Hypatia (2025), 1–31
doi:10.1017/hyp.2025.10025

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Feminist and Trauma-Informed Approaches to Teaching Formal Philosophy

Francisco Calderón[1], Thomas M. Colclough[2] and Helen Meskhidze[3]

[1]Department of Philosophy, University of Michigan, Ann Arbor, MI, USA, [2]Center for Knowledge, Technology, and Society, University of California, Irvine, CA, USA and [3]Departments of Philosophy and Physics, University of Cincinnati, OH, USA
Corresponding author: Francisco Calderón; Email: fcalder@umich.edu.

(Received 25 September 2024; revised 19 May 2025; accepted 2 June 2025)

Abstract
There has been significant interest in addressing the underrepresentation of various demographic groups in philosophy. Indeed, many have proposed remedies at the disciplinary level. However, underrepresentation is an issue that varies by subfield in philosophy. Women, for example, are especially underrepresented in subfields considered formal (e.g., logic). As has already been argued in the existing literature, addressing underrepresentation, even within subfields, is not as simple as recruiting more students from underserved populations. Instead, we advocate for a student-centered approach, promoting inclusive pedagogy. In this paper, we share a case study in which we implemented feminist and trauma-informed interventions in two undergraduate formal logic courses and investigated their impact with respect to elements of structural injustice. We found that our interventions successfully eliminated existing gender-based differences in perceptions of self-efficacy and largely diminished students' perceptions of the objectivity of logic, but were unsuccessful at changing students' impressions of the broader applicability of logic. By sharing our interventions, we hope to provide educators with practical tools and ideas for implementing similar approaches in their classrooms. By sharing our results, we invite educators to reflect on the potential impact of similar approaches in formal philosophy courses and on tools for measuring that impact.

Introduction

The underrepresentation of women in philosophy is now well-documented and has been quantified numerous times (Schwitzgebel and Jennings, 2017; Beebee and Saul 2011). Studies have tried to further quantify the underrepresentation of women at various points in the pipeline—from undergraduates first taking courses in philosophy, to the graduate student population, to the representation of women authors in major philosophy journals (Paxton et al. 2012; Baron et al. 2015; Thompson et al. 2016; Herfeld et al. 2022; Conklin et al. 2023). Small-scale interventions aimed at ameliorating the

© The Author(s), 2025. Published by Cambridge University Press on behalf of Hypatia, a Nonprofit Corporation. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

situation have also been proposed and investigated (see, e.g., Lockard et al. 2017). Finally, some have investigated the underrepresentation of racial and ethnic groups in philosophy (Schwitzgebel et al. 2021).

Most recently, studies addressing the underrepresentation of women in philosophy have begun to consider the distribution by subfield. They have found that women tend to be especially underrepresented in subfields considered "technical" or "formal." As noted by Kings, "In philosophy itself, you are more likely to find women working in ethics, aesthetics, feminism, and applied philosophy than you are to find them specializing in such areas as logic, metaphysics, epistemology, consciousness, and philosophy of science" (2019, 221). Paxton quantifies this trend in a recent report. She surveyed faculty in leading philosophy departments and found that 10 percent of faculty in "Logic and Philosophy of Logic" and 5 percent of faculty in "Philosophy of Physical Science" self-identify as female (Paxton 2015). These figures are significantly lower than the already low proportion of women faculty (about 20–40 percent, depending on rank, according to Paxton et al. 2012, n. 3). Schwitzgebel and Jennings similarly find that "Science, Logic, and Math" has the fewest women—faculty and graduate students—compared to the other subfields investigated (2017).

The dearth of minority representation in these subfields seems more than accidental. As Thompson conjectures, "it seems a reasonable hypothesis that the underrepresentation of women and Black students in STEM fields as well as in philosophy compounds for philosophy of science" (2021). Writing about her experiences as a philosopher of physics, Ruetsche (2020) reflects on the perceived relationship between feminism and "technical" fields in philosophy. She discusses the presumed incompatibility or, even more strongly, hostility between feminism and these technical fields. Though Ruetsche refutes this incompatibility, it is conceivable that undergraduate women perceive it early in their philosophical education. After all, consider the message being sent by the readings assigned in philosophy of science courses: Thompson (2021) reports from an informal examination of 25 philosophy of science course reading lists that texts from women philosophers of science are rarely assigned, and no texts from Black women philosophers were on any of the syllabi. If women perceive such hostility, it would not be surprising if they avoided such subjects. It is clear that more research is needed on the subfield-specific underrepresentation of women and how to address it.

