International Journal of
Gender, Science and Technology
http://genderandset.open.ac.uk

Selected papers presented at the 5th Network Gender & STEM Conference, 29–30 July 2021, in Sydney, Australia

In association with

NETWORK GENDER & STEM
educational and occupational pathways and participation

Gender-Inclusive Instructional Practices in University Mathematics Classes

Oksana Kavatsyuk, Maria Ioannou & Lucy Avraamidou
University of Groningen, The Netherlands

ABSTRACT

Women have been underrepresented in STEM fields around the world (UNESCO, 2018). Prior research identified some of the reasons for this gender disparity such as systemic barriers, lack of confidence, lack of female role models, and cultural and gendered science stereotypes. These issues have been framed in contemporary literature within the construct of STEM identity. Building upon this literature, our project explored the role of the university classroom in supporting the development of a strong STEM identity and specifically the view of self as a competent science person. The project consisted of two parts. In the first part student-led desktop and empirical research focused on generating evidence-based recommendations for how an introductory Calculus course could be redesigned to be more gender-inclusive. The second pertained to the evaluation of the redesigned Calculus course centered around two main indicators of success: (i) students' confidence as mathematics learners, and (ii) their intention to continue with STEM education. The project has scientific and practical implications as it contributes evidence towards understanding the kinds of activities that might support university students' STEM identity development and provides a set of concrete, evidence-based, gender-inclusive instructional practices.

KEYWORDS

gender; STEM identity; educational practices; higher education; the Netherlands

Gender-Inclusive Instructional Practices in University Mathematics Classes

INTRODUCTION – THE PROBLEM

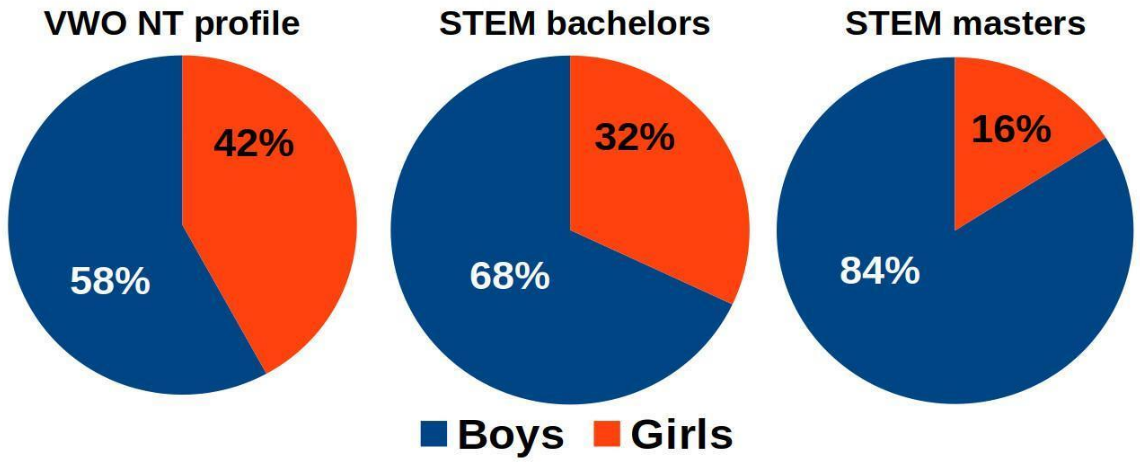
The underrepresentation of girls in Science, Technology, Engineering, and Mathematics (STEM) has been a concern for social scientists and policy makers worldwide (Microsoft, 2017; UNESCO, 2018; World Economic Forum, 2015). Girls have proven to be as competent as boys in STEM worldwide (OECD, 2003) and also in the Netherlands (Hyde et al., 1990), the context of this study. In Dutch education, however, a much lower percentage of girls as opposed to boys choose STEM studies in high school and university (Statistics Netherlands -CBS-, 2019), as illustrated in Figure 1[1].

The Netherlands, as indicated in UNESCO's (2018) latest report on gender equality, has a very low proportion of female STEM researchers when compared to other countries with only one in four researchers being a woman. The Dutch government has invested in attracting girls to STEM in the last decade, through various interventions aiming at higher enrollments of girls into STEM bachelor programs. These interventions yielded some positive changes in involving diverse talents in STEM especially at the level of primary and secondary education (e.g., projects "DigiVita", "Talent Viewer", "Girlsdays", Role model database).

One of the initiatives focusing on higher education was by the national (Dutch) expert organization on girls/women and science/technology (VHTO) who developed "Gender Scan" (Booij et al., 2011) to advise research and applied sciences universities on how to attract more women in STEM. From available reports, however, one can conclude that very few changes were introduced at the research universities as the focus remained exclusively to enrol more girls but not to retain them. Little to no attention was paid to changing the working or studying culture and the (not gender-responsive) content. Perhaps the fact that these programmes are typically "made by men to teach men" (van Uffelen, 2018), that they are, in other words, not sensitive to gender and gender differences, may be one reason to explain the even greater gender disparity in postgraduate STEM programmes, as the CBS cohort data we present in Figure 1 illustrate.

