







Applying the Delphi method with early-career researchers to explore a gender-issues agenda in STEM education

Paul Hernandez-Martinez ¹ , Angeles Dominguez ^{2,3*} , Genaro Zavala ^{2,3} , Maria Kambouri ⁴ ,
Judith Zubieta ⁵ , Robin Clark ⁶ 

¹ Department of Mathematics, Swinburne University of Technology, Melbourne, AUSTRALIA

² Institute for the Future of Education, Tecnologico de Monterrey, Monterrey, MEXICO

³ School of Engineering, Universidad Andres Bello, Santiago, CHILE

⁴ Institute of Education, University of Reading, Reading, UK

⁵ Instituto de Investigaciones Sociales, Universidad Nacional Autonoma de Mexico, Mexico City, MEXICO

⁶ International Manufacturing Centre, WMG, University of Warwick, Coventry, UK

Received 20 March 2022 ▪ Accepted 27 August 2022

Abstract

The Delphi method (DM) was initially conceived as a forecasting technique whose results are based on the consensus of a panel of experts. It has been used in many fields, assisting researchers, policymakers, and others in setting directions and future agendas. This study presents an application of the DM, with a broader interpretation of the notion of “expert” as a qualitative tool to explore gender issues in science, technology, engineering, and mathematics (STEM) education. Its aim is to analyze the use of the DM as a technique to achieve agreement among a group of early-career researchers from the UK and Mexico, who are not considered “experts” in the traditional sense. The other aim was to explore the advantages and disadvantages of using the DM in this context. We used three stages to collect information and reach a consensus. The first two were online activities, and the last consisted of a five-day face-to-face workshop. In the first stage, participants sent research questions, and organizers categorized them into themes. Participants responded to a survey ranking all research questions in the second stage. In the last, organizers analyzed the highest scored questions and arranged them into research topics in which participants worked on research proposals. The DM worked successfully with this group of participants by combining their interest in the field and engagement with the activities. The research strands and proposals of using this method are usable. The method used in this paper can serve as a model to develop research graduate courses to develop students’ skills.

Keywords: Delphi method, educational innovation, gender research agenda, graduate research training, higher education, STEM education

INTRODUCTION

The Delphi method (DM) takes its name after the ancient Greek oracle to whom people would travel seeking divine advice about the future. The DM was originally conceived as a forecasting technique whose results are based on the consensus of a panel of experts (Grime & Wright, 2016). It has been used in many fields of study, assisting researchers, scientists, policymakers, and others in setting directions and future agendas (e.g.,

Alcock et al., 2016; Gonzalez-Garcia et al., 2021; Guglyuvatyy & Stoianoff, 2015).

This paper presents a case of the use of the DM as a qualitative research tool to explore gender issues in the field of science, technology, engineering, and mathematics (STEM) education, a young research field. However, the use of the DM in this paper involves a broader interpretation of the notion of “expert.” In DM studies, an expert is broadly defined as someone who has considerable knowledge and/or expertise in a

Contribution to the literature

- The success of using the Delphi method to reach a consensus on a group of early-career researchers who are not considered “experts” in the traditional sense.
- Use of the Delphi method as a qualitative tool to generate research ideas for grant proposals.
- The contribution of this paper to capacity building in the form of faculty development and networking, and as a model to develop graduate research courses to develop students’ research skills and competencies and help them in their academic careers.

relevant field. In this paper, “expert” refers to a group of early-career researchers.

Interest in using the DM with such “experts” arose from the increasing number of calls to improve the quality of undergraduate STEM education (Henderson et al., 2017; Talanquer, 2014). As part of this concern, universities worldwide now offer graduate programs in STEM education. Some of these programs focus on the interrelation between the STEM disciplines, while others emphasize research on one particular subject (e.g., physics education or computer science education). It is expected that students in these programs should acquire different research abilities and competencies. Most of these programs provide training on methodological issues or data analysis (Kilburn & Earley, 2015), for example, but it is still uncommon to have courses or workshops that explicitly prepare students for the process of applying for research funding. This aspect becomes vital for early career researchers, who are expected to develop a research line and to conduct independent research that will, in turn, become a sustainable research trajectory. Gaining research grants is crucial in developing a sustainable research program and is considered an essential scholarly output by recruitment, tenure, and promotion committees (Kamerlin et al., 2019). In addition, bringing external funding has a beneficial impact on the researcher and the institution (Whittaker & Montgomery, 2014), as it contributes to knowledge, increases research productivity, provides opportunities for students, strengthens a network, and increases the institution’s prestige, among other benefits.

This paper analyzes the outcomes of using the DM as a qualitative tool to explore, agree, discuss, and refine research proposals by early career researchers and mentors. To contribute to capacity building and generate research ideas that could become fundable research proposals in an important area of STEM education (i.e., gender studies), the authors of this paper designed and implemented a workshop, taking advantage of a binational program of collaboration. We now describe context and establish the procedures in using the DM.

Context

As part of an international program for collaboration between the UK and Mexico, a five-day workshop took place in Playa del Carmen, Mexico. The theme of the workshop was “Gender issues in STEM education.” The aim was to bring together early-career researchers¹ (ECRs) from both countries to delve into the workshop topic through different activities (e.g., discussing readings, presenting, researching funding bodies). With the guidance and leadership of six experienced researchers (the mentors who are the authors of this paper), and the advice of a funding agency consultant, the ECRs would develop a series of research proposals to submit to relevant international funding bodies.

However, the participants’ various backgrounds, knowledge, interests, and career stages posed a potential barrier to fulfilling the workshop’s aims. There was a risk of information bias, in the sense that the mentors might impose their agendas, given their research programs and expertise, and that the voices and interests of the ECRs would be silenced. There was also a risk that the relative inexperience and limited expertise of the ECRs would mean that the outcome proposals were not of sufficient quality or significance to be considered for submission to funding bodies. The mentors desired a process that would ensure that all the participants’ interests would be considered and that, as a result of the workshop, there would be an agreement on the priority areas of gender research in STEM education and which of these areas could be developed as research projects. The mentors decided that the DM was a suitable technique to reach a consensus that would then allow the participants to develop fundable research proposals.

Delphi Method

DM was developed in the 1950s as part of the RAND (Research and Development) Corporation in California, USA, to forecast research into science and technology that could be used by the military (Gordon & Helmer, 1964). The method is now considered a “futures research” tool that explores alternative futures in a

¹ We defined an early-career researcher as a doctoral student (2nd/3rd year of studies in the UK or 3rd/4th year of studies in Mexico) or a post-doctoral researcher that was within five years of having earned their PhD by the time of the workshop.

complex and rapidly changing world (Aubusson et al., 2016).

The Delphi method's primary purpose is to reach a consensus among a group of experts about essential questions for future development and decision-making in a particular field of knowledge. Fowles (1978) argued that experts' testimony is permissible in fields that have not developed sufficiently to have scientific laws. Thus, the Delphi method attempts to address not the "what is" but the "what could be/should be" (Miller, 2006).

