures	76	100.0
Retention or pursuit of STEM	23	30.3

*Counts and percentages may not equal study *n* as categories were not mutually exclusive.

3.2. RQ#1: What is the relationship between mathematical identity and girls' participation, engagement, and achievement in mathematics?

Thirty-four articles addressed our first question. Among these articles, we were most interested in identifying the relationship of mathematical identity with academic achievement in mathematics and girls' retention in the STEM pipeline. For inclusion in this review, studies were not required to specifically reference mathematical identity. Instead, they needed to provide one or more measures that met our criteria for mathematical identity. As such, studies addressed student mathematical identity in varying ways. The most commonly included measures that, when combined, met our definition of mathematical identity were mathematics self-concept, mathematics self-efficacy, mathematics interest, mathematics anxiety, and perceived usefulness and/or importance of mathematics. Individually, these concepts measure student self-perceptions related to mathematics, but when measured together, combinations were operationalized as a measure of mathematical identity.

Mathematics self-concept, which is considered a broader assessment of the skills and abilities one possesses, was positively related to mathematics achievement in several studies (Bachman, 1970; Else-Quest et al., 2013; Liu, 2009; Simpkins et al., 2006). Mathematics self-efficacy, or one's confidence in successfully applying their abilities, was also positively correlated with performance. As self-efficacy increased, so did mathematics achievement and performance (Fennema & Sherman, 1977, 1978; Simpkins et al., 2006). Interest in or liking mathematics was also consistently positively correlated with mathematics performance (Benbow & Stanley, 1982; Simpkins et al., 2006). However, measures of self-concept, self-efficacy, and interest only met half of our criteria for mathematical identity. To meet our definition of mathematical identity, measures of perceived usefulness or importance were also necessary. For example, a person may like and be good at mathematics, but not necessarily want to engage or actively participate in mathematic fields and pursuits. As an example, Benbow and Stanley (1982) found that girls' liking of mathematics was significantly correlated with their mathematics performance but not their ratings of the importance of mathematics in their future careers.

Several studies looked at the perceived usefulness of mathematics, which was positively correlated with mathematics achievement (Fennema & Sherman, 1977, 1978; Pedro et al., 1981). Importance placed on mathematics was also positively linked to performance and continue engagement in mathematics (Meece et al., 1990; Simpkins et al., 2006; Watt et al., 2006). For example, Meece et al. (1990) found that perceived mathematics importance had a direct positive effect on students' performance expectations, and an indirect positive effect on student performance. This suggests that perceived importance of mathematics may be more important to girls. For example, Simpkins et al. (2006) found that 10th grade girls' feelings about the importance of mathematics were more strongly associated with mathematics grades than they were for male peers.

3.2.1. Academic achievement

In total, 22 studies provided a measure that met our definition of mathematical identity alongside a measure of mathematics achievement or performance. Overall, the notion that positive mathematical identity is associated, directly or indirectly, with achievement was prevalent (Bachman, 1970; Bandalos et al., 1995; Benbow & Stanley, 1982; Brown & Leaper, 2010; Elliott, 1990; Else-Quest et al., 2013; Ercikan et al., 2005; Ethington, 1992; Fennema &

Sherman, 1977, 1978; Greene et al., 1999; Gresky et al., 2005; Iben, 1991; Lawrence & Charbonneau, 2009; Liu, 2009; Meece et al., 1990; Pedro et al., 1981; Sherman, 1979, 1980b; Simpkins et al., 2006; Tartre & Fennema, 1995; Watt et al., 2006). Despite most studies having an unclear or high ROB, the overwhelming majority of studies found measures of mathematical identity to be significantly positively correlated with mathematics achievement (Bachman, 1970; Else-Quest et al., 2013; Fennema & Sherman, 1977), computational test (Fennema & Sherman, 1978), SAT-math scores (Benbow & Stanley, 1982), the number of math-items correct (Gresky et al., 2005), and mathematics grades (Pedro et al., 1981; Simpkins et al., 2006). Moreover, measures of mathematical identity within these studies were also found to significantly predict girls' mathematics achievement (Bandalos et al., 1995; Ercikan et al., 2005; Greene et al., 1999; Iben, 1991; Liu, 2009; Meece et al., 1990; Sherman, 1979; Tartre & Fennema, 1995; Watt et al., 2006), mathematics grades (Brown & Leaper, 2010; Sherman, 1980b), and the number of math-items correct (Lawrence & Charbonneau, 2009). Combined, these studies point to the link between higher mathematical identity and better mathematics achievement and performance in girls and women. Only one study failed to find a positive relationship between mathematical identity and mathematics achievement (Elliott, 1990).

3.2.2. Retention or pursuit of STEM pipeline

Sixteen studies explored the role of mathematical identity in girls' intent, retention, or pursuit of STEM-related fields (Ackerman et al., 2013; Burkley et al., 2010; Eccles & Wang, 2015; Ercikan et al., 2005; Ethington, 1991; Fredricks et al., 2018; Good et al., 2012; Meece et al., 1990; Pedro et al., 1981; Sherman, 1981, 1982, 1983a; Sherman & Fennema, 1977; Wang et al., 2015; Watt et al., 2006; Whigham, 1988). While the role of mathematical identity in girls' and women's retention or pursuit of STEM fields is complex, and most studies were of unclear ROB, studies suggest that positive mathematical identities influence students' intent to pursue mathematics. For example, measures of math identification were significantly correlated with pursuing mathematics coursework, majors, and career (Ackerman et al., 2013; Burkley et al., 2010; Ercikan et al., 2005; Ethington, 1991; Fredricks et al., 2018; Meece et al., 1990; Pedro et al., 1981; Sherman, 1981, 1982, 1983a; Sherman & Fennema, 1977; Watt et al., 2006). Moreover, three studies also reported that girls may disengage from mathematics due to lower mathematical identities (Ackerman et al., 2013; Fredricks et al., 2018; Wang et al., 2015).

For example, Watt et al. (2006) found that girls' 10th grade mathematical identity (measured by expectancy beliefs, intrinsic value, and importance) directly influenced 11th and 12th grade mathematics course participation. Sherman and Fennema (1977) found that 10th and 11th grade girls intending to continue taking mathematics course work had higher mathematical identity measures (confidence, attitudes, usefulness, and effectance motivation) than girls not intending to continue. Meanwhile, Fredricks et al. (2018) reported that girls' lower perceived enjoyment, importance, usefulness, and cost of participating in mathematics resulted in them reporting less interest in mathematics careers or a desire to participate in mathematics outside of school. Indeed, perceived usefulness or importance of mathematics was specifically highlighted as one of the most influential predictors for girls' persistence in mathematics (Meece et al., 1990; Pedro et al., 1981; Sherman, 1981, 1982; Sherman & Fennema, 1977). Additionally, Whigham (1988) reported that women in introductory engineering college courses were more likely than

men to have stronger attitudes towards success in mathematics, more positive perceptions of their teacher attitudes towards them as learners of mathematics, and were less likely to see mathematics as a male domain.

3.2.3. Risk of bias and quality of evidence for RQ#1

We assessed the risk of bias and quality of evidence for the thirty-four studies examining the relationship between girls' mathematical identity and our main outcomes of academic achievement and pursuit of the STEM pipeline (see Table 3). After qualitatively weighing the various limitations of each study we assigned a low to moderate quality of evidence for our RQ#1 findings. For studies examining academic achievement, just one was considered low risk, 13 had an unclear risk of bias, and an additional 8 had a high risk of bias. Taking this into consideration, we consider this evidence to be of low to moderate quality. There was a similar breakdown for studies examining the relationship between mathematical identity and the pursuit of the STEM pipeline. In both cases, the primary cause for an unclear or high risk of bias was due to the inclusion of studies that employed non-experimental designs. For example, the majority of studies under this outcome were either cross-sectional studies examining correlations or longitudinal cohort studies. Despite some advanced analytic techniques, many studies did not adequately control for confounding variables nor include a validated measure of mathematical identity.

3.3. RQ#2: What factors moderate and mediate the complexities of girls' mathematical identity?

Our review suggests that the development of girls' mathematical identity is complex and interacts with numerous factors as girls progress through school. These factors include the role of parental attitudes and beliefs; the effects of sex roles and mathematics stereotypes; influence of race and ethnicity; the effects of mathematics anxiety; influence of age groups; societal and cultural influences; and the influence of student disabilities. In total, 46 articles met our study criteria and addressed this research question.

Table 3. Risk of bias and quality of evidence assessments for RQ#1*.

Research question	Count	Design (#)	Risk of Bias (#)			Quality of evidence
			Low	Unclear	High	
All RQ1	34	• Cross-sectional (19) • Longitudinal cohort (12) • Mix-methods (1) • Pre-post, no control (2) • Randomized controlled Trial (1)	1 (3%)	21 (62%)	12 (35%)	Low to Moderate
Academic achievement	22	• Cross-sectional (13) • Longitudinal cohort (6) • Pre-post, no control (2) • Randomized controlled Trial (1)	1 (5%)	13 (59%)	8 (36%)	Low to Moderate
Retention or pursuit of STEM pipeline	16	• Cross-sectional (8) • Longitudinal cohort (8) • Mix-methods (1)	–	11 (69%)	5 (31%)	Low to Moderate

*Counts will not sum to total as categories were not mutually exclusive.

3.3.1. The role of parental attitudes and beliefs

One theme that constantly surfaced was the importance and weight of parental attitudes and beliefs, with 16 studies including measures of mathematical identity and parental influences (Casad et al., 2015; Dickens & Cornell, 1993; Eisenberg & Schneider, 2007; Ercikan et al., 2005; Ethington, 1991, 1992; Fennema & Sherman, 1977, 1978; Frome & Eccles, 1998; Gniewosz et al., 2015; Leedy et al., 2003; McGee, 2015; Sherman, 1983b; Simpkins et al., 2012; Stevenson & Newman, 1986; You & Sharkey, 2012). Parental attitudes and beliefs were strongly related to students' mathematical achievement and mathematical identity (Dickens & Cornell, 1993; Ercikan et al., 2005; Ethington, 1991; Frome & Eccles, 1998)—so much so that student perception of parental attitudes was found to have a significant direct effect on the value of mathematics for girls (Ethington, 1991). However, studies have repeatedly reported that parents demonstrate lower expectations of their daughters' mathematics ability (Ethington, 1991; Fennema & Sherman, 1977, 1978; Frome & Eccles, 1998; Gniewosz et al., 2015). For example, Gniewosz et al. (2015) found that parents' perception of student mathematics competences (i.e. their belief that their child is good at mathematics, finds mathematics easy, and is doing well in mathematics) was significantly lower for girls than for boys. Moreover, Fennema and Sherman (1978) reported these differences in girls' perception of their parent attitudes began as early as middle school and continued to grow into high school.

Notably different were the dynamics involved in mother-daughter and father-daughter relationships. Mother-daughter attitudes and beliefs were repeatedly linked to girls' mathematical identity (Casad et al., 2015; Frome & Eccles, 1998; Simpkins et al., 2012; Stevenson & Newman, 1986). Although the direction of this relationship was complicated, studies suggested that mothers' perceptions were positively related to their daughter's mathematics attitudes and achievement (Casad et al., 2015; Frome & Eccles, 1998). For example, Casad et al. (2015) found that when mothers' and daughters' mathematics anxiety were both high, daughters had more negative outcomes compared to mothers and daughters whose mathematics anxiety were both low. Complicating this relationship is the study by Stevenson and Newman (1986), which found that elementary aged girls (grades 2, 3, and 5) given the highest mathematics ability ratings by their mothers were least likely to perceive mathematics as a valuable subject and more likely to perceive mathematics as a difficult subject by tenth grade. Father-daughter relationships were equally interesting, with girls more likely to consult their fathers than their mothers for help with mathematics (Sherman, 1983b). You and Sharkey (2012) also underscored the importance of father expectations in that they were related to advanced mathematics course-taking patterns of girls in 10th grade. Perhaps somewhat related, Leedy et al. (2003) found that fathers held higher expectations for their daughters in mathematics, compared to fathers' expectations of their sons in mathematics.