Prior research has identified a number of factors that lead to the under-representation of girls/women in STEM, including girls' lower self-confidence (belief in own talents and qualities), lack of female role models, non-stimulating learning environments, cultural stereotypes in society about girls/women and STEM (Bøe et al., 2011; Ceci & Williams, 2010; Eccles, 2007; Watt et al., 2006). Recent research by Microsoft (Microsoft, 2017) revealed that the Netherlands is one of the few countries where cultural stereotypes about gender and science are very strong. Popular stereotypical beliefs are: "STEM is not for women, too complicated, boring and dirty", "Girls are hard-working, but do not have talent for STEM" (Gender4stem Erasmus project, 2017). Such stereotypes can be already be internalised during childhood (Olsson & Martiny, 2018).

Figure 1
Gender breakdown for high school (VWO type) "Nature and Technology" (NT) profile and STEM Bachelor's and Master's degree holders in 2017-2018. Numbers derived from CBS cohort data (Statistics Netherlands (CBS), 2019).



Since there is evidence that gender disparity becomes even larger at the university level (Figure 1) this project focuses on how university classrooms can become more gender-inclusive in an attempt to address the problem of under-representation and retention of girls in STEM fields. More specifically, our study aimed at re-designing a mathematics course (Calculus 1) using evidence-based recommendations derived by a student-led project. In redesigning this mathematics course, students, under their supervisors' guidance, reviewed the literature to study the problem of gender disparity in STEM and to identify best practices for possible interventions. They coupled their desktop research with primary research through surveys and semi-structured interviews with students who took the specific course in the past, the course's teacher, and other mathematics' teachers. Through this input the course was redesigned, implemented, and evaluated.

In this paper we present this process of redesigning and evaluating a university calculus course starting with the theoretical framework and moving on to a more detailed presentation of the different steps this inquiry followed and the results of each step before concluding with a critical discussion.

METHOD

Context of the Study

The setting of our inquiry is a Liberal Arts and Sciences college in a university in the Netherlands. The course Calculus 1 is a typical Year 1 undergraduate course in many STEM programs and in our context of study it is offered mainly to students who opt for a Sciences major. The original design of the Calculus course did not facilitate gender inclusivity in any way. For example, only white male scientists were mentioned in the textbook and the practical exercises were very abstract and without links to real-world problems.

The inquiry comprised of the following three steps: Step 1 was carried out as a project-based inquiry for which student interns were given the problem (under-representation of women in STEM courses and careers) and the research question of how a mathematics course can change to foster inclusivity with special attention to STEM identity. Step 1 resulted in research-informed and evidence-based recommendations for how the course can be re-designed. Thereafter, proposed interventions were adopted and trialed in the next iteration of the course (Step 2) upon which we performed an evaluation of the intervention based on identified success indicators (Step 3).

Step 1: Project-based inquiry

For the project-based inquiry we recruited five students of the University College Groningen as part of their third-year internships who had a reported interest in gender-science stereotypes or in the under-representation of women in STEM fields. Three of these students had taken the Calculus course and two students did not. This gave us a mixture of insiders and outsiders to our specific case study.

Students were given the general prompt: "Why are women under-represented in STEM?" and the more specific prompt: "Which factors compromise students' STEM identity development and how these could be tackled to foster inclusivity and a stronger sense of belonging in the mathematics classroom". To address these, students were guided to perform a thorough literature review to understand the problem of underrepresentation of women in STEM fields. They were also asked to zoom into the concept of STEM identity in order to design tools (surveys and interview protocols) for performing primary research in order to examine more closely the case study in question (the specific course and its students).

Following the development of these tools, an online survey was administered to the students taking the Calculus course prior to the course redesign. The survey questions can be found in Appendix A. The first survey was completed by 13 students out of 40 (54 % female, 46 % male) and in-depth interviews were conducted with six of these students (four females, one male, one non-binary) and three teachers (the course teacher and two teachers teaching Calculus 1 in other faculties). Interviews were conducted as semi-structured interviews; the questions can be found in Appendix B (student interviews) and Appendix C (mathematics teachers' interviews). The open-ended survey questions and interviews were analyzed using thematic analysis (Braun & Clarke, 2006) with a combination of top-down (codes drawn from Hazari et al.'s model) and bottom-up coding.

Simultaneously, the internship students were asked to register projects or interventions implemented to increase participation of underrepresented groups (especially girls) in science and identify best practices. As a team, we proceeded to evaluate the results of the survey and interviews and in consultation with our bank of identified best practices we proceeded to produce a series of evidence-based recommendations to redesign the course Calculus I.

Step 2: Intervention

Step 1 resulted in evidence-based recommendations on pedagogical methods aiming at a gender-inclusive classroom that would foster STEM identities. These recommendations were applied to redesign the Calculus 1 course, taught the following academic year. The planned interventions were discussed within our team (Calculus teacher, one social psychologist, and one science education expert), with educational professionals embedded at the faculty, and mathematics teachers from other faculties. Suggested changes and interventions were critically reviewed and the most suitable selected.