One of the most important aspects of the DM is the selection of a panel of experts. However, as Hallowell and Gambatese (2010, p. 102) observed in their review of studies that applied the DM in the construction engineering and management research field, "the characteristics required to define an individual as an 'expert' are equivocal." They found that some studies clearly defined the criteria for expertise while others did not indicate specific requirements. Despite this ambiguity in the definition of expertise, and since the DM does not depend on a sample of experts that is representative of a population but relies on the informed opinion of a homogenous or heterogeneous group. Most researchers using the technique would agree that the experts should have a deep understanding of the field or problem at hand (Okoli & Pawlowski, 2004).

In a review of the literature for this paper, only two cases were found where the researchers used participants that would not normally be considered "experts" as described above. The first case was the study of Wynekoop and Walz (2000), who used nine MBA students to investigate the traits and behaviors of top-performing software developers. Their paper reported the use of MBA students as a limitation of their study. The second case was the study of Garavalia and Gredler (2004), who used a group of undergraduate students and a group of doctoral students to document students' perceptions of the effectiveness of an academic program. However, the main purpose of their study was to teach the DM to their students and not, as is the case of the present paper, to use the DM as a methodological tool and to evaluate its usefulness in a particular context.

After a panel of experts is selected, the goal of the DM is accomplished through an iterative survey process with anonymous feedback in which the expert participants reassess their initial judgments on each iteration. The anonymity of the participants' feedback is an essential aspect of the method because it can reduce undesirable effects such as negative influence due to the participants' status or personality. The DM also has two other significant characteristics. First, controlled feedback: a coordinator controls the flow of information to filter the irrelevant information. Second, a group response reflects the opinions of all participants (Landeta, 2006), even if "at the end of the exercise there

may still be a significant spread in individual opinions" (Dalkey, 1969, p. 414).

Since the 1960s, the DM became better known publicly, and it has been used, developed, and adapted to various contexts and topics. For example, it has been used in STEM education to identify scientific competencies for citizenship (Gonzalez-Garcia et al., 2021), to gain consensus on how to teach linear algebra (Rensaa et al., 2020), or to develop an agenda for mathematical research cognition (Alcock et al., 2016). Consequently, its scope and definition have changed to accommodate various necessities, particularly relevant to this paper, its use in qualitative research (Brady, 2015). In line with Linstone and Turoff (2002, p. 3), we take the DM as "a method for structuring a group communication process so that the process is effective in allowing a group of individuals, as a whole, to deal with a complex problem."

Aims and Research Questions

The primary aim of this paper was to analyze the use of the DM as a technique to achieve agreement on important topics in gender issues in STEM education. A significant characteristic of our study was the use of early career researchers as "experts". We also wanted to evaluate if these topics could be used to produce research proposals that had the potential to be submitted for funding. Our use of the DM as a tool implied qualitative analysis of the data at crucial stages of the process, where synthesis and collaborative interpretation of the data were needed.

A secondary aim of the paper was to document this process considering our interest in developing ECRs' skills and competencies that are important and valuable to them. Hence, we were interested in knowing the advantages and limitations of using the DM in this way, so that others might benefit from this experience and might take this as a model for their own training programs. Therefore, our research questions are:

1. **RQ1.** How can the Delphi Method be used by a group of early-career researchers and experienced mentors to investigate priority areas in STEM gender education?
2. **RQ2.** What are the advantages and disadvantages of using the Delphi Method in this context?

METHOD

We now describe the participants in the workshop and how we used the DM to explore possible research agendas on gender issues in STEM education.

Participants

The two organizers of the workshop (the first two authors of this paper) invited four experienced researchers (two from the UK and two from Mexico) to

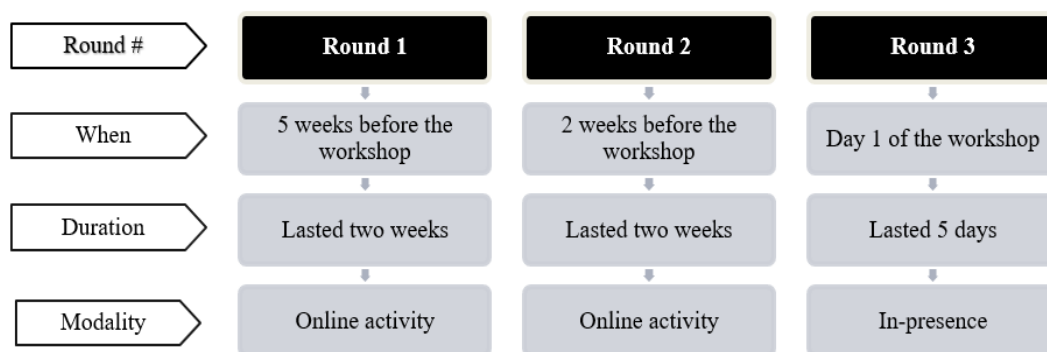


Figure 1. Implemented stages of the Delphi method to identify priority areas in gender issues in STEM education

participate as mentors and to lead teams of early-career researchers to develop various proposals to be submitted for funding after the end of the workshop. All six mentors are academics with expertise in researching STEM education and are knowledgeable about gender and other socio-cultural dimensions in education.

The organizers published a call in the UK and Mexico for ECRs to apply for the workshop. There was a maximum of 24 available places. The intention was to have a balance of gender and nationality in the group. The participants submitted their curriculum vitae and a brief application form explaining why they wanted to attend the workshop and what they would bring to it. It was not a prerequisite for acceptance a background in educational research, in fact, some of them came from STEM backgrounds. In addition, the Mexican applicants were required to have a reasonably level of oral and written English.

There were 10 UK applicants and 19 Mexican applicants. The organizers reviewed all applications and rejected those who did not have an adequate background (e.g., no experience in education research or gender issues) and accepted those with a strong background and interest in the topic. The other mentors were asked to comment on seven applications that were not accepted or rejected in the first round to help decide on these applications. In the end, the organizers accepted seven UK participants (two males and five females; four post-doctoral, three PhD students) and fourteen Mexican participants (five males and nine females; three post-doctoral and eleven PhD students). Hence, of all early-career research participants, 33% were male, and 66% female; 33% had a post-doctoral level education, while 66% were at the doctoral level.

Exploring Gender Issues in STEM Education

Following Hsu and Sandford (2007, p. 2), we used “the Delphi process as a data collection technique” and considered, as some other authors suggest (Custer et al., 1999; Ludwig, 1997), that three iterations are sufficient to collect the information required and to reach a consensus. The following paragraphs present details on those three rounds (Figure 1); the first two were carried

out online, while the third was completed during the workshop.