3.3.2. Sex roles and stereotypes

Fifteen studies discussed the importance of mathematics, sex, and racial stereotypes on student mathematical identity (Allen & Eisenhart, 2017; Brown & Leaper, 2010; Casad et al., 2015; Eccles & Wang, 2015; Ethington, 1991, 1992; Fennema & Sherman, 1977, 1978; Gilbert et al., 2014; Good et al., 2012; Greene et al., 1999; Gresky et al., 2005; Leedy et al., 2003; McGee, 2015; Steinberg et al., 2012). Uniformly, these studies agreed that perceptions of academic sexism or stereotypes were linked to lower mathematics attitudes and achievement. The more girls endorse perceived stereotypes, the stronger the negative relationship.

Casad et al. (2015) found that a stronger endorsement of math stereotypes was associated with greater anxiety; lower predicted mathematics GPA, intentions, attitudes; and greater mathematics devaluing. The more girls perceived mathematics as a male field, the less likely they were to continue in the area (Ethington, 1991), or expend effort (Greene et al., 1999). Fennema and Sherman (1978) found that sex-role attitudes were a more important predictor of girls' learning of mathematics than boys' learning of mathematics, in both middle and high school.

Moreover, Gallagher and De Lisi (1994) suggested that implicit academic stereotypes based on sex accounted for a significant portion of the variance in girls' mathematical identity. Steinberg et al. (2012) even found a main effect for stereotype threat on a 15-item mathematics performance measure above and beyond the role of mathematics identification. Findings were not limited to sex-based stereotypes but also those related to mathematics and STEM fields. For example, Eccles and Wang (2015) found that girls' entry into STEM fields was undermined by the stereotype that STEM fields require more time and commitment toward work over family needs than do non-STEM fields. However, Allen and Eisenhart (2017) found that girls were able to successfully identify with mathematics by accepting the validity of negative stereotypes and using those features to distinguish themselves as better compared to their peers. Minority girls may face a double threat when pursuing mathematics, in that they face stereotypes related both to their sex and their race. For example, Brown and Leaper (2010) found that perception of academic sexism was strongly linked with lower perceived competence of young Latinx girls (aged 13–15 years), and that older Latinx girls (aged 16–18 years) who perceived more than a few instances of academic sexism felt less competent in mathematics than similarly aged girls with fewer perceptions of academic sexism, controlling for grades in mathematics.

3.3.3. Influences of race and ethnicity

Eleven studies examined the influence of student race/ethnicity (Allen & Eisenhart, 2017; Brown & Leaper, 2010; Campbell, 1991; Catsambis, 1994; Else-Quest et al., 2013; Hendy et al., 2014; Iben, 1991; Ing & Nylund-Gibson, 2017; McClendon & Wigfield, 1998; McGee, 2015; Navarro et al., 2007). Findings on the influence of race/ethnicity on girls' mathematical identity were inconsistent yet pointed to an increased risk for Black or Latinx girls. For example, Brown and Leaper (2010) found significant ethnic differences between Latinx and White girls, with Latinx girls reporting less perceived competence and valuing of mathematics than White girls, and Catsambis (1994) found that Black and Latinx girls were least likely to aspire to careers in mathematics compared to their male counterparts. In contrast, Else-Quest et al. (2013) and Hendy et al. (2014) found no significant differences by sex between ethnic groups on measures of mathematical identity.

Several studies also looked at the inequities within ethnicities. When compared with male counterparts, by the eighth grade, Black and Latinx girls were the least likely to aspire to careers in mathematics (Catsambis, 1994) and Latinx girls had significantly less confidence in their mathematics skills (Catsambis, 1994; Navarro et al., 2007). Moreover, seventh through ninth grade Black girls had lower beliefs and expectations about their mathematics ability than Black males (McClendon & Wigfield, 1998). Iben (1991) found that for a small sample ($n = 41$) of Black girls aged 13 to14 years, beliefs on the usefulness of mathematics and its importance to their future career and education plans was a significant predictor of mathematic performance.

3.3.4. The effects of mathematics anxiety

Nine studies explored the influence of mathematics anxiety (Bandalos et al., 1995; Brown & Leaper, 2010; Casad et al., 2015; Eisenberg & Schneider, 2007; Good et al., 2012; Lent et al., 1996; Liu, 2009; Meece et al., 1990; Pedro et al., 1981) on girls' beliefs, engagement, and achievement in mathematics. Increased mathematics anxiety was linked to deficits in performance (Casad et al., 2015; Liu, 2009; Pedro et al., 1981). For example, Liu (2009) found that 15 year old girls reported greater levels of anxiety compared to male peers, which was in turn negatively related to their performance on a mathematics assessment. Lent et al. (1996) found that girls in an introductory college course were more likely than males to mention anxiety reactions when asked what they considered when making confidence ratings in relation to mathematics courses. However, perceived mathematics ability, importance, and self-concept were found to have direct negative effects on mathematics anxiety (Bandalos et al., 1995; Meece et al., 1990), suggesting that greater mathematical identity may reduce anxiety in girls.

3.3.5. Influences of age group

Eight studies looked at age group differences (Brown & Leaper, 2010; Casad et al., 2015; Ganley & Lubienski, 2016; Ing & Nylund-Gibson, 2017; Lauermann et al., 2017; Sherman, 1980a, 1982; Simpkins et al., 2006). These studies presented a nuanced view of mathematical identity, suggesting that differences in mathematical identity began as early elementary school and continued to grow through middle school, becoming somewhat stable as students enter high school (particularly for White girls). For example, Ganley and Lubienski (2016) found differences by sex in mathematics confidence, interest, and performance as early as third grade, and that differences continued into fifth and eighth grades. Casad et al. (2015) found that mathematics attitudes became more negative as girls progressed from sixth to eighth grade. Simpkins et al. (2006) found age group differences in mathematics grades and perceived competence with Latinx girls aged 13 to 15 years reporting higher grades and more perceived competence in mathematics and science than Latinx girls aged 16 to 18 years. Not only do mathematics attitudes become more negative as girls progress through elementary and middle school, these attitudes demonstrate high stability once they enter high school. For example, Sherman (1980a) and Simpkins et al. (2006) found that measures of girls' mathematical identity (e.g. self-concept, confidence, interest, usefulness, and importance of mathematics) evidenced moderate to high stability across 9^{th} to 12^{th} grade and 6^{th} to 10th grade, respectively. It is important to note that Brown and Leaper (2010) focused on Latinx populations, and both Sherman (1980a) and Simpkins et al. (2006) studied predominantly White samples.

3.3.6. Societal and cultural influences

Eight studies discussed the importance of peer, school, and cultural influences (Allen & Eisenhart, 2017; Eisenhart & Allen, 2016; Good et al., 2012; Navarro et al., 2007; Simpkins et al., 2006; Walton et al., 2012; Wang et al., 2015; You & Sharkey, 2012). Whether it is a matter of cultural orientation influencing performance (Navarro et al., 2007), the influence of out-of-school time experiences (Simpkins et al., 2006), or sense of belonging (i.e. feelings of membership and acceptance; Good et al., 2012; Walton et al., 2012; Wang et al., 2015) the context in which a girl found herself had an impact on her mathematical identity.

Reviewed studies suggested that girls may place stronger value on the importance of others. For example, You and Sharkey (2012) found that peer influences (i.e. among your close friends, how important is it that they attend classes regularly, study, continue education past high school) for grades and education were significantly related to girls' advanced mathematics course-taking but not for boys. Wang et al. (2015) found that girls' altruism values (i.e. the importance attributed to having an active role in helping others in their communities) mediated the effect on STEM occupations. Specifically, girls' higher altruism values were negatively correlated with occupations in STEM.

Local school practices and beliefs may inadvertently affect students' mathematical identities. For example, school practices or histories may perpetuate existing stereotypes regarding mathematics: that mathematics is a typically masculine domain, that advanced mathematics courses are reserved for specific students, and that minority students are typically underachieving (Allen & Eisenhart, 2017). Schools may also inadvertently encourage weak mathematical identities that make students vulnerable to future disruptions and challenges. Eisenhart and Allen (2016) found that schools that stressed being good at mathematics (i.e. making good grades, finishing assignments quickly, and taking advanced courses) created hollowed-out mathematical identities that could be easily disrupted as mathematics courses increased in difficulty. For example, one girl in their qualitative study reported that her placement in the most advanced 11^{th} grade mathematics course originally reinforced her mathematical identity. However, she soon found the coursework difficult and time consuming, and she also felt out of favour with her mathematics teacher. Taken together, these affected her mathematical identity and future career interest. Moreover, Eisenhart and Allen (2016) suggested these practices (i.e. emphasizing performance and enrolment in advanced courses), which are more likely in large, urban public schools, may severely compromise high-achieving, well-behaved, aspiring low-income minority students' chances of success.

3.3.7. Influences of student disability

Lastly, only one study was found on the influence of student disability on mathematical identity. Eisenberg and Schneider (2007) found that for third grade girls diagnosed with attention-deficit/hyperactivity disorder (ADHD), mathematical identity (measured as self-perceptions of aptitude and interest) was not significantly different than non-ADHD diagnosed girls. However, both their parents' and teachers' perceptions of student mathematics skills were substantially more negative.

3.3.8. Risk of bias and quality of evidence for RQ#2

The abovementioned evidence should be interpreted with the following risk of bias assessments in mind. For evidence included under this research question, 30 studies (nearly two-thirds) were determined to have an unclear risk of bias, while 11 were rated as high risk of bias and 5 were rated as low risk of bias (see Table 4). In general, this pattern held for each area of interest, with literature on sex roles and stereotypes as well as societal and cultural influences having some low risk of bias evidence. Similar to the RQ#1, we determined this evidence base to be of low to moderate quality. For example, the majority of studies were non-experimental, including cross-sectional or pre-and post-test studies; however, many did apply advanced analytic techniques to reduce bias.

Table 4. Risk of bias and quality of evidence assessments for RQ#2*.

Research question	Count	Design (#)	Risk of Bias (#)			Quality of Evidence
			Low	Unclear	High	
All RQ2	46	• Cross-sectional (20) • Longitudinal cohort (16) • Mix-methods (1) • Pre-post, no control (2) • Quasi-experimental design (1) • Qualitative (3) • Randomized controlled Trial (3)	5 (11%)	30 (65%)	11 (24%)	Low to Moderate
The role of parental attitudes and beliefs	16	• Cross-sectional (9) • Longitudinal cohort (5) • Mix-methods (1) • Pre-post, no control (1) • Quasi-experimental design (1) • Qualitative (1)	–	10 (63%)	6 (38%)	Low to Moderate
Sex roles and stereotypes	15	• Cross-sectional (9) • Longitudinal cohort (1) • Quasi-experimental design (1) • Qualitative (2) • Randomized controlled Trial (2)	3 (20%)	6 (40%)	6 (40%)	Low to Moderate
Influences of race and ethnicity	11	• Cross-sectional (7) • Longitudinal cohort (2) • Qualitative (2)	1 (9%)	6 (55%)	4 (36%)	Low to Moderate
The effects of mathematics anxiety	9	• Cross-sectional (7) • Longitudinal cohort (1) • Pre-post, no control (1)	–	7 (78%)	2 (22%)	Moderate
Influences of age group	8	• Cross-sectional (2) • Longitudinal cohort (6)	–	7 (88%)	1 (13%)	Moderate
Societal and cultural influences	7	• Cross-sectional (2) • Longitudinal cohort (2) • Qualitative (2) • Randomized controlled Trial (1)	3 (43%)	3 (43%)	1 (14%)	Moderate
Influences of student disability	1	• Cross-sectional (1)	–	1 (100%)	–	Moderate

*Counts will not sum to total as categories were not mutually exclusive.