Step 3: Evaluation

The effectiveness of implemented interventions was evaluated by means of a post-intervention survey and interviews with students and teachers. The post-intervention survey was modelled after the survey that was given to students who had taken the course before the intervention. As such, the results of this first survey functioned as baseline (pre-intervention) measures. The identified indicators of success were: (i) changes in confidence in doing mathematics, and (ii) changes in intention to enroll in advanced STEM courses.

RESULTS

Step 1: Project-based inquiry

Primary research: surveys and interviews with former students. The results of the quantitative part of the survey with students who had taken the original Calculus 1 course (prior to the intervention) are presented as part of Step 3 (evaluation) since they functioned as point of comparison for the purposes of the evaluation of the redesigned course. What we present here instead, are key findings from the open-ended part of the survey and the interviews with a smaller sample of these students. Our analysis was guided by the questions pertaining to factors affecting students' confidence in doing mathematics and/or the identification as math people (see STEM identity). Specifically, we used the responses of participants to questions 8 and 10 of the survey (see Appendix A), namely Q8: 'Can you explain what has led you to be confident or not (that) confident about your ability in mathematics?' and Q10: 'Were there specific elements of the course or the way/ environment it was taught (in) that either enhanced or compromised your confidence?'. We also used the responses of students in two of the interview (see Appendix B) questions: Q1: 'What is your relationship to mathematics? Do you consider yourself a Math person?'; Q2: 'How did the course Calculus I affect the way you see yourself (or not) as a math person?'[3]

The factors below were identified using bottom-up coding. We present them in two distinct categories: (i) factors contributing to, and (ii) factors hindering, students' confidence in doing mathematics and/or the extent to which they see themselves as math people:

(i) Factors that facilitate

Teaching materials: students mentioned how continuous and scaffolded practice with e.g., exercise sheets, was important to them. There was one student who also commented on how important it is to work towards these exercises within group work. According to this student, this contributed to a better understanding of the material:

> "I particularly enjoyed the fact that we had to prepare before class and do the exercises together during class. I believe that doing exercises together with my peers contributed to a better understanding of the material." *(Student 4, female, interview).*

Recognition from peers. Interestingly another factor identified as contributing to one's confidence or to perceiving oneself as a math person was the moments when their peers asked them for their help. We assume that being asked for help constitutes an act of recognition as a competent learner and therefore as someone who can do mathematics, and it also highlights another benefit of group work in a mathematics classroom as shown in the following quote:

> "After few weeks of the course other students started to ask me for help. I realized that I am good in math. Something I did not think before." (Student 3, female, interview).

Formative feedback. Students mention the importance of receiving formative feedback from the instructor which helped them master higher levels of difficulty (more difficult problems) as time went by, as this student illustrates:

> "The teacher and the teaching assistants were helpful and guided me to figure out the solution on my own while also helping with the topics that were difficult for me. So, by the end of the course, I could do difficult problems on my own" (Student 5, female, interview).

Safe learning environment. Moreover, *s*tudents stressed the importance of the existence of a space (created by the instructor) where it was safe to ask questions without feeling intimidated or insecure about their abilities for doing so. This quote is indicative:

> "I think it's a space where everyone feels safe to ask questions to try and understand and nobody feels stupid that they don't know something. An open environment where you can ask questions, you can engage and talk to the teacher and talk to your peers. You do these tasks that help you understand the material and to come out of the course as a more developed and skilled or understanding whatever" (Student 3, female, interview).

A motivated instructor. Students also mentioned the instructor's enthusiasm and especially their ability to render the course fun which seemed to help balance out the perceived difficulty of the course:

> "The teacher made the course fun to attend despite it being difficult." (Student 5, female, interview).

(ii) Factors that impede

Unfavorable comparisons. Direct comparison with people who were obviously better than oneself seemed to work against students' confidence which is why some students preferred small group work than working in uneven dyads. One student characteristically mentions:

> "I don't like that competition feeling. If we are there to just do our work individually, and to just help each other, and not to be like "oh look, I'm doing this so I can show everyone that I'm really fast. I'm good at this." It's not that" (Student 1, female, interview).

Fear of failing. Looming failure was an important factor that led to feelings of inadequacy and the sense that they are not good enough or not 'cut' for mathematics.

Pre-existing beliefs. Pre-existing beliefs, and specifically, the belief that it takes a natural talent to do mathematics, which some students gathered they did not have. The following quote is interesting as it points to how these beliefs are generated or engrained via the lack of recognition by influential others but it also shows how an affirming experience can counteract these beliefs:

> "In high-school some suggested math is not for me … after taking calculus I realized that maybe I am better at math than I thought and maybe I can do math. I would say it increased my confidence in math for sure." (Student 4, female, interview).

If anything, the findings above show that confidence in doing mathematics is something that preoccupies the students and that they can reflect on moments in their lives or in the course that acted as contributors or obstructors to seeing themselves as people who can do mathematics. The factors identified vary but they essentially boil down to good teaching practices and teachers' attitudes which contribute to a safe classroom environment where it is ok to fail and to increased perceived competence. Also, the importance of being recognized as a competent later by their peers surfaced as in important factor contributing to students' confidence in doing mathematics, thus confirming the importance of recognition by others in line with prior literature (e.g., Avraamidou, 2020).