ROUND ONE took place five weeks before the workshop started and consisted of an online questionnaire asking participants (including mentors but not the organizers) to formulate at least four questions that they thought were essential to pursue further advances in the research field. Since the ECRs were not experts, as defined in the classic DM, we added a preparation step consisting of mandatory readings for all the participants. Specifically, they read three literature review papers (Galeshi, 2013; Kulturel-Konak et al., 2011; Wang & Degol, 2017), a short report on the gender gap in STEM disciplines (UNESCO, 2016), and three to four additional research papers of their choice reflecting their current research or interests in the discipline. We wanted to ensure a good number of questions to build a common core of shared topics while allowing, at the same time, enough questions to cover a wide range of research interests. Participants were given two weeks to send in their questions. Once all the answers were received, the organizers analyzed, amalgamated, removed duplicates, and tried to allocate them into independent themes.

ROUND TWO started two weeks before the workshop when the participants were asked to rate the questions raised in round one. The themes were not revealed to the participants. The questions were presented sequentially by themes in the order described below (see Results section). In hindsight, this might not have been the best decision since it is well known that respondents are often inclined to rate higher the first options they read. However, we did not find any trend showing participants ranking the first questions higher or the last ones lower. In other words, there was no evidence that a random presentation of the questions would have yielded substantially different results.

Participants were asked to rank each question on a four-point Likert-type scale (1-not a priority, 2-rarely a priority, 3-somewhat a priority, 4-definitely a priority), assessing their perceptions of the importance of a research agenda for gender studies in STEM education.

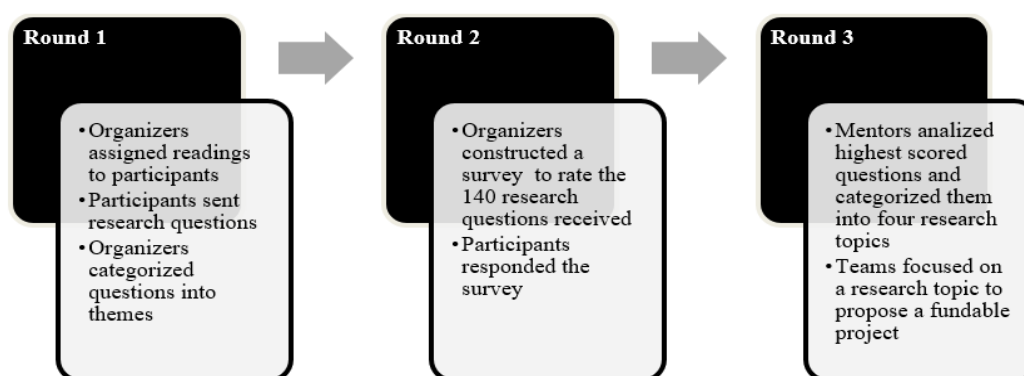


Figure 2. Actions taken in each round of the Delphi process

ROUND THREE took place during the workshop. All six mentors gathered to analyze questions with the highest scores, interpreting them and determining sensible research strands that could be used as topics to develop research proposals. Four topics arose from this round, and these were presented to the group. The participants were asked to choose the strand they would like to pursue, and all found an interest team to join.

The mentors agreed that the four research topics that emerged from the DM exercise constituted interesting and current research areas, and that projects based on these topics would advance the field of gender studies in STEM education. However, to validate that these topics were current and had the potential of being funded, we searched literature review articles published from 2015 onwards in the top three journals dedicated to this type of publications. Literature review articles, in general, reflect state-of-the-art research and point towards future agendas. We were interested to see if our four topics were represented in these journals, hence providing support to our currency claims. We searched the following top ranked journals: Educational Researcher, Review of Educational Research, and Review of Research in Education.

RESULTS

We now elaborate on the actions taken and the results obtained from each of the rounds of the Delphi process (Figure 2). Each round explains how we followed the methodology that allow us to identify the four research topics described below.

We received 140 questions in round one. The exercise showed various points of view and interests, posing a challenge to organize all 140 questions into sensible themes. The questions are not presented since they were classified and ranked in the following sections. However, the organizers felt all the questions were relevant, and none should be discarded or amalgamated with another. Some were similar but had slight interesting differences to keep for possible later discussion. The two organizers interpreted the questions and negotiated between them, resulting in the following **themes**:

1. School/university practices and cultures, including:
 - a. Cultural and pedagogical issues in the early years of men and women.
 - b. Particular pedagogical practices in schools and universities related to gender inequities.
 - c. Institutional programs and courses where gender participation is unequal and their characteristics.
 - d. Diverse pedagogical resources (e.g., textbooks) that affect men and women in different ways.
 - e. The influence of teachers in men's and women's educational experience.
 - f. Issues with special learners (e.g., indigenous students, LGBTIQ+ students, etc.).
2. STEM professional cultures and their impact on gender inequalities (e.g., salary gap, women representation, etc.).
3. Government/institutional/societal initiatives that promote gender equity.
4. The influence of family and other socio-cultural and political issues that affect opportunities, aspirations, choices, and access of men and women to STEM education and careers.
5. Particular characteristics of individuals in STEM (e.g., scientists, mentors, etc.) and the importance of role models in inspiring men's and women's participation in STEM careers.
6. The role of social media and particular technologies in encouraging STEM participation in men and women.
7. Measures to predict retention and achievement in STEM education (e.g., self-efficacy, the gender inequality index, etc.).
8. Individual perceptions, attitudes, and interests that determine the participation of men and women in STEM and how these arise and develop.
9. Theoretical perspectives for the study of gender issues in STEM.

Table 1. Questions from research topic 1 indicating percentage received on “definitely a priority” and “somewhat a priority”, corresponding theme and role of participant

Research questions topic 1	Percentage	Theme	Role
Most of the factors determining the retention or desertion in STEM fields are detected in high school or university. In consequence, what kind of expertise in STEM education should be offered in elementary school to prevent factors like stereotypes, female discrimination, and academic sexism from occurring in the next academic levels?	84.6%	T1	S1
How does research-based teaching in early school help overcome gender stereotypes of male and female students?	80.8%	T1	S2
How do early years experiences affect children’s aspirations to follow a STEM career?	76.9%	T1	M1
Do parents’ professions and education (especially mothers) influence young girls’ and boys’ educational choices?	76.9%	T4	S3
How can parents and early-year practitioners work together to ensure that both girls and boys have equal opportunities to follow a STEM career?	73.1%	T4	M1

Note. T# indicates the theme number from round 1 & Role denotes the participant who formulated the question, where S# for a PhD student, P# for a post-doctoral researcher, & M# for a mentor

Table 2. Questions from research topic 2 indicating percentage received on “definitely a priority” and “somewhat a priority”, corresponding theme and role of participant

Research questions topic 1	Percentage	Theme	Role
What more could be done in teacher education to dispel stereotyping in math and science education?	80.8%	T1	P1
What practices persist in schools despite awareness of their negative effects on female recruitment to STEM degrees? What do gender-neutral teaching strategies look like?	76.9%	T1	P1
Do males and females differ in what they look for in a mentor?	76.9%	T5	S4
How do teaching styles impact students’ learning styles from a gender perspective, and what kinds of activities can help break the gender gap?	73.1%	T1	S2
What are the educational features that discourage female pursuit of a STEM career?	73.1%	T10	M2
In which ways are gender stereotypes perpetuated in the Mexican STEM curricula and learning materials?	73.1%	T1	S5
How does an educational intervention (an activity that allows students to apply their learning in real-life situations through creative and hands-on experiments) impact the quality of student learning and interest in STEM-related subjects?	73.1%	T1	S5

Note. T# indicates the theme number from round 1 & Role denotes the participant who formulated the question, where S# for a PhD student, P# for a post-doctoral researcher, & M# for a mentor

10. Other general issues not fitting into any of the previous themes and could not be classified as a unique theme, for example, Why are there so few female professionals in areas that traditionally appeal to women, such as ecology and global health? Where do STEM professionals rank compared to other professions?