3.4. RQ#3: What are effective strategies for developing mathematical identities in girls?

There have been many attempts to find the most effective strategies for developing a positive mathematical identity in girls. In all, 18 studies discussed strategies to develop a positive mathematical identity in girls; eight focused on middle school, five on high school, and six on post-secondary.

3.4.1. Strategies promoting retention or pursuit of STEM pathways

Eight studies examined potential mathematical identity strategies promoting the retention or pursuit of STEM and math related fields (Anderson, 2005; Cordero et al., 2010; Fredricks et al., 2018; Gallagher & De Lisi, 1994; Genshaft & Hirt, 1980; Piatek-Jimenez, 2015; Schmader et al., 2004; Weinberg et al., 2011). As empowered learners of mathematics who perceived themselves as confident, capable, independent, and persistent, girls have an improved chance of participating and remaining in mathematics (Anderson, 2005). However, they need to see the link between what they are learning in mathematics and the careers they are planning (Fredricks et al., 2018). Girls desired to learn how mathematics can be applied to other fields and to everyday life activities. Enjoyment of mathematics also

played a role in women's persistence in the field (Piatek-Jimenez, 2015). However, a survey of women majoring in math-related fields revealed that women who believed that status differences between the sexes are legitimate were more likely to endorse stereotypes about women's mathematics abilities. This, in turn, predicted more negative self-perceptions of mathematics competence and less interest in continuing study in one's field (Schmader et al., 2004).

3.4.2. Constructivist learning strategies

A recurring theme in the development of girls' mathematical identity was a focus on constructivist learning techniques. Fives studies examined the effects of constructivist learning programmes (Anderson, 2005; Fredricks et al., 2018; Frost & Wiest, 2007; Redmond et al., 2011; Weavers et al., 2011), where learning occurs as learners are actively involved in a process of meaning and knowledge construction as opposed to passively receiving information. Anderson (2005) discussed this in terms of agency and authorship based on girls' experiences in a feminist mathematics class, which valued feminine voices and connected to girls' experiences. Agency is the process by which the learner acts on her own behalf to learn mathematics, and authorship refers to what and how mathematics learning is valued. Through this, girls in this study saw that they had a different relationship to mathematics, one in which the classroom became a site of empowerment.

Several studies (Anderson, 2005; Fredricks et al., 2018; Weavers et al., 2011) showed that girls became more engaged in student-centered classrooms with hands-on learning, group work, real-world applications and problem-solving activities. Collaborative environments, with cooperative group work and authentic instruction were particularly beneficial for girls' engagement (Fredricks et al., 2018; Frost & Wiest, 2007). Fun was also a factor (Weavers et al., 2011). However, in studies involving both girls and boys, interesting differences emerged. Expectancy beliefs – whether they thought they understood the content and could do well – were more strongly linked to boys' cognitive engagement in science but were a stronger predictor of girls' emotional engagement in mathematics (Fredricks et al., 2018). In studies where self-efficacy for both improved, the gain was greater for males (Fredricks et al., 2018). In another, girls' confidences in their own skills and potential was significantly more positive than the boys' confidence in the girls (Redmond et al., 2011).

Girls relied more heavily than boys on conventional solution strategies to mathematics problems (Gallagher & De Lisi, 1994) which could be caused by lowered confidence. Combining performance accomplishment and belief-persistence had a greater effect on self-efficacy for boys than girls (Cordero et al., 2010), indicating that ingrained messages for girls and women about mathematics may need a lengthier intervention. Weinberg et al. (2011) studied the effectiveness of experiential learning at residential summer programmes and found that interest gain, utility gain, cost value gain, attainment value gain and expectancy for success were equally successful for both sexes.

3.4.3. Single versus mixed-sex classrooms

Four studies analyzed the effect of single-sex vsersus mixed-sex classrooms (Brown & Ronau, 2012; Else-Quest & Peterca, 2015; Kombe et al., 2016; Reid & Roberts, 2006). There was mixed evidence on the usefulness of single-sex versus mixed-sex classrooms.

Whether classes were single-sex or mixed-sex had little influence on self-concept and self-efficacy (Brown & Ronau, 2012; Kombe et al., 2016). However, Reid and Roberts (2006) found that after participating in a single-sex programme that integrated mathematics and social science research, girls demonstrated greater confidence in their mathematics ability. Single-sex classes were linked to negative effects for boys' and positive effects for girls' mathematics confidence and achievement, with authors concluding that both girls and boys tend to fare better in the company of girls (Else-Quest & Peterca, 2015).

3.4.4. Mentors and role models

Three studies looked at mentoring and role modelling strategies (Redmond et al., 2011; Reid & Roberts, 2006; Stout et al., 2011). These studies showed that a mentoring relationship improved girls' mathematics confidence (Redmond et al., 2011; Reid & Roberts, 2006; Stout et al., 2011). Contact with same-sex experts enhanced women's self-concept and self-efficacy and boosted implicit positive attitudes toward STEM (Dickens & Cornell, 1993). Genshaft and Hirt (1980) found that mathematics anxiety in girls during testing was alleviated with positive modelling and verbal reinforcement.

3.4.5. Combatting stereotype threat

Three studies examined the negative effect of stereotype threat on girls' mathematical identity, in particular their self-perception, self-confidence, and career intentions (Christy & Fox, 2014; Gresky et al., 2005; Schmader et al., 2004). In a study of a virtual classroom with leaderboards that had either female or male majority names, Christy and Fox (2014) found that a social comparison process had a stronger influence than stereotype threat. In another study, Gresky et al. (2005) found that the influence of stereotype threat can be lessened through exercises in self-concept that underscore women's many social roles and identities.

These studies also looked at the connection between stereotype threat and academic performance. Christy and Fox (2014) found that women exposed to a female majority leaderboard actually performed more poorly on a subsequent mathematics test. Another study found that highly math-identified women scored significantly worse than highly math-identified men on a subsequent Graduate Record Examination-like mathematics test, but when they participated in an exercise that reminded them of their many roles and identities, they scored as well as those men (Gresky et al., 2005). Schmader et al. (2004) revealed that endorsing stereotypes can also have negative implications for women's performance on tests of their mathematics ability by making them more susceptible to stereotype threat. More specifically, they found that making gender identity salient while taking a mathematics test led to lower test performance only among women who tended to endorse the stereotype and had no effect on women who firmly rejected the stereotype.

3.4.6. Parental and teacher encouragement

Two studies examined the relationship between parental or teacher attitudes or beliefs and student mathematical identity (Anderson, 2005; Piatek-Jimenez, 2015). In her description of a feminist mathematics class, Anderson (2005) highlights the importance of redefining power relations. Educators in the programme redefined their power and became guides in the process instead of instructors. Their role was not to impart knowledge, but to guide the girls as they took power to construct their own knowledge. Piatek-Jimenez (2015) reported

that individual encouragement and recognition by parents and teachers continued to be instrumental in some women's interest and persistence in mathematics.

3.4.7. Risk of bias and quality of evidence for RQ#3

By necessity, this research question called for more robust designs to examine effective strategies for developing girl's mathematical identity. Here the overall risk of bias was lower signalling a higher quality body of evidence (see Table 5). Nine studies were rated as unclear risk of bias, and 7 were rated as having a low risk of bias. The remaining 2 studies were considered high risk of bias. Within these studies, we saw a greater number of quasi- or experimental studies, which typically had a lower risk of bias. Ultimately, we determined the body of evidence for all of RQ#3 to be of moderate to high quality, although quality of evidence varied across reported strategies.

3.5. Limitations

We applied a systematic approach to reduce bias in identifying relevant literature; however, certain limitations should be considered. First, mathematical identity is a complex latent construct that has been defined and measured inconsistently over time. We built our search strategy around commonly used terminology in the peer-reviewed literature, and although our search yielded over 11,000 hits, it is possible certain articles were not captured. Additionally, our review team was extensively trained to identify mathematical identity measures, but some articles may have been inadvertently excluded on this criterion due to reviewer bias. We consider this risk to be low, given that all articles were reviewed by two reviewers at both review stages and final decisions on all conflicts were made by an education expert.

Table 5. Risk of bias and quality of evidence assessments for RQ#3*.

Research question	Count	Design (#)	Risk of Bias # (%)			Quality of Evidence
			Low	Unclear	High	
All RQ3	18	• Cross-sectional (1) • Mix-methods (5) • Quasi-experimental design (4) • Qualitative (4) • Randomized controlled Trial (3)	7 (39%)	9 (50%)	2 (11%)	Moderate to High
Strategies promoting retention or pursuit of STEM pathways	8	• Cross-sectional (1) • Mix-methods (2) • Quasi-experimental design (1) • Qualitative (3) • Randomized controlled Trial (2)	3 (38%)	4 (50%)	1 (13%)	Moderate to High
Constructivist learning strategies	5	• Mix-methods (3) • Quasi-experimental design (2)	1 (20%)	2 (40%)	2 (40%)	Low to Moderate
Single versus mixed sex classrooms	4	• Mix-methods (1) • Quasi-experimental design (3)	–	4 (100%)	–	Moderate
Mentors and role models	3	• Mix-methods (2) • Randomized controlled Trial (1)	1 (33%)	2 (67%)	–	Moderate to High
Combatting stereotype threat	3	• Cross-sectional (1) • Quasi-experimental design (1) • Randomized controlled Trial (2)	2 (67%)	1 (33%)	–	Moderate to High
Parental and teacher encouragement	2	• Qualitative (2)	1 (50%)	1 (50%)	–	Moderate to High

*Counts will not sum to total as categories were not mutually exclusive.

Second, we limited this review to literature from the United States to strengthen the conclusions to be drawn for domestic research and practice; therefore, we have not incorporated potentially insightful findings from other countries, including a volume of work from European and Australian researchers. We note that works from some prominent authors in the field were not incorporated, primarily due to studies failing to meet one or more of our study criteria.

Third, while we specifically attempted to highlight representative populations identified within each study, studies overwhelmingly focused on cisgendered middle school girls and boys through post-secondary education in Midwestern states, primarily with rural white populations. Given our own study highlights the intersection of sex and race in the development of positive mathematical identity, it is important that future studies both quantitively and qualitatively explore the development of girls mathematical and gender identity, specifically, as well as how findings, interventions, and strategies differ by race, sex, gender identity, geographic locations, and urbanicity.

Fourth, in our desire to provide a comprehensive review of mathematical identity we did not time bound our review and therefore have most likely included areas of contention or disagreement amongst researchers. For example, forty percent of studies included are before 1998, many of which included the use of Fennema-Sherman mathematics attitude scales, mathematics as a male domain, which researchers today may find no longer valid (Forgasz, Leder, & Gardner, 1999). Another example, our review noted the role of stereotype threat in moderating and mediating girls' mathematical identity, the importance of which has been contested by current researchers (Stoet & Geary, 2012). However, given our intent to support the development of a consistent definition of mathematical identity we felt it important to adhere to a systemic methodology and not time bound our review.

The goal of this review was to generate a comprehensive understanding of girls' mathematical identity, not limited to a specific time period, intervention, or the effect of mathematical identity on specific outcomes of interest. As such, we have generated a large body of evidence that details the relationship between girls' mathematical identity with mathematics performance and persistence, potential moderator and mediators, and effective strategies, but only on girls' mathematical identity. Given our focus, this review does not examine the role of identity development in general or current theoretical work on the importance of identity development in education. This limits our inferences solely to girls and mathematical identity; yet remains true to the goal of this study which was to better understand the complexities of mathematical identity, including any variables that researchers consider as influencing or being influenced by it.