Identified good practices and interventions. As part of their desktop research, our internship students searched for papers that mention "gender inclusive classroom", "gender and STEM", "gender inclusive practices" and which described recommended practices in STEM courses in higher education (or in high-school, but also applicable in a research university setting). They then listed these interventions and discussed some of the key interventions with former students and teachers in their interviews with them in order to determine among other things their feasibility (ease of implementation) and envisioned impact.

The proposed recommendations are presented under Step 2, however, Step 1 investigations essentially yielded three pillars for the recommended interventions.

(i) Formal (teaching) practices

Formal practices relate to how to teach mathematics so as to foster students' science/mathematics' identity (sense of belonging and confidence in doing mathematics) which would lead them to continue studying in STEM fields. Targeted practices included the interaction of the teacher with students (the way of asking and answering questions in class), modelling respect, and providing support.

(ii) Informal practices

Informal practices refer to extracurricular experiences that would help foster students' science/mathematics identity (encouraging students to see themselves as science persons). Targeted practices included providing opportunities to meet and interact with a diverse set of researchers representing a variety of STEM careers and who can serve as role models.

(iii) Systemic interventions
Systemic interventions relate to ways of making teachers and faculty aware of dysfunctional gender stereotypes pertaining to science/mathematics. Targeted practices included changes in teaching materials: inserting short information pieces about female scientists, real-life applications to illustrate mathematical methods, different assessment forms and formative feedback.

Step 2: Course redesign
Before deciding on concrete changes to the course, we asked former students of the course as well as other calculus teachers what their reaction to the aforementioned areas of intervention was by isolating certain practices falling under each area. We provide indicative responses of our interviewees to showcase their reactions.

Formal teaching practices: Learning activities directly connected to real-world problems. Showing the creative side and societal relevance of STEM with hands-on classes and project-based learning since there is literature showing that girls/ women in particular are specifically interested in social implications of STEM topics (Ceci & Williams, 2010). This student shares the impression that applying knowledge to real-world situations would make the course -content- more attractive:

> "... when you explain that you can use it in a number of jobs, and then it is easier for you to deal with taxes for example, I think then people would be more attracted." (Student 1, female, interview).

Informal practices: Role models (as recommended in Gender4stem Erasmus project, 2017): featuring more female lecturers, teaching assistants, and guest lecturers as well as opportunities to meet and interact with female STEM researchers. This student comments on the importance of (realistic) role models, with whom they can relate:

> "Because usually you need a role model, like people need to expect something and look up to someone. So okay, maybe people look up to "Einstein" or something. But like Einstein is like very far from… So I can be someone who's a little bit closer. I think a professor is like the closest thing to a smaller role model." (Student 1, female, interview).

Informal practices: Peer-mentoring (Dennehy & Dasgupta, 2017). An example of peer mentoring is having a first-year student being mentored by a second- or third-year student to support their educational choices and share learning strategies. This is a practice which was already being trialed out in the course, one that was well-received as this student testifies:

> "We had Q&A sessions for people that have troubles with this class, and I attended them after that class. It was two hours of just me and a teaching assistant or an instructor. They just explained how everything is and I just started to understand everything". (Student 2, female, interview).

Systemic interventions: Removing gender bias from learning materials. As suggested by the UNESCO training module "*Girls into Science: A training Module" (UNESCO, 2007).* Practically, this means mentioning equally both genders in lectures and instruction materials. In case of historical facts (e.g., discoveries by male scientists) additional examples of women in science should be given. Pronouns she/her should be used in the first place, when referring to a general reader.

After receiving feedback from students and teachers on the aforementioned practices, we proceeded to purposefully redesign the course in the following ways: (i) adopted a problem-based approach centered around real-world problems; (ii) introduced female mentor/role-models; (iii) engaged students in hands-on activities that exemplify the application of science to everyday life, and (iv) incorporated explicit discussions about gender-STEM stereotypes as well as the role of women in science. A detailed account of these practices can be found in Appendix D.

Step 3: Evaluation

A post-intervention survey ($N = 7$; 6 females, one male) was used to evaluate the effectiveness of the redesigned course in fostering more confidence in doing mathematics and a desire to continue with STEM studies. Additionally, these students were interviewed to give their qualitative feedback on the redesigned course.

For the survey we focused on two main indicators of success both of which were included in the pre-intervention and the post-intervention survey: (i) confidence in doing mathematics which was measured through the following question combination: Question A: When you were at high school did you think you were good at mathematics and Question B: while taking the calculus course did you think you were good at Mathematics? and (ii) enrollment in more (advanced) STEM courses with the Question: Do you plan to study subjects where mathematics is involved in future? These two success indicators were chosen based on the following rationale: confidence in doing mathematics as a significant contributor to STEM identity (seeing oneself as a math person) served as a proxy for whether the redesigned course contributed to a boost in students' self-concept as math people. The second indicator was a proxy to our over-arching goal, namely attracting more students (especially female) in STEM. The results were as follows.