Theme 1 had the highest percentage of questions, with 25.7% of the 140 questions. Next came theme 4 with 18.6%, general questions (theme 10) with 13.6%, and themes 5 and 8 with 11.4% each. The other five themes added up to 19.3%, less than 6% each.

ROUND TWO. All participants, including all mentors, ranked the questions in this round. We were pleasantly surprised to find that everybody took the time to rate all 140 questions because sometimes the literature reports that many first-time users of the DM find it disappointing due to the effort and difficulty involved (Landeta, 2006). Also, expert participants feel that they are asked to do much work without objective justification (Landeta, 2006). 25 questions that received 70% or above on “definitely a priority” and “somewhat a priority” were selected in this round.

ROUND THREE. In this round, all the mentors gathered to interpret and discuss the questions obtained in **ROUND TWO** and grouped them into four sensible strands to develop research proposals:

RESEARCH TOPIC 1. *Factors influencing attitudes toward STEM fields in early years and elementary schools.* The mentors observed sufficient interest in early-year issues (when many attitudes towards STEM begin to influence boys and girls) to integrate a research strand to investigate these issues. **Table 1** presents the questions belonging to this strand.

RESEARCH TOPIC 2. *Influence of teachers and teaching on the STEM education gender gap and how to address it in teacher education.* The mentors thought these questions fitted into a research strand where teachers, teaching, and teacher education were the focus. **Table 2** indicates the questions that formed this strand.

RESEARCH TOPIC 3. *The role of the job market and industry in STEM career choices of males and females.* The mentors thought this research strand was about society’s role in promoting gender equality, including industry. **Table 3** presents the questions that formed this strand.

Table 3. Questions from research topic 3 indicating percentage received on “definitely a priority” and “somewhat a priority”, corresponding theme and role of participant

Research questions topic 1	Percentage	Theme	Role
What are the implications of society, culture, education, and the labor market in the male and female career choices in STEM?	84.6%	T4	P2
Which strategies and actions are taken by researchers, science centers, and science communicators to influence the social imagery of STEM to make visible women in these fields?	77.0%	T3	S6
What is the role of industry in addressing gender issues in STEM?	76.9%	T2	M2
Who are the gatekeepers of STEM professions, and what can be done to encourage more women to enter these fields (e.g., female-friendly gatekeepers)? What are the policies currently in place that, if changed, would encourage women to consider STEM professions?	76.9%	T3	P1
To what extent and how can institutions such as schools, universities, and governmental bodies influence closing the gender gap in STEM at the institutional level?	73.1%	T3	P1

Note. T# indicates the theme number from round 1 & Role denotes the participant who formulated the question, where S# for a PhD student, P# for a post-doctoral researcher, & M# for a mentor

Table 4. Questions from research topic 4 indicating percentage received on “definitely a priority” and “somewhat a priority”, corresponding theme and role of participant

Research questions topic 1	Percentage	Theme	Role
What are the characteristics of a role model for greater gender equality in STEM education?	80.8%	T10	M2
How do female role models in STEM fields influence interest and persistence in STEM fields in our context?	76.9%	T5	S7
How do stereotypes act as barriers to entry into the STEM fields?	76.9%	T4	P3
How do the STEM aspirations of girls in different ethnic groups (especially indigenous girls) change as they progress from primary school through middle school and high school?	76.9%	T1	S8
What factors influenced current Mexican female scientists to study a STEM major in science at the university level?	76.9%	T10	P4
How do cultural and ethnical backgrounds influence stereotypes related to gender in the STEM fields?	73.1%	T4	P3
How do STEM graduates describe their identities? To what extent is this influenced by gender?	73.1%	T8	S9
Does the possible influence of parents, siblings, and teachers on STEM attitudes vary according to different social groups?	73.1%	T4	M3

Note. T# indicates the theme number from round 1 & Role denotes the participant who formulated the question, where S# for a PhD student, P# for a post-doctoral researcher, & M# for a mentor

RESEARCH TOPIC 4. The effect of “identity” and “role models” on the gender gap in STEM career participation. The mentors thought that this strand was about the influences of particular individuals (including family members) on men’s and women’s identities and how they influence their aspirations, choices, and participation in STEM careers. A previous question referring to the mentor’s role was repeated because it was felt that it also fitted within this strand. **Table 4** shows the questions of this strand.

The 25 questions on the four research topics were formulated by nine PhD students, four post-graduate researchers, and three mentors. Specifically, Ph.D. students developed 40% of the top 25 questions, postgraduate researchers formulated 36%, and 24% by mentors. The questions came from original themes 1, 2, 3, 4, 5, 8, and 10. The themes that were not considered in the final 25 questions were “role of social media/technologies” (6), “measures to predict retention/achievement” (7), and “theoretical perspectives” (9). Therefore, we achieved a good spread of participants and themes in the final four research topics; no participant (even mentors) or theme

dominated the DM exercise’s results. We found external validation of all four topics when we searched for literature review articles in these areas. Within the top three journals that publish literature review articles, we found at least one article closely related to each of our four topics (**Table 5**).

After round three, the group was divided into teams to further explore the topics and to propose a viable research study that could be fundable. At this stage, a funding agency consultant helped the teams identify the most appropriate funding bodies for their proposals. Participants and mentors were given the opportunity to decide which team they would like to be in, and everyone found a team that suited their interests. Each team had at least one mentor on it, and there were very few participants that moved teams afterwards.

We now briefly describe how each team conducted their work, and we synthesize key points that we consider useful for others looking at pursuing similar work.