4. Implications for research and practice

Findings suggest that girls' mathematical identities are malleable and susceptible to intervention. These efforts should emphasize the importance of encouraging and supporting early positive attitudes toward mathematics, particularly for ethnically underrepresented girls, because of the impact on outcomes such as achievement and STEM career attainment (Ing & Nylund-Gibson, 2017). Girls pursuing mathematics must overcome gender, domain, and racial stereotypes before successfully identifying with mathematics. Moreover, Black and Latinx girls and women face a double bind in pursuing mathematics, as members of two minority groups. Future research is needed to explore the

intersectionality of gender and race and its effects on the development of mathematical identity

It is clear that gender differences in mathematical identities are formed early and deepen with age, along with the impact of sexism and stereotypes. Ingrained messages about mathematics for girls and women need early, sustained, and proactive intervention. This stereotype threat affects both identity and achievement. Therefore, it is important to begin early to address mathematics confidence and interest and as negative attitudes continue to form. With this in mind, it is interesting to note that only seven studies in this review focused on elementary school age children, while none focused on children in pre-K. Moreover, none of those studies examined effective strategies for developing mathematical identity. There clearly is a need for further research on the strategies and practices that would be effective in promoting mathematical identity in young girls, from pre-K through third grade.

This review underscored the importance and weight of parental attitudes and beliefs, particularly in terms of self-perception. Parental attitudes and beliefs were found to be strongly related to mathematics achievement and identity. In particular, mothers' expectations may be more closely related to girls' mathematical identities (Casad et al., 2015; Frome & Eccles, 1998; Stevenson & Newman, 1986), but fathers' expectations to performance (You & Sharkey, 2012). Teachers' attitudes and school practices can either support a positive identity or perpetuate existing stereotypes. In light of this, some studies (Anderson, 2005; Fredricks et al., 2018) emphasized the importance of professional development to provide training on gender equity issues related to teaching, including how to recognize and prevent gender stereotypes and how to implement more student-centered instructional practice. In addition, findings show that it is important for teachers to provide messages about the relevance of mathematics and science to girls' lives outside the classroom as well as information on the diverse career options that are related to mathematics and science (Fredricks et al., 2018). It is important for parents, teachers, professors, and advisors to encourage and support girls in their interests in mathematics. Ideally this encouragement would begin at a young age and continue throughout their educational careers. Additionally, peer groups may be more important to girls but not boys (You & Sharkey, 2012), strengthening research that suggest girls may avoid mathematics as it is viewed as an individualistic/solitary field.

Moving forward, future studies in girls' mathematical identity should employ stronger and more diverse analytical and methodological designs that consider the use of validated and clearly defined measures of mathematical identity to better understand the links between mathematical identity and improved outcomes, participation, and persistence in mathematics. One of the most challenging parts of the review was identifying when a study did in fact include a measure of mathematical identity. Often studies would include one or more measures of student attitudes or beliefs, self-concept, self-efficacy, and task or value orientation, but mathematical identity, as we have come to define it, is more than any one of these measures. A uniform measure of math identify would then strengthen study rigour and help to support the field as it seeks to improve young girls' and women's mathematics achievement and performance. Lastly, while comparisons between girls and boys may still be of interest, studies designed to improve girls' mathematical identity should focus specifically on female populations and its intersection with race and ethnicity.

5. Conclusion

Studies consistently linked a positive girls' mathematical identity with increases in mathematics performance. There were a large number that measured mathematical identity, and they repeatedly indicated that higher mathematical identity led to better mathematics achievement and performance, increased usefulness or importance of mathematics, and persistence in related fields.

These studies clearly pointed to mathematical identity as a complex pathway with multiple inputs (e.g. from the numerous measures of mathematical identity to the factors influencing girls' mathematical identity) that affects girls' participation and persistence in the STEM pipeline. For example, findings from our review suggest that only women with high levels of mathematical identity will persist in mathematic fields, even though they may have similar cognitive and attitudinal beliefs compared to male peers. Conversely, our study also indicates that girls and women with high levels of mathematical identity are more strongly affected by gender stereotypes, a finding which only increases as they continue to persist in mathematic fields. A similar dynamic is also found regarding parental attitudes, where parents' perceptions of mathematics competency were lower for girls than boys. Addressing these imbalances is crucial to any attempt to promote positive mathematical identities and increase the participation of girls and women in STEM-related fields.

Moreover, studies have also shown that implementing practices that support a positive mathematical identity can lead to improved outcomes, participation, and persistence in mathematics. These studies suggested that numerous intervention programmes have been developed to encourage women to pursue the field of mathematics. Potential strategies that have succeeded to some extent in making many women feel welcome in the discipline include constructivist learning techniques that encourage hands-on, real world applications and problem solving; the use of positive same-sex mentors and role models; and reminding women of their many roles and identities to combat the negative influence of stereotype threat. Despite these findings, our systematic review of mathematical identity indicated several gaps in the current body of literature. First, we did not find a uniform measure of mathematical identity that could help to support the fields as it seeks to improve young girl's and women's mathematics achievement and performance. Second, we found no studies that examined the role of mathematical identity in early education (early childhood learning and Pre-k), and very few that that looked at elementary grade (K-5). Lastly, most of the current studies reviewed employed cross-sectional or longitudinal designs which limit casual attributions and validity. Moving forward, we suggest that research and practitioners develop and validate a uniform measure of mathematical identity and related tools, that can recognize students at-risk for low of negative mathematical identities, determine the root causes of said risk, and identify appropriate and effective strategies or interventions. Additionally, future studies should rigorously examine which factors of mathematical identity (e.g. academics, motivation, societal influences, gender/racial stereotypes) are most important and which strategies and practices are most effective in promoting positive mathematical identity in young girls and women.

Lastly, in reviewing mathematical identity related articles, we were encouraged by the changes in tone and context of how mathematical identity and gender differences were discussed over the last 50 years (Lubienski & Ganley, 2017). While not universal, in many early articles, mathematics was discussed as a possible 'male domain' (Sherman, 1983a),

and questions were raised as to the impact this idea might have on girls' self-perceptions (Fennema & Sherman, 1977; Iben, 1991; Sherman, 1983a). The research has moved from locating the 'problem' within girls, to shifting the lens to teaching practices and the field of mathematics. In more recent literature, the language has changed, and the discussion is around 'sexism' (Brown & Leaper, 2010) as it applies to mathematics and its consequences. Differences are no longer studied in terms of impaired or lower cognitive abilities but rather attitudinal, motivational, and societal factors that influence performance and persistence. Most recently, our understanding of mathematical identity has been broadened through an intersectional analysis of gender and its interplay with race, class, and other socially constructed identities (Joseph, 2017; Leyva, 2017). Until recently, the experiences of underrepresented minorities in mathematics has been largely unexplored (Joseph, 2017), but current research on Black women and girls suggest the importance of identity, structural supports, and community and family influences as key themes in their persistence and resilience in STEM fields (Ireland et al., 2018; Joseph et al., 2017). These shifts are important, as the language we use to discuss and study the performance and persistence of girls and women in STEM fields is critical; it can influence the context for mathematics education and strengthen or weaken perceptions. For example, conversation has shifted from mathematics as a male domain, where boys typically perform better than girls, to mathematics as a gender-neutral domain where girls and boys are equally capable, to what is it about the field of mathematics or how instruction is delivered that perpetuates gender and racial differences and how can these differences be removed.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Felix Fernandez http://orcid.org/0000-0003-2678-8222
Lara Lorenzetti http://orcid.org/0000-0003-4758-7390

References

Ackerman, P. L., Kanfer, R., & Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, *105*(3), 911–927. https://doi.org/10.1037/a0032338

Allen, C. D., & Eisenhart, M. (2017). Fighting for desired versions of a future self: How young women negotiated STEM-related identities in the discursive landscape of Educational opportunity. *Journal of the Learning Sciences*, *26*(3), 407–436. https://doi.org/10.1080/10508406.2017.1294985

Anderson, D. L. (2005). A portrait of a feminist mathematics classroom: What adolescent girls say about mathematics, themselves, and their experiences in a "unique" learning environment. *Feminist Teacher*, *15*(3), 175–194. http://www.jstor.org/stable/40546037

Bachman, A. M. (1970). The relationship between a seventh-grade pupil's academic self-concept and achievement in mathematics. *Journal for Research in Mathematics Education*, *1*(3), 173–179. https://doi.org/10.2307/748335

Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology*, *87*(4), 611–623. https://doi.org/10.1037/0022-0663.87.4.611

Benbow, C. P., & Stanley, J. C. (1982). Intellectually talented boys and girls: Educational profiles. *Gifted Child Quarterly*, *26*(2), 82–88. https://doi.org/10.1177/001698628202600208

Boaler, J., & Greeno, J. G. (2000). Identity, Agency and Knowing in Mathematical Worlds. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 171–200). Ablex.

Brown, C. S., & Leaper, C. (2010). Latina and European American girls' experiences with academic sexism and their self-concepts in mathematics and science during adolescence. *Sex Roles*, *63*(11-12), 860–870. https://doi.org/10.1007/s11199-010-9856-5

Brown, S. L., & Ronau, R. R. (2012). Students' perceptions of single-gender science and mathematics classroom experiences. *School Science and Mathematics*, *112*(2), 66–87. https://doi.org/10.1111/j.1949-8594.2011.00121.x

Burkley, M., Parker, J., Paul Stermer, S., & Burkley, E. (2010). Trait beliefs that make women vulnerable to math disengagement. *Personality and Individual Differences*, *48*(2), 234–238. https://doi.org/10.1016/j.paid.2009.09.002

Butler, M. B. (1999). Factors associated with students' intentions to engage in science learning activities. *Journal of Research in Science Teaching*, *36*, 455–473. https://eric.ed.gov/?id=EJ590245

Campbell Collaboration. (2019). *Campbell systematic reviews: policies and guidelines*. Campbell Policies and Guidelines Series No. 1, Version 1.6. The Campbell Collaboration. https://doi.org/10.4073/cpg.2016.1.

Campbell, J. R. (1991). The roots of gender inequity in technical areas. *Journal of Research in Science Teaching*, *28*(3), 251–264. https://doi.org/10.1002/tea.3660280306

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*, 1597. https://doi.org/10.3389/fpsyg.2015.01597

Catsambis, S. (1994). The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high school. *Sociology of Education*, *67*(3), 199–215. https://doi.org/10.2307/2112791

Christy, K. R., & Fox, J. (2014). Leaderboards in a virtual classroom: A test of stereotype threat and social comparison explanations for women's math performance. *Computers & Education*, *78*, 66–77. https://doi.org/10.1016/j.compedu.2014.05.005

Cobb, P., Gresalfi, M., & Hodge, L. L. (2009). An interpretive scheme for analyzing the identities that students develop in mathematics classrooms. *Journal for Research in Mathematics Education*, *40*(1), 40–68. https://doi.org/10.5951/jresematheduc.40.1.0040

Cordero, E. D., Porter, S. H., Israel, T., & Brown, M. T. (2010). Math and science pursuits: A self-efficacy intervention comparison study. *Journal of Career Assessment*, *18*(4), 362–375. https://doi.org/10.1177/1069072710374572

Cribbs, J. D., Hazari, Z., Sonnert, G., & Sadler, P. M. (2015). Establishing an explanatory model for mathematics identity. *Child Development*, *86*(4), 1048–1062. https://srcd.onlinelibrary.wiley.com/doi/abs/10.1111/cdev.12363. https://doi.org/10.1111/cdev.12363.