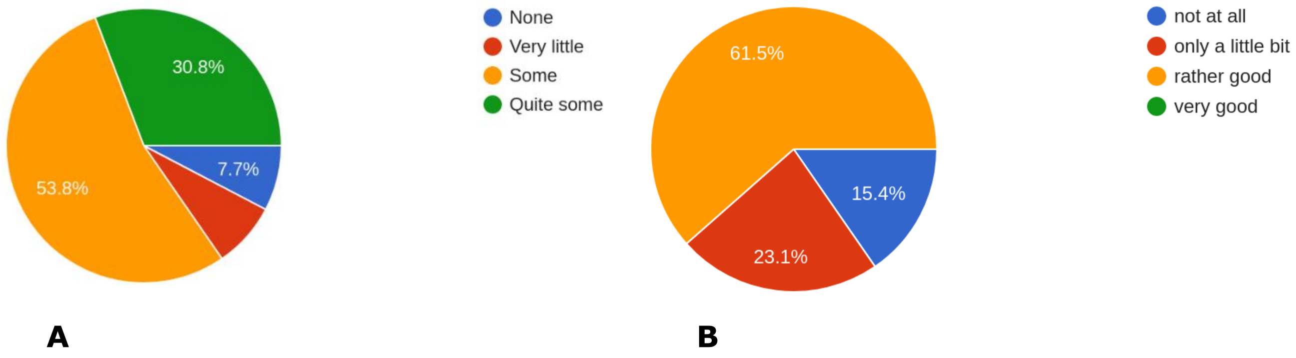
(i) Confidence in doing mathematics

Before the course redesign only a little change was observed between how students evaluated their confidence in doing mathematics when at high school (54% reported being rather good in mathematics, Figure 2) and after the Calculus 1 course (62% reported being rather good in mathematics, Figure 2). The number of students who thought they were not good at mathematics increased slightly and the number of students who thought they were very good decreased slightly.

After the intervention we observed an increase in self-confidence. A similar percentage of students 57 % thought they had "some" talent for mathematics in high-school. After the course all respondents (100%, 7 students) believe they are "rather good" in mathematics.

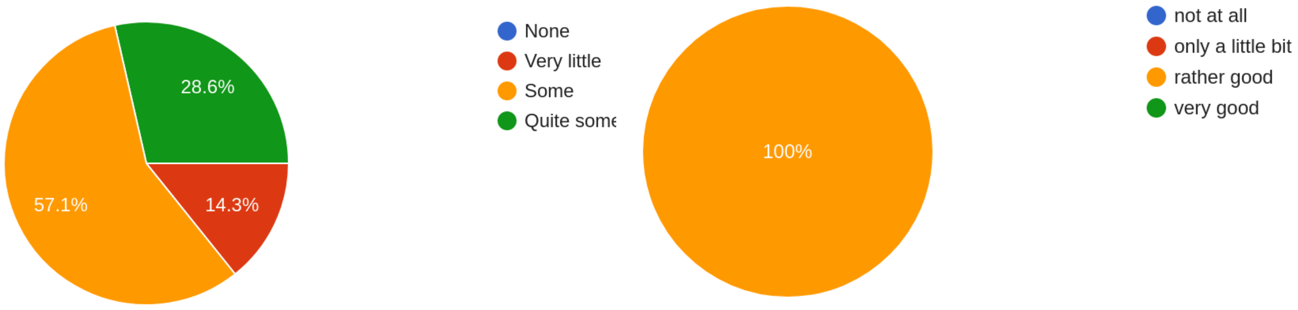
Importantly, the first (Survey 1) and the second (Survey 2) cohorts named different reasons as important to be successful in mathematics. In Survey 1 (pre-redesign) 54% considered "talent" as important, in contrast to just 14 % in Survey 2 (post-redesign). After the Calculus 1 course redesign, students indicated “supportive environment” and “collaborative classroom” as key-factors for success (38% in Survey 1 versus 71% in Survey 2). These are important findings as they speak to a change in perspective on who is/ can be a mathematics person. The post-redesign cohort were less likely to think that only talented people can be good at and do mathematics.

Figure 2
Results of the survey before the course redesign.
Question A: When you were in high school did you think you had a talent for mathematics?
Question B: While taking the Calculus I course, did you think you were good at mathematics?



A **B**

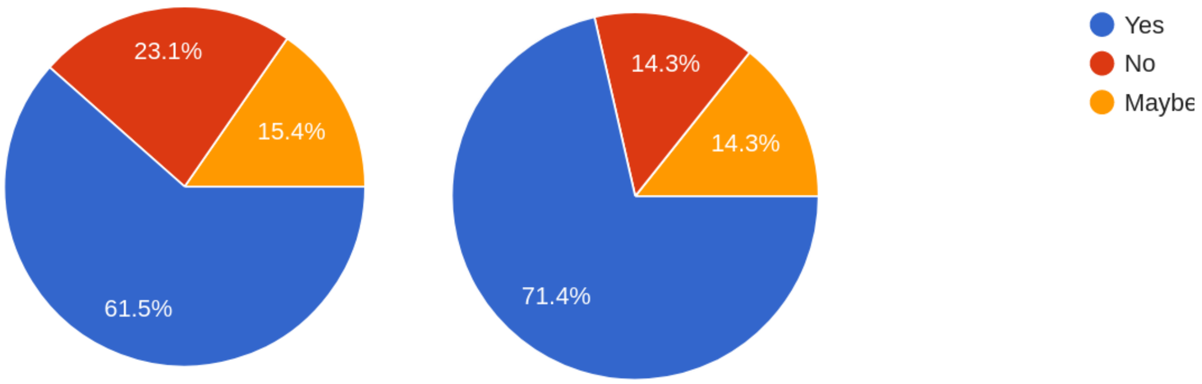
Figure 3
Results of the survey after the course redesign.
Question A: When you were in high school did you think you had a talent for mathematics?
Question B: While taking the Calculus I course, did you think you were good at mathematics?