TEAM 1 (early years). Considering the questions that formed research topic 1, this team started by examining

Table 5. Literature review articles that have been published since 2015 in the four research topic areas in three top journals

Topic	Published articles in the topic area
Topic 1-Early years	Currant and Kellogg (2016), and Quinn and Cooc (2015)
Topic 2-Influence of teachers & teaching	Copur-Gencturk et al. (2020), Gregory et al. (2017), and Rowan et al. (2021)
Topic 3-Career choices	Saw et al. (2018)
Topic 4-Identity & role models	Ahmad (2017), Cabrera et al. (2021), Ireland et al. (2018), Joseph et al. (2017), and Kim et al. (2018)

previous research conducted in different countries that focused on young people's attitudes and aspirations towards STEM. The team discussed what its members already knew about gender stereotypes and how they develop at a young age. During this process, the team gathered more sound evidence on the emergence of the gender gap, how it is defined by different scholars, at what age it is identified, and how it defines or prevents vocational aspirations. The team noted that the literature tends to focus on older students; little evidence was found regarding early childhood, especially in non-developed countries where studies of this kind are hard to find (Allen, 2011).

This literature review allowed the team to identify those interventions focusing on tackling gender stereotypes at the secondary school level that were considered unsuccessful because it was believed already too late to change (Archer & DeWitt, 2014). This highlighted the need to focus efforts during earlier years, when aspirations and attitudes are being formed, and when children's preferences and aversions are still malleable. While discussing how the gender gap develops in the early years of life, the team briefly looked at actors and activities that seemed essential to nurture interest in STEM disciplines and prevent gender stereotypes.

Once they agreed on the general issues, the team negotiated various options to determine the project's aim, scope, focus, and geographical coverage. Teachers and parents were included due to their influence on children's attitudes and preferences.

Considering the team's strength, networks (access to new participants), and backgrounds (expertise), the team finally conceived a project to identify when gender stereotypes emerge and then develop and implement a parent-teacher partnership workshop. Its goal was to learn and practice the critical characteristics of an inclusive pedagogy that celebrates diversity. The study sought to explore the situation in three countries (UK, Mexico, and Argentina) comparatively to identify the age at which gender stereotypes emerge (through structured interviews with children aged 5, 6, and 7). The study sought to determine the characteristics of a pedagogy that would discourage such stereotypes from developing (through a partnership workshop to promote an inclusive approach). The data collected would be analyzed and compared at national and international levels; findings would be used to make

recommendations for policymaking and practices in schools in the three participating countries.

TEAM 2 (influence of teachers and teaching). The team thought that the questions from research topic 2 were thoughtful enough to become a project with objectives and expected results that transcend research on gender issues. However, they thought that they needed to focus their work on identifying themes within the research questions and establish attainable goals. Because the team members' backgrounds were from different STEM subjects (physics, mathematics, biology, and computer science), they decided to follow a collaborative strategy that leveraged each individual's potential. They spent some time investigating the task at hand in their respective fields, and then worked together to discuss and find common ground to establish their objectives.

The team agreed that much of the work by Baker (2013), Kim (2016), Lorenzo et al. (2006), and Valla and Williams (2012) were essential and agreed to work in two sub-teams, one of them working on the literature about what has been done in STEM pedagogy addressing gender issues (mostly active, inquiry-based, or hands-on learning) while the other was concerned about how a gender-inclusive STEM pedagogy should appear.

The results of the two-team activity allowed them to identify some problems on which they based their research questions and design:

1. Evidence in the literature is insufficient to assert that active/ inquiry-based learning (EBL)/hands-on learning strategies decrease the gender gap in STEM outcomes.
2. All studies used a gender perspective to analyze the results, but not for activities design, curricula, or assessments.
3. Studies from the literature work under the assumption that these strategies are value-free and "objective."
4. Active/EBL approaches work by simulating real-world science in the classroom. However, not reflecting on gender issues in science can reproduce its biases and problems in the schoolroom.

Based on the literature's deficiencies, the team finally agreed on a project to identify the characteristics of an inquiry-based learning (IBL) strategy that could decrease the gender gap in STEM performance. The team decided that the project should include the effect of EBL

with a gender perspective on the gap in STEM learning, considering women's attitudes, motivations, aspirations, and self-efficacy for learning science and on retaining women in STEM.

TEAM 3 (STEM career choices). After revisiting the questions from research topic 3, the team kicked off with the idea-generation phase. This team initially discussed where the research work should focus. The feeling was that the path ahead should consider the context in which STEM career choices were made. This discussion led to the formulation of an overarching research question: What contextual characteristics promote STEM equality? The team adopted a working definition of equality to provide a lens through which other literature could be explored: "equality allows all people regardless of gender, ethnicity, ... to achieve their aspirations to the fullest." This definition ultimately resulted in some initial ideas having the potential to be developed into research proposals.

Reaching consensus on a definition for equality allowed the team to work synergistically throughout the rest of the project. Small-group brainstorming helped to determine the characteristics for evaluating any idea for a research proposal. In this process, the team built upon the ideas rather than discard them. The main ones selected were:

1. Value/impact—does it have the potential to bring about sustainable change?
2. Practicality—can it be achieved within the funding framework?
3. Does it lend itself to an international investigation?
4. What has been done so far in the space?
5. Will it excite the potential funders/fit a particular call?

With a straightforward research question, a process built on the initial DM, and stated parameters to guide an idea's selection, the team explored the literature and identified potential funding sources. Baker's (2016) work was fundamental in driving this part of the work, and a conceptual model for gender equality in STEM was developed. This model's goal was to go beyond current thinking to promote sustainable change in career opportunities for both males and females.

Through a series of iterative discussions, the team identified a revised question for a research proposal: How can industry contribute to developing a pedagogy that promotes gender equality in STEM employment? Moreover, the team looked for an industry funding source interested in participating in conducting a pilot project. The team concluded that efforts are underway to reduce the gender gap in STEM areas in both the university and industrial contexts. However, these efforts seem not to be aligned or connected. Therefore, the team decided to construct a proposal that would

tackle this problem and create a link between the university and industry to reduce the gender gap in STEM areas.

TEAM 4 (role models). This team started by reviewing the questions in research topic 4 and finding what the team members knew about the mechanisms by which "others" influence someone's STEM identity. The team then made a brief review of the literature on this topic. The reading of the literature brought out a focus on the pivotal role that family has on career choices. Several factors were considered a determinant of those decisions, including economic or social stratification, the image of a profession held by family members, and the satisfaction of family needs (Razo, 2008). The team decided to use the science capital concept (Archer et al., 2015) to assess these influencing factors. They aimed to develop a project that:

- (1) investigates the science capital of parents, particularly mothers, who can influence their children's STEM aspirations, and
- (2) determines what students, particularly females, look for in a role model (i.e., someone that can potentially provide them with important science capital).

To pursue these goals, the team agreed that developing an engaging information campaign about women in STEM, directed toward parents in lower socio-economic backgrounds, would be an excellent way to investigate these issues. Focus groups with parents exposed to this campaign would be used to collect data. The information gathered would include their contexts, circumstances, and thoughts about encouraging their daughters to pursue STEM careers and what would make it possible for them to consider supporting their children to follow a STEM profession. In addition, interviews with these families' children about their aspirations and how they see science and scientists would allow an understanding of how their identities develop and what can influence them to study a STEM subject.