Critical Appraisal Skills Programme. (2018). *CASP qualitative checklist*. https://casp-uk.net/wp-content/uploads/2018/01/CASP-Qualitative-Checklist-2018.pdf

Darragh, L. (2016). Identity research in mathematics education. *Educational Studies in Mathematics*, *93*(1), 19–33. https://doi.org/10.1007/s10649-016-9696-5

Doepken, D., Lawsky, E., & Padwa, L. (1993). *Modified Fennema–Sherman attitude scale*. Retrieved August 16, 2006, from http://www.woodrow.org/teachers/math/gender/08scale.html

Dickens, M. N., & Cornell, D. G. (1993). Parent influences on the mathematics self-concept of high ability adolescent girls. *Journal for the Education of the Gifted*, *17*(1), 53–73. https://doi.org/10.1177/016235329301700106

Eccles, J. S., & Wang, M.-T. (2015). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, *40*(2), 100–106. https://doi.org/10.1177/0165025415616201

Eccles, J., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, *21*, 215–225. https://doi.org/10.1177/0146167295213003

Eisenberg, D., & Schneider, H. (2007). Perceptions of academic skills of children diagnosed with ADHD. *Journal of Attention Disorders*, *10*(4), 390–397. https://doi.org/10.1177/1087054706292105

Eisenhart, M., & Allen, C. D. (2016). Hollowed out: Meaning and authoring of high school math and science identities in the context of neoliberal reform. *Mind, Culture, and Activity*, *23*(3), 188–198. https://doi.org/10.1080/10749039.2016.1188962

Elliott, J. C. (1990). Affect and mathematics achievement of nontraditional college students. *Journal for Research in Mathematics Education*, *21*(2), 160–165. https://doi.org/10.2307/749143

Else-Quest, N. M., Mineo, C. C., & Higgins, A. (2013). Math and science attitudes and achievement at the intersection of gender and ethnicity. *Psychology of Women Quarterly*, *37*(3), 293–309. https://doi.org/10.1177/0361684313480694

Else-Quest, N. M., & Peterca, O. (2015). Academic attitudes and achievement in students of urban public single-sex and mixed-sex high schools. *American Educational Research Journal*, *52*(4), 693–718. https://doi.org/10.3102/0002831215591660

Ercikan, K., McCreith, T., & Lapointe, V. (2005). Factors associated with mathematics achievement and participation in advanced mathematics courses: An examination of gender differences from an international perspective. *School Science and Mathematics*, *105*(1), 5–14. https://doi.org/10.1111/j.1949-8594.2005.tb18031.x

Ethington, C. (1991). A test of a model of achievement behaviors. *American Educational Research Journal*, *28*(1), 155–172. https://doi.org/10.3102/00028312028001155

Ethington, C. A. (1992). Gender differences in a psychological model of mathematics achievement. *Journal for Research in Mathematics Education*, *23*(2), 166–181. https://doi.org/10.2307/749499

Fennema, E., & Sherman, J. (1977). Sex-related differences in mathematics achievement, spatial visualization and affective factors. *American Educational Research Journal*, *14*(1), 51–71. https://doi.org/10.3102/00028312014001051

Fennema, E., & Sherman, J. (1978). Sex-related differences in mathematics achievement and related factors: A further study. *Journal for Research in Mathematics Education*, *9*(3), 189–203. https://doi.org/10.2307/748997

Fredricks, J. A., Hofkens, T., Wang, M.-T., Mortenson, E., & Scott, P. (2018). Supporting girls' and boys' engagement in math and science learning: A mixed methods study. *Journal of Research in Science Teaching*, *55*(2), 271–298. https://doi.org/10.1002/tea.21419

Forgasz, H. J., Leder, G. C., & Gardner, P. L. (1999). The fennema-sherman mathematics as a male domain scale reexamined. *Journal for Research in Mathematics Education*, *30*(3), 342–348. https://doi.org/10.2307/749839

Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of Personality and Social Psychology*, *74*(2), 435–452. https://doi.org/10.1037/0022-3514.74.2.435

Frost, J. H., & Wiest, L. R. (2007). Listening to the girls: Participant perceptions of the confidence-boosting aspects of a girls' summer mathematics and technology camp. *The Mathematics Educator*, *17*(2), 31–40. https://www.researchgate.net/publication/237579725

Gallagher, A. M., & De Lisi, R. (1994). Gender differences in scholastic aptitude test: Mathematics problem solving among high-ability students. *Journal of Educational Psychology*, *86*(2), 204–211. https://doi.org/10.1037/0022-0663.86.2.204

Ganley, C. M., & Lubienski, S. T. (2016). Mathematics confidence, interest, and performance: Examining gender patterns and reciprocal relations. *Learning and Individual Differences*, *47*, 182–193. https://doi.org/10.1016/j.lindif.2016.01.002

Genshaft, J. L., & Hirt, M. L. (1980). The effectiveness of self-instructional training to enhance math achievement in women. *Cognitive Therapy and Research*, *4*(1), 91–97. https://doi.org/10.1007/BF01173358

Gilbert, P. N., O'Brien, L. T., Garcia, D. M., & Marx, D. M. (2014). Not the sum of its parts: Decomposing implicit academic stereotypes to understand sense of fit in math and English. *Sex Roles*, *72*(1-2), 25–39. https://doi.org/10.1007/s11199-014-0428-y

Gniewosz, B., Eccles, J. S., & Noack, P. (2015). Early adolescents' development of academic self-concept and intrinsic task value: The role of contextual feedback. *Journal of Research on Adolescence*, *25*(3), 459–473. https://doi.org/10.1111/jora.12140

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, *102*(4), 700–717. https://doi.org/10.1037/a0026659

Greene, B. A., DeBacker, T. K., Ravindran, B., & Krows, A. J. (1999). Goals, values, and beliefs as predictors of achievement and effort in high school. *Sex Roles*, *40*(5/6), 421–458. https://doi.org/10.1023/A:1018871610174

Gresalfi, M., & Hand, V. M. (2019). Coordinating situated identities in mathematics classrooms with sociohistorical narratives: A consideration for design. *ZDM*, *51*(3), 493–504. https://doi.org/10.1007/s11858-019-01034-y

Gresky, D. M., Ten Eyck, L. L., Lord, C. G., & McIntyre, R. B. (2005). Effects of salient multiple identities on women's performance under mathematics stereotype threat. *Sex Roles*, *53*(9-10), 703–716. https://doi.org/10.1007/s11199-005-7735-2

Hendy, H. M., Schorschinsky, N., & Wade, B. (2014). Measurement of math beliefs and their associations with math behaviors in college students. *Psychological Assessment*, *26*(4), 1225–1234. https://doi.org/10.1037/a0037688

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M., & Welch, V. (2019). *Cochrane handbook for systematic reviews of interventions*. Version 6.0 (updated July 2019). Cochrane, 2019. www.training.cochrane.org/handbook.

Iben, M. F. (1991). Attitudes and mathematics. *Comparative Education*, *27*(2), 135–151. https://doi.org/10.1080/0305006910270203

Ing, M., & Nylund-Gibson, K. (2017). The importance of early attitudes toward mathematics and science. *Teachers College Record*, *119*(5), 1–32. https://doi.org/10.1177/016146811711900507

Ireland, D. T., Freeman, K. E., Wintrin-Proctor, C. E., DeLaine, K. D., Lowe, S. M., & Woodson, K. M. (2018). (Un)Hidden figures: A synthesis of research examining the intersectional experiences of black women and girls in STEM education. *Review of Research in Education*, *42*(1), 226–254. https://doi.org/10.3102/0091732X18759072

Joseph, N. (2017). The invisibility of Black girls in mathematics. *Virginia Mathematics Teacher*, *44*(1), 46–52. https://www.researchgate.net/publication/320268286

Joseph, N., Hailu, M., & Boston, D. (2017). Black women's and girls' persistence in the P–20 mathematics pipeline: Two decades of children, youth, and adult education research. *Review of Research in Education*, *41*(1), 203–227. https://doi.org/10.3102/0091732X16689045

Kombe, D., Che, S. M., Carter, T. L., & Bridges, W. (2016). Student academic self-concept and perception of classroom environment in single-sex and coeducational middle grades mathematics classes. *School Science and Mathematics*, *116*(5), 265–275. https://doi.org/10.1111/ssm.12178

Langer-Osuna, J. M., & Esmonde, I. (2017). Identity in research on mathematics education. In J. Cai (Ed.), *Compendium for research in mathmetaics educaiton* (pp. 637–648). National Council of Teachers of Mathematics.

Lauermann, F., Tsai, Y. M., & Eccles, J. S. (2017). Math-related career aspirations and choices within Eccles et al.'s expectancy-value theory of achievement-related behaviors. *Developmental Psychology*, *53*(8), 1540–1559. https://doi.org/10.1037/dev0000367

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Lawrence, J. S., & Charbonneau, J. (2009). The link between basing self-worth on academics and student performance depends on domain identification and academic setting. *Learning and Individual Differences*, *19*(4), 615–620. https://doi.org/10.1016/j.lindif.2009.08.005

Leedy, M. G., LaLonde, D., & Runk, K. (2003). Gender equity in mathematics: Beliefs of students, parents, and teachers. *School Science and Mathematics*, *103*(3), 285–292. https://doi.org/10.1111/j.1949-8594.2003.tb18151.x

Lent, R. W., Brown, S. D., Gover, M. R., & Nijjer, S. K. (1996). Cognitive assessment of the sources of mathematics self-efficacy: A thought-listing analysis. *Journal of Career Assessment*, *4*(1), 33–46. https://doi.org/10.1177/106907279600400102

Lewin, S., Booth, A., Glenton, C., Munthe-Kaas, H., Rashidian, A., Wainwright, M., Bohren, M. A., Tunçalp, Ö, Colvin, C. J., Garside, R., Carlsen, B., Langlois, E. V., & Noyes, J. (2018).

Applying GRADE-CERQual to qualitative evidence synthesis findings: Introduction to the series. *Implementation Science*, *13*(Suppl 1), 2. https://doi.org/10.1186/s13012-017-0688-3

Leyva, L. A. (2017). Unpacking the male superiority myth and masculinization of mathematics at the intersections: A review of research on gender in mathematics education. *Journal for Research in Mathematics Education*, *48*(4), 397–433. https://doi.org/10.5951/jresematheduc.48.4.0397

Liberati, A., Altman, D., Tetzlaff, J., Mulrow, C., Gotzche, P., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Annals of Internal Medicine*, *151*, W-65. https://doi.org/10.3102/0091732X16689045

Liu, O. L. (2009). An investigation of factors affecting gender differences in standardized math performance: Results from U.S. and Hong Kong 15 year olds. *International Journal of Testing*, *9*(3), 215–237. https://doi.org/10.1080/15305050903106875.

Lubienski, S. T., & Ganley, C. M. (2017). Research on gender and mathematics. In J. Cai (Ed.), *Compendium for research in mathmetaics educaiton* (pp. 649–666). National Council of Teachers of Mathematics.

Major, B., & Schmader, T. (1998). Coping with stigma through psychological disengagement. In J. Swim, & C. Stangor (Eds.), *Prejudice: The target's perspective* (pp. 219–241). The Academic Press. https://doi.org/10.1016/B978-012679130-3/50045-4

Martin, D. B. (2000). *Mathematics success and failure among African-American youth: The roles of sociohistorical context, community forces, school influence, and individual agency* (1st ed.). Routledge.