A B

(ii) Opting for more STEM courses
Changes in confidence and a perception of who can be a mathematics person seemed to positively influence the interest in following higher level mathematics and STEM courses (Figure 4). After the course redesign we observed about a 10% increase in the number of students that plan to study subjects where mathematics is involved in the future.

Figure 4
Answers to the question "Do you plan to study subjects where mathematics is involved in future?"
A: Results of the survey 1 (before the course redesign)
B: Results of the survey 2 (after the course redesign)



A B

Student interviews focusing on soliciting more concrete feedback on what worked well in the redesigned Calculus 1 course and what specific and systemic changes can be implemented to teach mathematics in a way that would increase students' confidence in doing mathematics yielded the following:
1) teachers' support and encouragement: positive feedback is immensely helpful as it boosts self-confidence but it should be very specific, avoiding general statements;
2) teachers' gender is not important as long as they teach with enthusiasm and excitement;
3) teachers of all genders should try to serve as role models, sharing personal experiences regarding their way into science (including difficulties and successes);
4) creating a safe atmosphere in the classroom where students feel it is ok to fail (e.g., encourage students to try solving problems, even if they do not get the correct answer);
5) including real-world problems and applications underlining the importance of mathematics for society (e.g. Calculus methods are used in MRI imaging to get images from the measured data);
6) working in small groups where students feel at ease to ask for help but also feel good about being asked for help;
7) checking for gender biases in learning materials (e.g., are both female and male scientists mentioned? are the problems given interesting for all students?).

DISCUSSION

In our paper we describe the journey of redesigning a university mathematics course to render it more gender inclusive in an attempt to respond to the challenge of underrepresentation of females in STEM worldwide but also in NL (UNESCO, 2018). The context of our study was a Liberal Arts and Sciences faculty in a Dutch university and more specifically the Calculus I course offered to its science major students. Our research was guided by the question of how a mathematics course, in the given context, can change to become more gender inclusive?

Drawing from research in STEM identity we narrowed this question further to specifically ask what type of changes would enhance students' perception of themselves as math people which was mainly approached through the construct of confidence (competence in doing mathematics; Hazari et al., 2010). As part of the empirical part of this inquiry we tried to understand which in the students' experience were the main contributors and obstructors to considering themselves as competent in mathematics. This knowledge was coupled with desktop research aiming at registering prior interventions that addressed the problem of inclusivity in a mathematics classroom. This combination of primary and secondary research ended up informing the recommendations which were implemented in Step 2 to redesign the course. The new course was then evaluated based on two indicators: confidence in doing mathematics as a proxy of STEM identity (seeing oneself as a math person), and in intention to register in advanced STEM course, a proxy of interest in STEM.

In this discussion section we highlight the most important points of this inquiry and leave recommendations for researchers and educators pursuing similar endeavors.

Evidence-based and literature-informed interventions

Recent developments in teacher education underline the importance of evidence-based practices as a way of connecting research and practice in teaching (Bauer and Prenzel, 2012; Cutspec, 2004). These developments are grounded in the idea that empirical evidence should be an important source of knowledge for teachers in classrooms (Niemi, 2008). Davies (1999) described evidence-based education as one that is integrating individual teaching and learning expertise with the best available external evidence from systematic research. While educational research cannot change the practices of educators and teachers it does provide foundational knowledge on teaching and learning (Niemi, 2008). A lesson from our experience is that educators undertaking educational research should ideally combine empirical evidence and secondary research to identify, customize, and evaluate best practices in teaching.

More specifically, the process we followed to redesign this course was one that relied on simultaneous top-down and bottom-up processes where the literature was our go to place for best practices but these practices were tailored to the needs of our course. This customization would not have been possible without the input from key stakeholders namely students who had undertaken the course, other Calculus I teachers, and the very teacher of the course.

For example, students' experiences helped to illustrate what was known to us from the literature such as the importance of recognition (Avraamidou, 2020) by mentioning without being prompted in which occasions they derived recognition (when their peers ask them for help) which then reinforced our intuition that small group work is important in a mathematics classroom. Students were also explicit about what is the kind of feedback they want (positive, formative, and specific) as opposed to generic positive feedback, which helped us inform our teaching practices especially as regarding modes of communication with students. They also helped us compliment identified best practices by, for example, placing emphasis on the importance of a safe environment and what constitutes such an environment (feeling safe to ask questions).

Gender specific recommendations for a gender-inclusive classroom?