The four teams developed their ideas and proposals independently of each other, and it can be seen that their processes differ. Some of the teams started with a literature review while others formulated an initial research question based on their corresponding Research Topic and then conducted a review based on key research studies. Some of the teams placed a greater effort in developing a proposal that was practical and that might have an impact while others considered important to have a proposal that fitted certain funders. However, we identified some similarities in their approaches, namely:

1. Revisiting questions from its corresponding research topic,

2. Conducting a review of the literature to better understand the area, and what has been done already,
3. Building on the team members' backgrounds, strengths, and knowledge, and,
4. Collaborative decision making to formulate consensus, aims and a final proposal.

These approaches provide, in our view, models that can be used to train early-career researchers to develop research proposals—based on identified priorities—that have the potential to be funded.

We now discuss our results in light of our use of the DM.

DISCUSSION AND CONCLUSIONS

The use of the DM with a group of early-career researchers, who are not considered as “experts” in the traditional DM literature, posed a challenge that in our view is worth being studied and reported. Our research questions were:

1. **RQ1.** How can the Delphi method be used by a group of early-career researchers and experienced mentors to investigate priority areas in STEM gender education?
2. **RQ2.** What are the advantages and disadvantages of using the Delphi Method in this context?

To answer our first research question, we observed that the results showed, in general, that our use of the DM was successful in providing a consensus on essential topics about gender issues in STEM education. These topics arose from the participants' interests and what they considered priorities in the field. They offered practical avenues to pursue various research agendas in the form of funding proposals. The fact that our “experts” were early-career researchers and that the group was non-homogeneous with multiple expertise levels and viewpoints was not a significant obstacle in producing exciting and researchable strands with open questions in the field. In this study early-career researchers were representative of a population with a necessarily informed opinion (Hallowell & Gambatese, 2010) and the requirement of having people with a deep understanding of the problem at hand (Okoli & Pawlowski, 2004). Our decision to ask participants to read a few comprehensive review articles and others of their choice was an excellent strategy to ensure a basic understanding of the state-of-the-art in the field, as shown by the quality of questions received. At the same time, this strategy allowed for enough diversity of views to appeal to everyone's interests. Therefore, we consider that the level of “expertise” achieved, combined with the interest and engagement of the participants, was sufficient to obtain good quality questions, arrive at a consensus of priorities in the field, and arrange those priority questions into sensible strands that formed a

basis for achievable and interesting research proposals. We then disagree with Wynekoop and Walz (2000), who used MBA students to investigate the traits and behaviors of top-performing software developers. They concluded that having MBA students was a limitation. In our case, using early-career researchers, doctoral students, and young researchers was not a limitation.

The results showed that no one group of participants, including the mentors, influenced the exercise's outcomes heavily. Most of the themes identified at the beginning of the process were represented at the end. This was an outcome of the application of the DM, rather than a purposeful decision by the organizers or mentors.

To answer our second research question, we reflected on our experience of using the DM by discussing some of its most relevant methodological weaknesses pointed out by Landeta (2006), to show the robustness and usefulness of our results:

1. *Its basic source of information (who is the expert, what biases each expert has, etc.).* The fact that the participants were not “experts” in the traditional sense did not detract from good and relevant questions being proposed, as we discussed above. The questions comprised a bank of ideas from which to discuss and form coherent, interesting research proposals.
2. *The use of consensus as a way to approach the truth.* We were not looking for truth in the sense that we did not want to find *the* questions that would lead the field of gender issues in STEM education for the following years. Instead, we were interested in questions relevant to the research field at the time and of interest to our participants, which could be formulated into research proposals that funding bodies would be interested in sponsoring. In this sense, our four research topics can be seen as areas where more research is needed to advance knowledge. Thus, what was important to us was consensus on the participants' interests and relevant research topics that would advance knowledge of gender issues in STEM education.
3. *The limitation of the interaction involved in written and controlled feedback.* In the classic DM, the coordinator usually controls information flow and decides what is relevant. In our case, the two organizers came together to interpret and agree if the participants posed any repeated or irrelevant questions. None of the 140 questions were discarded or amalgamated with others and were classified into coherent themes, showing that the organizers valued all the contributions. Thus, they did not exercise inadequate control over the flow of information despite participants not being actively involved in this stage of the process. Similarly, the six mentors interpreted and agreed on four research topics that made sense, given the

participants' ranking of questions. The mentors led the teams, but their approach was collaborative, building on the interests, knowledge, and expertise of the ECRs, as described in the results. Therefore, there was no single stage in the process where one person had total control; instead, we saw negotiation and agreement on participants' information as the main driver of the process.

4. *The restriction to the possibility of social compensation for individual contribution to the group (The reinforcement and motivation usually provided by the other expert group members' support and social approval are removed).* We did not encounter a lack of motivation from our participants. On the contrary, all the participants were fully engaged and intrinsically motivated by their interest in this topic. The ECRs were encouraged by the possibility of gaining experience in producing a research proposal. Their questions' quantity and quality and their participation in all activities during the workshop attest to this motivation. We did not encounter evidence of participants wanting to have a "free" holiday or having a secondary plan other than the workshop's stated and publicly accessible aims.
5. *The impunity of irresponsible actions by the experts conferred by anonymity.* Again, given the quantity and quality of the participants' information and their full participation in the workshop, we did not find evidence of any participant acting irresponsibly (by trying to impose their views or deceive others) or strategically restricting information. The mentors sought a fully collaborative approach with the ECRs in developing the research proposals. We believe that applying the method as described here made it quite improbable that a participant's negative actions would have biased the results.
6. *This methodology's inherent ease of data manipulation toward particular interests by the person running the study.* As said before, at any stage of the process, no one person could manipulate the information based on their interests. Manipulation towards particular interests would have been quite tricky. Even at the last stage of the process, the teams made decisions based on the discussion of literature reviews, funding bodies' requirements, and practical issues concerning the conduction of the research rather than personal opinions or particular interests.
7. *The difficulty of checking the method's accuracy and reliability.* This group of ECRs and a few experienced researchers had particular characteristics and interests. As already stated, the resulting research topics reflected these characteristics and interests because the aim was

not to reach a "truth". However, the consensus about what was important, given the parameters, achieved certain reliability through the process. We believe that if the exercise were to be repeated with different participants, there would be topics that would be similar to those in our study. Of course, the participants' interests were crucial, and, therefore, the wording of the questions and the final topics are bound to be different if the exercise was to be repeated.

8. *The time required to carry it out.* This project had the advantage of being fully funded and, therefore, there was sufficient time for planning and executing it. The timelines were tight, mainly before the workshop started, but everyone gave their time and effort to achieve a common goal.

Landeta (2006) also points to some researchers' problems using the DM, particularly in Social Sciences, whereby experts do not have an emotional or professional link or commitment to those running the study. However, we did not find this problem because all the participants had an intrinsic interest in being active in the workshop to make the best of their participation. On the contrary, we believe that because the participants were ECRs, they were highly motivated and were prepared to do an excellent job despite their short experience as researchers.