McClendon, C., & Wigfield, A. (1998). Group differences in African American adolescents' achievement-related beliefs about math and science: An initial study. *Journal of Black Psychology*, *24*(1), 28–43. https://doi.org/10.1177/00957984980241003

McGee, E. O. (2015). Robust and Fragile mathematical identities: A framework for exploring racialized experiences and high achievement Among Black college students. *Journal for Research in Mathematics Education*, *46*(5), 599–625. https://doi.org/10.5951/jresematheduc.46.5.0599

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

Munthe-Kaas, H., Bohren, M. A., Glenton, C., Lewin, S., Noyes, J., Tuncalp, O., & Carlsen, B. (2018). Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 3: How to assess methodological limitations. *Implementation Science*, *13*(Suppl 1), 9. https://doi.org/10.1186/s13012-017-0690-9

Nasir, N. S., & Hand, V. (2008). From the court to the classroom: Opportunities for engagement, learning, and identity in basketball and classroom mathematics. *Journal of the Learning Sciences*, *17*(2), 143–179. https://doi.org/10.1080/10508400801986108

Navarro, R. L., Flores, L. Y., & Worthington, R. L. (2007). Mexican American middle school students' goal intentions in mathematics and science: A test of social cognitive career theory. *Journal of Counseling Psychology*, *54*(3), 320–335. https://doi.org/10.1037/0022-0167.54.3.320

Pedro, J. D., Wolleat, P., Fennema, E., & Becker, A. D. (1981). Election of high school mathematics by females and males: Attributions and attitudes. *American Educational Research Journal*, *18*(2), 207–218. https://doi.org/10.3102/00028312018002207

Piatek-Jimenez, K. (2015). On the persistence and attrition of women in mathematics. *Journal of Humanistic Mathematics*, *5*(1), 3–54. https://doi.org/10.5642/jhummath.201501.03

Redmond, A., Scott, M., Thomas, J., Jordan, P., High, K., & Dockers, J. (2011). Enriching science and math through engineering. *School Science and Mathematics*, *111*(8), 399–408. https://doi.org/10.1111/j.1949-8594.2011.00105.x

Reid, P. T., & Roberts, S. K. (2006). Gaining options: A mathematics program for potentially talented at-risk adolescent girls. *Merrill-Palmer Quarterly*, *52*(2), 288–304. https://doi.org/10.1353/mpq.2006.0019

Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles*, *50*(11/12), 835–850. https://doi.org/10.1023/B:SERS.0000029101.74557.a0

Schunemann, H., Brozek, J., Guyatta, G., & Oxman, A. (2013). *GRADE handbook: Handbook for grading the qualit and the strength of recommendations using the GRADE approach*. https://gdt.gradepro.org/app/handbook/handbook.html.

Sherman, J. (1979). Predicting mathematics performance in high school girls and boys. *Journal of Educational Psychology*, *71*(2), 242–249. https://doi.org/10.1037/0022-0663.71.2.242

Sherman, J. (1980a). Mathematics, spatial visualization, and related factors: Changes in girls and boys, grades 8–11. *Journal of Educational Psychology*, *72*(4), 476–482. https://eric.ed.gov/?id=EJ235506

Sherman, J. (1980b). Predicting mathematics grades of high school girls and boys: A further study. *Contemporary Educational Psychology*, *5*(3), 249–255. https://doi.org/10.1016/0361-476X(80)90048-X

Sherman, J. (1981). Girls' and boys' enrollments in theoretical math courses: A longitudinal study. *Psychology of Women Quarterly*, *5*, 5. Supplement. https://doi.org/10.1111/j.1471-6402.1981.tb01092.x

Sherman, J. (1982). Continuing in mathematics: A longitudinal study of the attitudes of high school girls. *Psychology of Women Quarterly*, *7*(2), 132–140. https://doi.org/10.1111/j.1471-6402.1983.tb00825.x

Sherman, J. (1983a). Factors predicting girls' and boys' enrollment in college preparatory mathematics. *Psychology of Women Quarterly*, *7*, 3. https://doi.org/10.1111/j.1471-6402.1983.tb00840.x

Sherman, J. (1983b). Girls talk about mathematics and their future: A partial replication. *Psychology of Women Quarterly*, *7*, 4. https://doi.org/10.1111/j.1471-6402.1983.tb00848.x

Sherman, J., & Fennema, E. (1977). The study of mathematics by high school girls and boys: Related variables. *American Educational Research Journal*, *14*(2), 159–168. https://doi.org/10.3102/00028312014002159

Simpkins, S. D., Davis-Kean, P. E., & Eccles, J. S. (2006). Math and science motivation: A longitudinal examination of the links between choices and beliefs. *Developmental Psychology*, *42*(1), 70–83. https://doi.org/10.1037/0012-1649.42.1.70

Simpkins, S. D., Fredricks, J. A., & Eccles, J. S. (2012). Charting the Eccles' expectancy-value model from mothers' beliefs in childhood to youths' activities in adolescence. *Developmental Psychology*, *48*(4), 1019–1032. https://doi.org/10.1037/a0027468

Solomon, Y. (2007). Not belonging? What makes a functional learner identity in undergraduate mathematics? *Studies in Higher Education*, *32*(1), 79–96. https://doi.org/10.1080/03075070601099473

Sparks, P., Guthrie, C. A.., & Shepherd, R. (1997). The dimensional structure of the perceived behavioral control construct. *Journal of applied social psychology*, *27*, 418–438. https://doi.org/10.1111/j.1559-1816.1997.tb00639.x

Steinberg, J. R., Okun, M. A., & Aiken, L. S. (2012). Calculus GPA and math identification as moderators of stereotype threat in highly persistent women. *Basic and Applied Social Psychology*, *34*(6), 534–543. https://doi.org/10.1080/01973533.2012.727319

Stevenson, H., & Newman, R. S. (1986). Long-term prediction of achievement and attitudes in mathematics and reading. *Child Development*, *57*(3), 646–659. https://doi.org/10.2307/1130343

Stoet, G., & Geary, C. (2012). Can stereotype threat explain the gender gap in mathematics performance and achievemet? *Review of General Psychology*, *16*(1), 93–102. https://doi.org/10.1037/a0026617

Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). Steming the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, *100*(2), 255–270. https://doi.org/10.1037/a0021385

Tartre, L. A., & Fennema, E. (1995). Mathematics achivement and gender: A longitudinal study of selected cognitive and affective variables [grades 6–12]. *Educational Studies in Mathematics*, *28*(3), 199–217. https://doi.org/10.1007/BF01274173

Walton, G. M., & Cohen, G. L. (2011). Sharing motivation. In D.Dunning (Ed.), *Social motivation* (pp. 79–101). Psychology Press. https://doi.org/10.1037/a0025731

Walton, G. M., Cohen, G. L., Cwir, D., & Spencer, S. J. (2012). Mere belonging: The power of social connections. *Journal of Personality and Social Psychology*, *102*(3), 513–532. https://doi.org/10.1037/a0025731

Wang, M. T., Degol, J., & Ye, F. (2015). Math achievement is important, but task values are critical, too: Examining the intellectual and motivational factors leading to gender disparities in STEM careers. *Frontiers in Psychology*, *6*, 36. https://doi.org/10.3389/fpsyg.2015.00036

Watt, H. M. G., Wynarczyk, P., Eccles, J. S., & Durik, A. M. (2006). The leaky mathematics pipeline for girls. *Equal Opportunities International*, *25*(8), 642–659. https://doi.org/10.1108/02610150610719119

Weavers, L. K., Bautista, D. T., Williams, M. E., Moses, M. D., Marron, C. A., & La Rue, G. P. (2011). Assessing an engineering day camp for middle-school girls. *Journal of Professional Issues in Engineering Education and Practice*, *137*(3), 127–134. https://doi.org/10.1061/(ASCE)EI.1943-5541.0000046

Weinberg, A. E., Basile, C. G., & Albright, L. (2011). The effect of an experiential learning program on middle school students' motivation toward mathematics and science. *RMLE Online*, *35*(3), 1–12. https://doi.org/10.1080/19404476.2011.11462086

Whigham, M. A. (1988). *Gender-related differences in engineering students*, 35–45.

You, S., & Sharkey, J. D. (2012). Advanced mathematics course-taking: A focus on gender equifinality. *Learning and Individual Differences*, *22*(4), 484–489. https://doi.org/10.1016/j.lindif.2012.03.005

Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, *80*, 284–290. https://doi.org/10.1037/0022-0663.80.3.284

Appendix A. Summary table of included studies

Table A1. Study summary table.

Study	RoB	Study aim	Mathematical identity measure(s)	Findings
Ackerman et al. (2013)	Unclear risk	Investigate the selection and retention of STEM students and gender differences in STEM persistence	• Self-concept in math • Self-estimate of math abilities	Women who left STEM majors tended to have lower Math/Science Self-Concept, but relatively high levels of the Mastery/Organization trait complex, while men who left STEM majors tended to have lower levels of Mastery/Organization trait complex.
Allen and Eisenhart (2017)	Low risk	Illustrate challenges for young women of colour to construct and maintain STEM identities	• Student views of math and science; characteristics of those who are good at math and science; student views of themselves in relation to those views	Girls successfully identified with mathematics by accepting the validity of negative stereotypes and using those features to distinguish themselves as better compared to their racial and class peers.
Anderson (2005)	Unclear risk	Explore experiences of 7 adolescent girls in a feminist mathematics classroom	• Qualitative descriptions of confidence, agency, and authorship in math	Feminist approach to mathematics (i.e. valuing and connecting to female voices within the math classroom) increased students' sense of authorship and agency.
Bachman (1970)	High risk	Assess relationship between math self-concept and math achievement and strength of this relationship relative to other subjects	• General Self-Concept of Ability Scale (Specific Subject Scale for Math)	Mathematics self-concept was significantly correlated with math achievement.
Bandalos et al. (1995)	Unclear risk	Examine simultaneous impact of attributions for success or failure; self-concept; and perceived self-efficacy on statistics test anxiety	• Benson's math self-concept (1989) • Perceived self-efficacy	Mathematics self-concept was significantly related to women's perceived self-efficacy and test anxiety.
Benbow and Stanley (1982)	High risk	Describe and compare by sex the educational experiences and attitudes of participants in the 1976 Study of Mathematically Precocious Youth (SMPY) talent search	• Perceived status in math • Reported Liking for math • Rated importance of math	Liking mathematics and status in mathematics class were significantly correlated with girls' SAT mathematics score, but not with importance of mathematics in future career.
Brown and Leaper (2010)	High risk	Examine relationships between perceived academic sexism and adolescent girls' perceived competence and valuing of math and science	• Perceived competence in math from Ability Perception subscale of Children's Self- and Task Perception Questionnaire (Eccles & Wigfield 1995) • Valuing of math from Task Value subscale	Student self-reported grades were significantly related to valuing of mathematics and science.

Brown and Ronau (2012)	Unclear risk	Compare student perceptions of single- versus mixed-gender classrooms	• Fennema-Sherman Mathematics Attitude Scale	No significant difference in males' or females' attitudes toward learning math or science when separated by mixed- or single-gender class.
Burkley et al. (2010)	High risk	Examine the role that trait beliefs play in women's disengagement from math	• Mathematical identity (Smith & White)	Math identification significantly correlated with math trait beliefs (i.e. math ability is something you can't change), math importance (being good at math), math enjoyment, likelihood of pursuing a math major, and math career.
Campbell (1991)	Unclear risk	Examine the self-concepts and attributions of talented Asian American and Caucasian students	• Fennema-Sherman Mathematics Attitude Scale • Math Self-Concept Attribute Attitude Scale (SaaS)	Asian American females have higher levels of math self-concept than Caucasian females.
Casad et al. (2015)	Unclear risk	Explore antecedents of math anxiety: parent math anxiety, math-gender stereotypes, and mechanistic perspectives of math anxiety	• Math self-efficacy (Zimmerman & Martinez-Pons, 1988) • Math education behavioral intentions (Sparks et al., 1997, Butler 1999) • Math attitudes (Sparks et al., 1997, Butler 1999) • Math devaluing (Major & Schmader, 1998)	Mother-daughter dyads' math anxiety predicted math self-efficacy, math behavioral intentions, math attitudes, and math devaluing. Daughters with lower math anxiety had more positive math attitudes than daughters with higher math anxiety, and this relationship was strongest when mothers had higher anxiety.
Catsambis (1994)	Unclear risk	Examine the development of gender differences in learning opportunities, achievement, and choice in math	• Choices and attitudes toward mathematics (i.e. aspirations to a mathematics career, attitudes towards mathematics, participation in relevant extra-curricular activities)	Black and Latinx female students were least likely to aspire to mathematics and science careers than male classmates, with Latinx students responding most negatively than any other group of students.
Christy and Fox (2014)	Low risk	Investigate the impact of leaderboards on academic performance	• Mathematical identity (Marsh's Academic Skill Self-Assessment Questionnaire II)	Participants in a female majority leaderboard condition performed more poorly on the math test than those in the male leaderboard condition yet demonstrated a higher level of academic identification than those in the male and control conditions.
Cordero et al. (2010)	Low risk	Evaluate the utility of incorporating belief-perseverance techniques into an intervention to elevate math self-efficacy, interest in, and willingness to engage in STEM courses/careers	• Mathematics self-efficacy (Everyday Math Tasks Subscale) • Math/Science course self-efficacy survey	Prompting performance accomplishment and career belief-persistence had no effect on self-efficacy of females in comparison to a performance accomplishment only control condition.