In their final round of feedback students who had undertaken the revised course highlighted what for them were best practices. What was interesting about the highlighted practices is that they were not directly gender-focused. However, literature suggests that some educational best practices aiming at high quality of instruction have more impact on female students. The recent work by Blanchard and Riegle-Crumb (2017), for example, has shown that perceptions of the social relevance of science positively and significantly predict female students' plans to major in certain STEM fields (the biological sciences, the physical sciences, and engineering). For these fields male students' intentions are not impacted.

Our observations and conclusions agree with work by Skipper and Leman (2017) that concluded that teachers' feedback is more important for girls than boys to support their achievements in STEM. Even the best female students rely less on their grades but more on the feedback of their teacher when it comes to assessing their knowledge (Skipper & Leman, 2017).

The importance of the supportive environment for female students is also underlined in the literature. She (2000) observed that boy/girl differences in learning style and classroom participation are reinforced or sustained by her behavior. These differences include unequal feedback and encouragement for male and female students and different amounts of direct questions.

Limitations

Before concluding we wish to acknowledge some important limitations of our study. The first concerns a lack of complete coherence between our theoretical framework and the actual process. Our departure point, as noted in the first part of our paper, was the construct of STEM identity and its contributing components as researched by especially Hazari et al., (2010). While we tried to use Hazari et al.'s (2010) work as our analytical framework we failed to do that with great consistency. In fact, the construct of STEM identity more broadly and Hazari et al.'s work particularly mostly informed our inquiry by: (i) helping us to first and foremost understand what we were after when aiming at a gender-inclusive course and (ii) designing our survey and interview protocol; (iii) functioning as a filter when discussing possible interventions, and (iv) helping us inform the indicators of success of our intervention (confidence in mathematics and interest in pursuing more mathematics courses). We attribute the lack of complete coherence, however, to the fact that this endeavour was largely managed by students themselves whom we, as supervisors, could only advise but not entirely steer. On the other hand, we consider this student-led process to be one of the innovations of our project: a project led by students for students.

A second major limitation concerned the evaluation design which did not follow a rigorous methodology. For example, the pre- and post-test were done with different cohorts (not deploying a within-participant design); essentially students from the previous year functioned as our control group. More importantly though, the sample size was very small which does not warrant any safe conclusions about the comparisons performed and it did not allow for any statistical tests. To counter that though, we conducted interviews to generate qualitative data to support and nuance our quantitative data.

Finally, the reported research was conducted within a small-classroom-based Liberal Arts and Sciences program with a larger percentage of female students than in a typical STEM programme in sciences faculties. The group dynamic in a big classroom with very few females might be very different than in this course's environment. It is therefore important to examine the proposed best practices with a larger cohort of students in a typical STEM program (not small-classroom teaching and with a different balance of male-female students).

Conclusion

To conclude, the evaluation of our student-led intervention notwithstanding its methodological setbacks pointed towards some positive results on two indicators of positive change: one being confidence in doing mathematics, and the second being higher enrollment in advanced STEM courses. These results suggest that specific aspects of the course as well as the role of the teacher were instrumental in the development of the students' competence and interest, two important factors of STEM identity. The changes we implemented in a Calculus 1 course are easy to implement, especially in small-scale settings, and do not require a lot of the instructors' time. The results of our intervention indicate that these changes can yield positive impacts for both male and female mathematics students.

AUTHOR NOTE

The authors have no conflict of interest to disclose. We would like to thank E. E. Dryden Silva, J. A. Jacukowicz, A. M. Ostaf, M. Pokerznik and A. M. Punter for helping with student surveys and interviews and suggesting the most effective (in students' view) interventions. We thank I. M. Svevad for analyzing the teachers' interviews. We also thank the students of the Calculus and Mathematics courses especially those who took the time to respond to our surveys and participate in interviews. Finally, we thank the anonymous reviewers for their patience and very constructive comments. O. Kavatsyuk and M. Ioannou are from University College Groningen; L. Avraamidou from the Faculty of Science and Engineering. Please direct correspondence to:
Oksana Kavatsyuk at University College Groningen, University of Groningen, Hoendiepskade 23/24, 9718 BG Groningen, The Netherlands. E: o.kavatsyuk@rug.nl.

ENDNOTES

1 The data we present in Figure 1 are from publicly available datasets offered by the Statistics Netherlands Authority (CBS). We have considered the high-school profile "Nature and Technology" (NT) of the VWO (Voorbereidend Wetenschappelijk Onderwijs), the high school track that prepares pupils for university education. For the STEM Bachelor's and Master's we requested the dataset about the diplomas issued in 2017-2018 in the following fields: Category 1 "Mathematics, Natural Sciences", Category 2 "Informatics, Technology", Category 3 "Technology, Industry and Engineering".

2 We chose to use STEM as an umbrella term and we acknowledge that not all of the studies we cite in this section use the term STEM when studying identity. For example, Carlone and Johnson (2007) use science identity, while Hazari et

al.’s (2010) research focusing on physics undergraduate students, studies physics -and not STEM- identity.