We agree with Aubusson et al. (2016) since the method helped construct a future research agenda. The participants were part of the rapidly changing world, and their diversity created an excellent way to integrate different views. The result was a consensus on what could be done in STEM education regarding gender issues; following what Miller (2006) mentioned, the DM attempts to address what can be done instead of the current.

Finally, we see our contribution to knowledge, as follows:

1. Our use of the DM to reach a consensus on a group of ECRs, who are not considered "experts" in the traditional sense: To the best of our knowledge, this has not been studied before, as noted in our review of the DM above. We see the success in applying the DM with this particular group as a combination of their interest in the field, engagement with the activity, and sufficient preparation before beginning to produce questions and ranking them.
2. The outcomes of using this method (i.e., the research strands and proposals) are usable. They reflect essential issues in the area that need to be addressed by research, as judged by the mentors and by the fact that recent published reviews on the field of gender education are closely related to the four research topics that arose from the DM exercise. Other investigators might want to use

our topics or questions to produce their own research proposals or use them to suggest future research agendas.

- The method used in this paper can serve as a model to develop courses in graduate research studies to develop their students' skills and competencies in producing fundable proposals that can help them kickstart their academic careers.

The limitations of this study are related to its implications. The use of DM with early-career researchers was a success story, but the result is an agenda on which to work. DM is rich in producing consensus to integrate different or various points of view but is not meant to be a tool to produce research products. The technique is a start-point to continue working to tackle what was defined in the agenda. The limitation is that it cannot be used to follow up with the agenda. The implication related to this limitation is that the agenda is ready to be implemented. The different research topics resulting from the DM should be implemented after the DM has been used.

The opportunities that this work opens are related to using this method as a model for courses or workshops for graduate students. A course or an academic workshop will help them develop competencies to produce research proposals on essential issues in their fields. In terms of further research, it is necessary to conduct follow-up research to look at whether the proposals are being submitted and study the individual competencies of each individual who participated.

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

Funding: This study was supported by the Newton Fund Researcher Links Workshop grant number 2016-RLWK7-10109, through a partnership between the British Council, UK, and Consejo Nacional de Ciencia y Tecnología, Mexico.

Acknowledgements: The authors acknowledge the technical and financial support of Writing Lab, Institute for the Future of Education, Tecnológico de Monterrey, Mexico, in the production of this work.

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

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

REFERENCES

- Ahmad, S. (2017). Family or future in the academy? *Review of Educational Research*, 87(1), 204-239. <https://doi.org/10.3102/0034654316631626>
- Alcock, L., Ansari, D., Batchelor, S., Bisson, M.-J., De Smedt, B., Gilmore, C., Göbel, S. M., Hannula-Sormunen, M., Hodgen, J., Inglis, M., Jones, I., Mazzocco, M., McNeil, N., Schneider, M., Simms, V., & Weber, K. (2016). Challenges in mathematical cognition: A collaboratively-derived research agenda. *Journal of Numerical Cognition*, 2(1), 20-41. <https://doi.org/10.5964/jnc.v2i1.10>
- Allen, G. (2011). Early intervention: The next steps. *The Stationery Office*. <https://www.gov.uk/government/publications/early-intervention-the-next-steps--2>
- Archer, L., & DeWitt, J. (2014). Science aspirations and gender identity: Lessons from the ASPIRES project. In E. K. Henriksen, J. Dillon, & J. Ryder (Eds.), *Understanding student participation and choice in science technology and education* (pp. 89-102). Springer. https://doi.org/10.1007/978-94-007-7793-4_6
- Archer, L., Dawson, E., DeWitt, J., Seakins, A., & Wong, B., (2015). 'Science capital': A conceptual, methodological, and empirical argument for extending Bourdieusian notions of capital beyond the arts. *Journal of Research in Science Teaching*, 52(7), 922-948. <https://doi.org/10.1002/tea.21227>
- Aubusson, P., Panizzon, D., & Corrigan, D. (2016). Science education futures: "Great potential. Could do better. Needs to try harder". *Research in Science Education*, 46, 203-221. <https://doi.org/10.1007/s11165-016-9521-2>
- Baker, D. R. (2013). What works: Using curriculum and pedagogy to increase girls' interest and participation in science. *Theory Into Practice*, 52(1), 14-20. <https://doi.org/10.1080/07351690.2013.743760>
- Baker, D. R. (2016). *Understanding girls. Quantitative and qualitative research*. Sense Publishers. <https://doi.org/10.1007/978-94-6300-497-8>
- Brady, S. R. (2015). Utilizing and adapting the Delphi Method for use in qualitative research. *The International Journal of Qualitative Methods*, 14(5), 1-5. <https://doi.org/10.1177/1609406915621381>
- Cabrera, N. L., Karaman, A. K., Ballysingh, T. A., Oregon, Y. G., Gonell, E. A., Lopez, J. D., & Deil-Amen, R. (2021). Race without gender? Trends and limitations in the higher education scholarship regarding men of color. *Review of Educational Research*, 92(3), 331-369. <https://doi.org/10.3102/00346543211054577>
- Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T., & Thacker, I. (2020). Teachers' bias against the mathematical ability of female, Black, and Hispanic students. *Educational Researcher*, 49(1), 30-43. <https://doi.org/10.3102/0013189x19890577>
- Curran, F. C., & Kellogg, A. T. (2016). Understanding science achievement gaps by race/ethnicity and gender in kindergarten and first grade. *Educational Researcher*, 45(5), 273-282. <https://doi.org/10.3102/0013189x16656611>
- Custer, R. L., Scarella, J. A., & Stewart, B. R. (1999). The modified Delphi technique: A rotational