(*continued*).

Table A1. Continued.

Study	RoB	Study aim	Mathematical identity measure(s)	Findings
Dickens and Cornell (1993)	Unclear risk	Examine the effect of parent expectations and identification on females capable of doing well in math	• Mathematics Attitude Scale (Eccles) • Parent Attitude Scales	Parent expectations were significantly correlated with adolescent mathematics self-concept; moreover, parent mathematical self-concepts had an indirect effect on adolescent self-concept.
Eccles and Wang (2015)	Unclear risk	To identify competencies and motivational factors in high school years that predict future occupational choice	• Ability self-concept in math • Interest value in math	Occupational and lifestyle values, math ability self-concepts, family demographics, and high school course-taking more strongly predicted both individual and gender differences in the likelihood of entering STEM careers than math scores. Gender differences in 12th grade math ability self-concept and valuing a people-oriented versus thing-oriented career mediated the association of gender with STEM occupational choice.
Eisenberg and Schneider (2007)	Unclear risk	Investigate how academic skills of children diagnosed with ADHD are perceived by teachers, parents, and children themselves	• Parent and teacher perceptions of children's math skills • Self-perceptions in math (Marsh's Self-Description Questionnaire)	For ADHD-diagnosed girls compared to other girls, both parents' and teachers' perceptions of self-esteem, motivation, and performance, efforts were substantially more negative.
Eisenhart and Allen (2016)	Low risk	Identify external circumstances that organized student opportunities to become known as good in math or science; determine how high achieving students author identities	• Views of self • Views of math	Schools that stressed being good at math meant making good grades, finishing assignments quickly, taking advanced courses created hallow-out mathematical identities that could be easily disrupted as math courses increased in difficulty.
Elliott (1990)	Unclear risk	Determine if the relationship between selected affective variables and math achievement was different for non-traditional female and male college students compared with traditional	• Modified Confidence in Learning Mathematics Scale (Fennema & Sherman) • Modified Perceived Usefulness of Mathematics Scale (Fennema & Sherman)	After controlling for prior achievement, measures of confidence, attribution, and usefulness were not affective predictors of traditional age female college students.

Else-Quest and Peterca (2015)	Unclear risk	Assess differences in attitudes toward and achievement in reading, math, and science in single-sex schools	• Math self-concept; task value in math; expectations of success in math (Eccles) • Attitudes about math (Programme for International Student Assessment)	Girls in single-sex schooling outperformed girls in mixed math, science, reading, and writing.
Else-Quest et al. (2013)	High risk	Describe math and science attitudes and achievements at the intersection of gender and ethnicity among urban adolescents	• Math self-concept; task value in math; expectations of success in math (Eccles)	Girls' mathematics achievement was significantly correlated with math self-concept and expectations and positively correlated with mathematics value.
Ercikan et al. (2005)	High risk	Understand how personal and environmental variables are associated with achievement and participation in mathematics for females and males	• Beliefs and attitudes toward math; confidence in math	For females in the USA, attitude towards mathematics, confidence in mathematics, self-expectations, friends' expectations, home support for learning, and parents' highest education level were strong predictors of mathematics achievement.
Ethington (1991)	High risk	Determine the degree to which student expectation for success and task value influences achievement behaviour and serves as mediates for the indirect influence of prior constructs	• Value of mathematics; perceived difficulty of math; math as a male domain • Single item measures for math self-concept, goals, expectations for success	There were significant direct effects of stereotyping math, self-concept, goals, and value of math on girls' intention to take more mathematics.
Ethington (1992)	High risk	Describe the extent to which psychological determinants directly influence math achievement and serve as mediators for indirect influence of prior constructs	• Value of mathematics; perceived difficulty of math; math as a male domain • Single item measures for math self-concept, goals, expectations for success	There were significant indirect effects of prior mathematics achievement, stereotyping math, perceptions of parents' attitudes, self-concept, and difficulty of mathematics on girls' mathematics achievement.
Fennema and Sherman (1977)	High risk	Examine sex-related differences in math achievement controlling for previous study of math	• Attitude toward success in math; stereotyping of math as a male domain; perceived attitudes of mother, father, teach; effectance motivation in math; confidence in learning math; usefulness of math	Various factors were significantly correlated with mathematics achievement for girls in grade 9–11: confidence in learning mathematics; math as a male domain; attitudes towards success in math from mother, father, and teachers; usefulness of mathematics; effectance motivation in math; and math related course-taking.

(*continued*).

Table A1. Continued.

Study	RoB	Study aim	Mathematical identity measure(s)	Findings
Fennema and Sherman (1978)	High risk	Explore sex-related differences in math achievement and cognitive and affective variables related to such differences	• Attitude toward success in math; stereotyping of math as a male domain; perceived attitudes of mother, father, teach; effectance motivation in math; confidence in learning math; usefulness of math	Various factors were significantly correlated with math computation test scores: girls' confidence in math; math as a male domain; attitudes towards success from mother, father, teacher; usefulness of math; and effectance motivation.
Fredricks et al. (2018)	High risk	Explore and test student perceptions of motivational and contextual factors influencing engagement by gender	• Math and Science Engagement Scales (Wang) • Student motivational beliefs	Girls' lower perception of the task value of math and science resulted in girls being less likely to report an interest in a math and science career or a desire to learn about math and science outside of school.
Frome and Eccles (1998)	Unclear risk	Investigate (1) the mediating role of parental perceptions between grades and adolescents' self-perceptions, (2) gendered nature of parents' perceptions	• Adolescent self-concept of math ability; perception of difficulty in math; expectancies for success in math	Parents' perceptions were more strongly related to children's self- and task perceptions than children's grades in math (for mothers only).
Frost and Wiest (2007)	Low risk	Examine how girls of diverse racial/ethnic and socioeconomic backgrounds perceived the characteristics of an intervention to develop confidence in math skills	• Qualitative descriptions of confidence and perseverance in math	Cooperative group work and authentic instruction were beneficial for girls' engagement in mathematics.
Gallagher and De Lisi (1994)	Unclear risk	Examine whether male and female students of high mathematical ability use different solution strategies on math problems	• Fennema-Sherman Mathematics Attitudes Scale	High-ability girls relied more heavily than boys on conventional solution strategies for math problems.
Ganley and Lubienski (2016)	Unclear risk	Explore gender differences and reciprocal relations in student math confidence, interest, and achievement during elementary and middle school	• Math confidence • Math interest	Elementary school girls feel less confident in their math abilities than boys, with disparities in math confidence being greater than differences in achievement or interest at the 3rd and 5th grades.
Genshaft and Hirt (1980)	Low risk	Examine if the performance of high test-anxious individuals can be improved by directing attention to task-relevant cues and repressing critical self-evaluative ruminations	• Fennema-Sherman Mathematics Attitude Scale	Self-instructional training significantly increased 7th grade girls' preferences for mathematics and sciences.
Gilbert et al. (2014)	Unclear risk	Decompose math and English stereotypes to examine the relationship between four components of stereotypes and student sense of fit in math and English	• Math Identification Questionnaire (Brown and Josephs)	Gender, implicit academic stereotypes, and their interactions account for a significant portion of the variance in sense of fit in math.

Gniewosz et al. (2015)	Unclear risk	Investigate the extent to which achievement-related feedback predicts early adolescents' domain-specific academic self-concept and intrinsic task values	• Academic self-concept in math • Intrinsic task value	Parents' perception of student math competencies was significantly lower for females than for males.
Good et al. (2012)	Unclear risk	Explore Sense of Belonging to Math measure and confirm factor structure identified in exploratory analysis	• Sense of Belonging to Math • Psychological sense of school membership; social connectedness; stigma consciousness; gender-based rejection sensitivity; mathematical identity • Intent to pursue math; interest in math	Mathematical identity was not a significant predictor of math anxiety, math confidence, or intent to pursue mathematics after controlling for female students sense of belonging, psychological sense of school membership, gender-based rejection sensitivity, and anxiety. However, mathematical identity was still a significant predictor for usefulness of math.
Greene et al. (1999)	Unclear risk	Use a variation of the Eccles expectancy-value framework to explore different aspects of gender and motivation to explain motivation and performance in math	• Expectancy-value model: task-specific goals, task-specific values, task-specific beliefs, effort	For females in required math courses, goals, values, and beliefs accounted for significant proportions of variance in achievement scores. For females in elective math, only values and belie fs accounted for significant proportions of variance in achievement.
Gresky et al. (2005)	Low risk	Examine how women's math stereotype threat might be alleviated by reminding women of their multiple roles and identities	• Mathematical identity (Smith & White)	Women high in math identification got a larger percentage of attempted items correct and higher adjusted scores than did participants low in math identification.
Hendy et al. (2014)	High risk	Expand understanding of math beliefs in college students by developing three new psychometrically tested measures guided by the expectancy-value theory	• Math Value Scale • Math Confidence Scale • Math Barriers Scale	There was no significant interaction between gender and ethnicity on student math value scale, beliefs scale, and confidence scales.
Iben (1991)	Unclear risk	Examine whether attitudes towards math is a function of majority/minority status in a country, if attitudes towards math are a function of ethnicity or related to a combination of sex-ethnicity within a culture	• Math Attitudes Scale (Fennema-Sherman)	Mathematical identity was a significant predictor of girls' math ability, measures of abstract mathematical thought, and spatial relations.

(*continued*).

Table A1. Continued.

Study	RoB	Study aim	Mathematical identity measure(s)	Findings
Ing and Nylund-Gibson (2017)	Unclear risk	Describe heterogeneity in math and science attitudes and how these change through middle and high school	• Math and Science Attitudinal Survey	If there is a shift toward more positive attitudes prior to 10th grade, a large percentage of girls maintain a positive attitude in 12th grade.
Kombe et al. (2016)	Unclear risk	Investigate the relationship between all-girls, all-boys, and co-educational classes on student math self-concept and perception of classroom environment	• Math Attitudes Scale (Fennema-Sherman)	Single-sex education did not significantly influence student mathematics self-concept or student perception of the classroom environment.
Lauermann et al. (2017)	Unclear risk	Examine reciprocal associations between adolescents' math-related expectancy-value beliefs and career plans; examine the predictive role of these beliefs in math career attainment	• Expectancy and subjective task value beliefs in math • Career plans for math-related occupations	There were no significant gender differences in 9th and 12 grade math, utility and interest, but from 9–12 grade, girls' mathematics self-concept became more negative, as boys had higher self-concepts.
Lawrence and Charbonneau (2009)	Unclear risk	Assess how students' academic contingencies of self-worth and domain identification relate to academic performance depending on stereotyped status	• 3 items: perceptions of a strong mathematical identity; importance that others consider them as good in math; importance to score well in math	There was a marginally significant positive link between math identification and number correct on a mathematics test.
Leedy et al. (2003)	High risk	Explore attitudes toward and beliefs about the role of mathematics held by students	• Math Attitudes Scale (Fennema-Sherman)	The expectations of mathematics achievement for mothers of girls and mothers of boys were essentially equal, while the expectation of fathers of girls were higher than the fathers of boys.
Lent et al. (1996)	High risk	Explore sources of information that students employ in appraising their math self-efficacy	• Self-efficacy (adapted from Betz and Hackett) • Career aspirations (Goldman and Hewitt's science-non-science continuum)	Women were more likely than men to mention physiological states and reactions (e.g. anxiety, fatigue), and teaching quality, when asked to write down what they considered when making confidence rating in relation to mathematics courses.
Liu (2009)	Unclear risk	Investigate gender differences in affective factors, learning strategies, and preferred learning situations, and how these variables affect math achievement	• Motivational factors: intrinsic interest and motivation in learning math • Affective factors: self-efficacy belief, self-concept, and math anxiety	Girls reported higher levels of anxiety than boys, which was in turn negatively related to performance.