3 We acknowledge that the two sets of questions tap onto slightly different constructs: the first two focus on confidence in doing mathematics and was inspired by the performance/ competence factor in Hazari et al.’s work (2010); and the other two tap onto self-concept. Although the two (competence and self-concept) are not one and the same, and perhaps should not be used interchangeably; confidence/ competence was theorized and empirically shown to be a significant contributor to self-concept – seeing oneself as a science person (see Hazari et al., 2010). Also, the two sets of questions explicitly asked students to point at factors contributing or inhibiting their confidence in abilities and / self-concept.

REFERENCES

Avraamidou, L. (2020). Science identity as a landscape of becoming: Rethinking recognition and emotions through an intersectionality lens. *Cultural Studies of Science Education, 15*, 323-345. DOI : 1007/s11422-019-09954-7.

Bauer, J., & Prenzel, M. (2012). Science education. European teacher training reforms. *Science* 336, 1642–1643. doi: 10.1126/science.1218387

Blanchard Kyte, S., & Riegle-Crumb, C. (2017). Perceptions of the social relevance of science: exploring the implications for gendered patterns in expectations of majoring in STEM fields. *Social Sciences*, 6(1), 19.

Booij, C., Jansen, N., Joukes, G. & Van Schaik, E. (2011). *Trendanalyse gender in het bèta/technisch onderwijs*. Amsterdam: VHTO, Landelijk expertisebureau meisjes/vrouwen en bèta/techniek. https://www.vhto.nl/kennis/onderzoek/rapport-trendanalyse/

Bøe, M. V., Henriksen, E. K., Lyons, T., & Schreiner, C. (2011). Participation in science and technology: Young people’s achievement-related choices in late modern society. *Studies in Science Education, 47*(1), 7-37.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101.

Carlone, H., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching, 44(8)*, 1187–1218.

Ceci, S. J. & Williams, W.M. (2010). *The mathematics of sex. How biology and society conspire to limit talented women and Girls*. New York, USA: Oxford University Press.

Cutspec, P. A. (2004). Bridging the research-to-practice gap: evidence-based education. *Evidence Based Approach. Early Childhood. Dev.* 2.

Davies, P. (1999). What is evidence-based education? *Br. J. Educ. Stud.* 47, 108–121.

Dennehy T.C., Dasgupta N. (2017) Female peer mentors early in college increase women’s positive academic experiences and retention in engineering. *Proceedings of the National Academy of Sciences, 114 (23)*, 5964-5969.

Eccles, J. S. (2007). *Where are all the women? Gender differences in participation in physical science and engineering*’. In S.J. Ceci, & W.M. Williams (Eds.), Why aren't more women in science: Top researchers debate the evidence (pp. 199-210). Washington, DC, USA: American Psychological Association.

Hazari, Z., Sonnert, G., Sadler, P.M. & Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, 47(8), 978-1003.

Hyde, J. S., Fennema, E., Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*, 139-155.

Microsoft (2017). Why Europe’s girls aren’t studying STEM. White paper. https://news.microsoft.com/europe/features/dont-european-girls-like-science-technology/

Niemi, H. (2008). “Advancing research into and during teacher education,” in *Monographs on Journal of Research in Teacher Education. Teacher Education Policy in Europe: A Voice of Higher Education Institutions*, eds. B. Hudson and P. Zgaga (Umeå: University of Umeå, Faculty of Teacher Education), 183–208.

OECD (2003). PISA dataset 2003. Paris, France: OECD. https://www.oecd.org/pisa/pisaproducts/database-pisa2003.htm

Olsson, M., & Martiny, S. E. (2018). Does Exposure to Counter-stereotypical Role Models Influence Girls' and Women's Gender Stereotypes and Career Choices? A Review of Social Psychological Research. *Frontiers in psychology, 9*, 22-64.

She, H. C. (2000). The interplay of a biology teacher's beliefs, teaching practices and gender-based student-teacher classroom interaction. *Educ. Res*. 42, 100–111. doi: 10.1080/001318800363953

Skipper, Y., and Leman, P. J. (2017). The role of feedback in young people's academic choices. *Int. J. Sci. Educ*. 39, 453–467. doi: 10.1080/09500693.2017.1294783

Statistics Netherlands (CBS) data (2019). WO-cohorten, diploma. https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83285NED/table?ts=1568532577091

UNESCO (2007). *Girls into Science: A training Module*. Paris, UNESCO.

UNESCO (2018). *Women in Science*. Fact Sheet No.51. June 2018 FS/2018/SCI/51. http://uis.unesco.org/sites/default/files/documents/fs51-women-in-science-2018-en.pdf

Van Uffelen, C. (2018). How can TU Delft attract more female students? https://www.delta.tudelft.nl/article/how-can-tu-delft-attract-more-female-students

Watt, H. M. G., Eccles, J. S., Durik, A. M. (2006). The leaky mathematics pipeline for girls: A motivational analysis of high school enrollments in Australia and the USA. *Equal Opportunities International, 25*(8), 642-659.

World Economic Forum (2015). *The Global Gender Gap Report 2015.* Geneva, Switzerland: World Economic Forum. https://reports.weforum.org/global-gender-gap-report-2015/