- modification. *Journal of Vocational and Technical Education*, 15(2), 1-10. <https://doi.org/10.21061/jcte.v15i2.702>
- Dalkey, N. C. (1969). An experimental study of group opinion. *Futures*, 1(5), 408-426. [https://doi.org/10.1016/S0016-3287\(69\)80025-X](https://doi.org/10.1016/S0016-3287(69)80025-X)
- Fowles, J. (1978). *Handbook of futures research*. Greenwood Press.
- Galeshi, R. (2013). Women and nontraditional fields: A comprehensive review. *Journal of Sustainability Education*, 4(Winter), 1-13. http://www.susted.com/wordpress/content/women-and-nontraditional-fields-a-comprehensive-review_2013_02/
- Garavalia, L., & Gredler, M. (2004). Teaching evaluation through modeling: Using the Delphi technique to assess problems in academic programs. *American Journal of Evaluation*, 25(3), 375-380. <https://doi.org/10.1177/109821400402500307>
- Gonzalez-Garcia, F., Blanco-Lopez, A., Espana-Ramos, E., & Franco-Mariscal, A. (2021). The nature of science and citizenship: A Delphi analysis. *Research in Science Education*, 51, 791-818. <https://doi.org/10.1007/s11165-018-9817-5>
- Gordon, T. J., & Helmer, O. (1964). Report on a long-range forecasting study. RAND paper P-2982. RAND Corporation. <https://www.rand.org/pubs/papers/P2982.html>
- Gregory, A., Skiba, R. J., & Mediratta, K. (2017). Eliminating disparities in school discipline: A framework for intervention. *Review of Research in Education*, 41(1), 253-278. <https://doi.org/10.3102/0091732x17690499>
- Grime, M. M., & Wright, G. (2016). Delphi method. In *Wiley StatsRef: Statistics reference online*. Wiley. <https://doi.org/10.1002/9781118445112.stat07879>
- Guglyuvatyy, E., & Stoianoff, N. (2015). Applying the Delphi method as a research technique in tax law and policy. *Australian Tax Forum*, 30(1), 179-204. <https://ssrn.com/abstract=2597916>
- Hallowell, M. R., & Gambatese, J. A. (2010). Qualitative research: Application of the Delphi method to CEM research. *Journal of Construction Engineering and Management*, 136(1), 99-107. [https://doi.org/10.1061/\(asce\)co.1943-7862.0000137](https://doi.org/10.1061/(asce)co.1943-7862.0000137)
- Henderson, C., Connolly, M., Dolan, E. L., Finkelstein, N., Franklin, S., Malcom, S., Rasmussen, C., Redd, K., & John, K. (2017). Towards the STEM DBER alliance: Why we need a discipline-based STEM education research community. *International Journal of STEM Education*, 4(14), 1-4. <https://doi.org/10.1002/jee.20168>
- Hsu, C.-C., & Sandford, B. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12(12), 10. <https://doi.org/10.7275/pdz9-th90>
- Ireland, D. T., Freeman, K. E., Winston-Proctor, C. E., Delaine, K. D., McDonald Lowe, S., & 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. M., 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>
- Kamerlin, L., Yates, L., Kell, D., Donald, A., McCoy, R., & Tregoning, J. (2019). How to win a research grant. *THE World Ranking Universities*. <https://www.timeshighereducation.com/features/how-win-research-grant>
- Kilburn, D., & Earley, J. (2015). Disqus website-based commenting as an e-research method: Engaging doctoral and early-career academic learners in educational research. *International Journal of Research & Method in Education*, 38(3), 288-303. <https://doi.org/10.1080/1743727X.2015.1026253>
- Kim, A. Y., Sinatra, G. M., & Seyranian, V. (2018). Developing a STEM identity among young women: A social identity perspective. *Review of Educational Research*, 88(4), 589-625. <https://doi.org/10.3102/0034654318779957>
- Kim, H. (2016). Inquiry-based science and technology enrichment program for middle school-aged female students. *Journal of Science Education and Technology*, 25(2), 174-186. <https://doi.org/10.1007/s10956-015-9584-2>
- Kulturel-Konak, S., D'Allegro, M., & Dickinson, S. (2011). Review of gender differences in learning styles: Suggestions for STEM education. *Contemporary Issues in Education Research*, 4(3), 9-18. <https://doi.org/10.19030/cier.v4i3.4116>
- Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting & Social Change*, 73, 467-482. <https://doi.org/10.1016/j.techfore.2005.09.002>
- Linstone, H. A., & Turoff, M. (2002). *The Delphi method: techniques and applications*. <https://web.njit.edu/~turoff/pubs/delphibook/index.html>
- Lorenzo, M., Crouch, C. H., & Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics*, 74(2), 118-122. <https://doi.org/10.1119/1.2162549>
- Ludwig, B. (1997). Predicting the future: Have you considered using the Delphi methodology? *Journal of Extension*, 35(5), 1-4. <https://archives.joe.org/joe/1997october/tt2.php>

- Miller, L. E. (2006). Determining what could/should be: The Delphi technique and its application. In *Proceedings of the 2006 Annual Meeting of the Mid-Western Educational Research Association*.
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management*, 42(1), 15-29. <https://doi.org/10.1016/j.im.2003.11.002>
- Quinn, D. M., & Cooc, N. (2015). Science achievement gaps by gender and race/ethnicity in elementary and middle school. *Educational Researcher*, 44(6), 336-346. <https://doi.org/10.3102/0013189x15598539>
- Razo, M. L. (2008). La inserción de las mujeres en las carreras de ingeniería y tecnología [Women in engineering and technological degrees]. *Perfiles Educativos [Educational Profiles]*, 30(121), 63-96. <https://doi.org/10.22201/iisue.24486167e.2008.12.1.61038>
- Rensaa, R. J., Hogstad, N. M., & Monaghan, J. (2020). Perspectives and reflections on teaching linear algebra. *Teaching Mathematics and its Applications*, 39(4), 296-309. <https://doi.org/10.1093/teamat/hraa002>
- Rowan, L., Bourke, T., L'Estrange, L., Lunn Brownlee, J., Ryan, M., Walker, S., & Churchward, P. (2021). How does initial teacher education research frame the challenge of preparing future teachers for student diversity in schools? A systematic review of literature. *Review of Educational Research*, 91(1), 112-158. <https://doi.org/10.3102/0034654320979171>
- Saw, G., Chang, C.-N., & Chan, H.-Y. (2018). Cross-sectional and longitudinal disparities in STEM career aspirations at the intersection of gender, race/ethnicity, and socioeconomic status. *Educational Researcher*, 47(8), 525-531. <https://doi.org/10.3102/0013189x18787818>
- Talanquer, V. (2014). DBER and STEM education reform: Are we up to the challenge? *Journal of Research in Science Teaching*, 51(6), 809-819. <https://doi.org/10.1002/tea.21162>
- UNESCO. (2016). *Closing the gender gap in STEM: Drawing more girls and women into science, technology, engineering, and mathematics*. <http://unesdoc.unesco.org/images/0024/002457/245717E.pdf>
- Valla, J. M., & Williams, W. M. (2012). Increasing achievement and higher-education representation of under-represented groups in science, technology, engineering, and mathematics fields: a review of current K-12 intervention programs. *Journal of Women and Minorities in Science and Engineering*, 18(1), 21-53. <https://doi.org/10.1615/JWomenMinorScienEng.2012002908>
- Wang, M., & Degol, J. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, 29(1), 119-140. <https://doi.org/10.1007/s10648-015-9355-x>
- Whittaker, J. A., & Montgomery, B. L. (2014). Cultivating institutional transformation and sustainable STEM diversity in higher education through integrative faculty development. *Innovative Higher Education*, 39(4), 263-275. <https://doi.org/10.1007/s10755-013-9277-9>
- Wynekoop, J. L., & Walz, D. B. (2000). Investigating traits of top performing software developers. *Information Technology & People*, 13(3), 186-197. <https://doi.org/10.1108/09593840010377626>

<https://www.ejmste.com>