McClendon and Wigfield (1998)	Unclear risk	Examine group difference in academic achievement beliefs of Black adolescents	• Belief about ability in math; effort expenditure in school; expectancies for success; subjective values (Eccles)	In a population of 7th through 9th grade students, who were predominantly black (90%), female students valued math and science less than their male peers.
McGee (2015)	Unclear risk	Examine the development of mathematical identity for two high-achieving students and identity the approaches they adopted, created, or negotiated to maintain high achievement outcomes	• Fragile and Robust Mathematical Identity Framework	Fragile mathematical identity was a permanent aspect of Black female students, based largely on stereotypical assumptions about Black intelligence.
Meece et al. (1990)	Unclear risk	Identify predictors of math anxiety; asses the relative predict influence of math anxiety on students' math course enrolment plans	• Student Attitude Questionnaire (Eccles)	Perceived math ability and importance had a direct positive effect on current performance expectations and an indirect positive effect on performance.
Navarro et al. (2007)	High risk	Test portions of the social cognitive career theory in the math and science domains with Mexican American middle school students	• Math/Science Interests Scale (Fouad & Smith) • Math/Science Intentions and Goals Scale	Mexican girls had significantly lower self-efficacy than Mexican boys, but not outcome expectations, interest, or goal intentions. Additionally, Mexican girls had larger cultural orientation than boys.
Pedro et al. (1981)	Unclear risk	Identify variables related to election of math courses by gender; determine whether gender was predictive of election of math courses when controlling for achievement in math	• Math Attitudes Scale (Fennema-Sherman)	Measures of girls' math plans, success attributions, stereotyping math as a male domain, anxiety, usefulness and effectance were positively correlated with algebra achievement.
Piatek-Jimenez (2015)	Low risk	Identify what motivates women to choose math as an undergraduate major and investigate the factors influencing future career goals	• Qualitative descriptions of math as 'their subject'	The two predominant reasons women choose to study mathematics at the undergraduate level were a strong mathematical identity and enjoying the subject of mathematics.

(*continued*).

Table A1. Continued.

Study	RoB	Study aim	Mathematical identity measure(s)	Findings
Redmond et al. (2011)	Unclear risk	Explore the impact of an engineering-enrichment project on student ideas and attitudes towards math and science	• Modified Fenneme-Sherman Attitude Survey (Doepken et al., 1993)	Introducing real-world application of science and mathematics, problem-based learning, expanded opportunity for creativity, and after-school programming significantly increased girls' mathematics confidence compared to boys after a two-year intervention period.
Reid and Roberts (2006)	Unclear risk	To examine the effect of an intervention aimed to engage at-risk girls potentially talented in math on pursuit of math-related career options	• Fennema-Sherman Math Confidence subscale • Educational aspirations	Integrating mathematics and social science research in a single-sex, technology-rich environment, supported by college mentors significantly increased girls' mathematics confidence and achievement.
Schmader et al. (2004)	Unclear risk	Explore associations among legitimizing beliefs, stereotype endorsement, self-perceptions, and career intentions; test whether women who tend to endorse gender stereotypes about math ability exhibit a more negative impact of stereotype threat on math performance	• Continuation in math • Self-perceptions: confidence in future math performance, self-appraised math ability; performance self-esteem	Endorsing gender stereotypes had negative implications for women's math performance by making them more susceptible to stereotype threat.
Sherman and Fennema (1977)	Unclear risk	Determine the relationships between important affective, cognitive, and other variables and the decision to enrol in further math classes	• Math Attitudes Scale (Fennema-Sherman)	Girls intending to continue math had higher mathematical identity measures (confidence, attitudes, usefulness, and effectance motivation in math) than girls not intending to continue.
Sherman (1979)	Unclear risk	Determine whether cognitive and affective variables predict math performance and if these differ by sex	• Math Attitudes Scale (Fennema-Sherman)	Ninth grade cognitive and affective measures successfully predicted later math performance but weakened over time, with only cognitive 9th grade variables predicting 12th grade performance. Ninth grade confidence in learning mathematics was a significant predictor of 10th grade female math performance.
Sherman (1980a)	Unclear risk	Examine sex differences in cognitive skills, math performance, and spatial visualization; consider sociocultural influences	• Math Attitudes Scale II (Fennema-Sherman)	Girls' attitudes toward math became significantly less favourable from 8th to 11th grade, with the least stable attitudes toward mathematics found on the attitude toward success and perceived usefulness of mathematics.

Sherman (1980b)	Unclear risk	Predict grades in math courses	• Fennema-Sherman Aptitude Test • Math Attitudes Scale (Fennema-Sherman)	Eighth grade cognitive and affective measures successfully predicted girls' grades 1–3 years later, with confidence in learning math predicting student grades one year later and effectance motivation predicting student grades two years later.
Sherman (1981)	Unclear risk	Determine factors related to continuing to enrol in courses of theoretical mathematics	• Math Attitudes Scale (Fennema-Sherman)	Spatial visualization, vocabulary, perceived usefulness of mathematics, and confidence in learning mathematics of 9th grade girls were highest in discriminating those who had elected to participate in two, three, or four years of theoretical math in high school.
Sherman (1982)	Unclear risk	Assess if attitudes toward math behaviour is a causal factor affecting girls' enrolment in math and changes that occur in girls' attitudes toward math over high school years	• Math Attitudes Scale (Fennema-Sherman)	Ninth grade girls who enrolled in a 4th year of college preparatory mathematics were significantly more likely to consider mathematics as more of a male domain compared to females who had only enrolled in one, two, or three years of a college mathematics.
Sherman (1983a)	Unclear risk	Identify factors in the 8th grade which predicted whether 11th grade girls and boys had enrolled in one, two, three, or four years of a college preparatory math sequence	• Math Attitudes Scale (Fennema-Sherman)	Vocabulary, confidence in learning mathematics, attitude toward success in math, and effectance motivation in mathematics were the strongest predictive factors for girls' enrolment in college preparatory mathematics. Moreover, spatial visualization skill was a more important discriminator for females than for males.
Sherman (1983b)	Unclear risk	Gain insight into girls' feelings about how mathematics fits into their lives and future	• Math Attitudes Scale (Fennema-Sherman)	Girls continuing in math had more positive attitudes toward math, came from more favoured environments, were closer to parents and more influenced by them.
Simpkins et al. (2006)	Unclear risk	Test relations between math and science choices and beliefs from middle childhood through adolescence	• Expectancies-values in math and science (Eccles)	The intercorrelations among youths' ratings of interest, beliefs about importance, and self-concept in mathematics and science were high within 6th and 10th grade. Girls' mathematics grades and feelings about the importance of math in 10th grade was more strongly associated than boys' indicators.

(*continued*).

Table A1. Continued.

Study	RoB	Study aim	Mathematical identity measure(s)	Findings
Simpkins et al. (2012)	Unclear risk	Test a developmental model describing the processes by which parents' beliefs about their children were associated with adolescents' achievement-related behaviors	• Self-concept of ability • Task-value • Interests	Mothers' self-concept and value beliefs of mathematics significantly predicted their behaviors one year later, which significantly predicted youth beliefs and behaviors one year later.
Steinberg et al. (2012)	Low risk	Examine the moderating effects of mathematical identity and GPA on the relationship between stereotype threat and performance	• Gender identification measure (Schmader)	Women with higher math identification were more affected by stereotype threat than women with lower math identification.
Stevenson and Newman (1986)	Unclear risk	Assess the relative effectiveness in predicting high school attitudes of earlier ratings of general cognitive abilities made by students' mothers and teacher and measures of children's earlier achievement	• Math self-concept • Math expectancy • Math value • Math difficulty	Girls and boys did not differ in their self-ratings of either general cognitive abilities (learning, understanding discussions, and remembering) or specific academic abilities (mathematics and reading) in 2nd, 3rd, and 5th grades, but by 10th grade, there were differences in their attitudes about math and reading.
Stout et al. (2011)	Low risk	Test whether interactions a female or male peer support expert would influence women's self-conceptions in math and effort and performance on a math test	• Attitudes towards math • Identification with math • Stereotypes about math	When women pursuing STEM majors interacted with an advanced female peer who had expertise in math, they expressed more positive implicit attitudes toward math, showed more implicit identification with math, and increased their effort on a very difficult math test compared with others who interacted with an advanced male peer.
Tartre and Fennema (1995)	Unclear risk	Determine whether there are consistent patterns of gender differences for cognitive/affective variables and if these patterns predict math achievement for each gender	• Math Attitudes Scale (Fennema-Sherman)	Cognitive and affective factors (e.g. spatial and verbal skill, confidence, and perceived usefulness) accounted for 65% to 85% of the variance for female mathematics achievement at 6th, 8th, 10th, and 12th grade.
Walton et al. (2012)	Low risk	Examine the effect on achievement motivation of mere belonging, a minimal social connection to another person or group in a performance domain	• Math motivation (i.e. time persisting on an insoluble math puzzle) • Self-reported math motivation (Walton & Cohen, 2011): possible selves in math, math identification, self-efficacy with math, interest in math, potential to succeed in math, social fit in math	Minimal cues of social connectedness or sense of belonging increased women's persistence in an insoluble math puzzle. Additionally, women's altruism values negatively mediated the gender effect on STEM occupations.

Wang et al. (2015)	Unclear risk	Assess which intellectual and motivational factors in high school predict gender differences in career choices and whether students' motivational beliefs mediated the pathway of gender on STEM career via math achievement	• Math ability self-concept (Bleeker) • Academic self-concept (Eccles) • Math task value	Math task value partially explained the gender differences in STEM career attainment: women had lower math task values than men, and lower math task value was associated with lower math achievement and lower likelihood of pursuing STEM careers.
Watt et al. (2006)	Unclear risk	Examine female adolescents' choices for math participation during high school	• Math self-concept, expectancies, intrinsic value, and importance value (Eccles)	Girls' 10th grade expectancy beliefs, intrinsic value, and importance value directly influenced 11th grade math achievement.
Weavers et al. (2011)	High risk	Describe results from an evaluation of the Future Engineers' Summer Camp	• Math attitudes and confidence	A week-long day summer camp containing hands-on, minds-on activities spanning a range of engineering disciplines, significantly improved girls' knowledge of engineering, affirmed positive attitudes, and reinforced aspirations in mathematics and science.
Weinberg et al. (2011)	Unclear risk	Evaluate the effect of four experiential learning programmes on the interest and motivation of middle school students towards math and science	• Student Motivational Survey (Conley)	An experiential summer learning programme significantly increased mathematics interest and expectancy for success of both girls and boys, while also significantly lowering their attainment value (e.g. sense of importance of mathematics on who they are).
Whigham (1988)	High risk	Assess the academic achievement traits of academic females; describe how these traits compare with male students; and determine the differential attitude and anxiety toward math	• Measures adapted from • Fennema-Sherman Math Attitudes Scale • Dutton's Attitude Scale • Math Anxiety Rating Scale • Betz' anxiety scale	First-year female engineering college freshman reported significantly more positive attitudes toward success, attitudes of teachers, and stronger disagreement representing math as a male domain, compared to male peers. No differences were found on level of parental encouragement, confidence levels, and attitudes about the usefulness of mathematics.
You and Sharkey (2012)	Unclear risk	Investigates the impact of social and cognitive variables on math course-taking for girls and boys	• Math affection • Math self-concept	There were differential relations of parent and peer influences across gender groups. Peer values for grades and education were significantly related to taking an advanced mathematics course for girls only, whereas parental participation and mother's expectation were significant for boys only.