There are many ways to respond to the issue of underrepresentation. One approach, prominent in STEM, focuses on retention and finds that students' sense of belonging plays a major role in their choice to persist as STEM majors. Comparing various demographic groups, a recent study finds that leavers from underrepresented groups (women and students of color) report a lower sense of belonging (Rainey et al. 2018). Based on interviews with 201 college students of diverse gender and racial backgrounds, Rainey and colleagues find four factors contributing to students' sense of belonging: interpersonal relationships, perceived competence, personal interest, and science identity. The leavers tended to cite "a lack of interpersonal relationships and weak sense of competence" (2018, 11). If these results transfer to philosophy—and indeed, since we are focused here on "technical/formal" philosophy, we have reason to think they do—they suggest that intervening on these factors (i.e., trying to increase students' interpersonal relationships and improve their sense of their own competence) might help retain students.

The above suggestions may work for retaining those students who are already in our classrooms. What about those who are not? One might be tempted to focus on recruitment efforts targeted at underrepresented groups. While important, philosophers Jacquart, Scott, Hermberg, and Bloch-Schulman (2019) suggest it is time to move beyond

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

recruitment efforts. They argue that adopting a student-centered approach in the classroom shifts the focus away from efforts to improve diversity in the classroom and towards inclusive pedagogy.[1] They advocate for fostering a growth mindset and learning-centered approach to authority in the classroom. They also argue that transparency, flexibility, and continuous self-reflection are important for students and teachers.

It was with this proposed shift toward inclusive pedagogy in mind that we designed our study. Here, we investigate two types of inclusive pedagogy—feminist and trauma-informed pedagogy—and ask: To what extent can feminist and trauma-informed interventions change students' perceptions of formal philosophy and/or their abilities? How are these changes different for different demographic groups? We performed a study on two introductory logic courses at two universities with distinct undergraduate populations: the University of California, Irvine and the University of Michigan, Ann Arbor. We designed and administered a diagnostic tool to analyze whether students (women and other underrepresented social groups) at the introductory level exhibit the perceptions of formal philosophy often attributed to them. Our survey addresses interest in and perceptions of the subdiscipline and students' perceptions of their own abilities. The course then proceeded, guided by feminist and trauma-informed practices, which impacted the assessment structure and classroom activities. Finally, we redistributed the diagnostic survey to assess whether the interventions had a differential impact on students' perceptions. We provide evidence that the interventions in this course close gaps in students' perceptions of self-efficacy along gender lines and, generally, change students' perceptions of the objective nature of logic. They do not alter students' perceptions of the broader applicability of formal tools. As a result, and overall, we argue that feminist and trauma-informed pedagogies should be implemented in (at least introductory) formal courses as part of inclusive pedagogy.

Below, we discuss the design and implementation of the diagnostic tool and describe some of the practices we adopted. Then, we analyze the effects of the practices, as evidenced by student responses to the surveys. We highlight the differences between the institutional settings throughout. We hope this discussion is fruitful for those interested in understanding the structural aspects of the underrepresentation of various social groups in various subfields of philosophy and those interested in bringing feminist and trauma-informed practices to their classrooms.

Context and theoretical lenses

We begin by discussing the typical methods used when teaching logic.[2] Following this discussion of the context, we will turn to the theoretical underpinnings of our study, presenting the motivations and principles of feminist and trauma-informed pedagogy. Throughout, we underscore the importance of the social location of our students and their past experiences, even in contexts where they might not initially seem salient, like a logic course.

Context

To understand the typical approaches to teaching logic, we have collected and reviewed various logic syllabi. Though such an approach is limited—syllabi offer only a glimpse into a course—they nonetheless offer a useful starting point. Many of the larger interventions we performed would be evident on a syllabus, so we think an analysis of syllabi will offer worthwhile comparisons to the interventions we suggest.[3]

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

For this analysis, we searched Google for logic syllabi. We considered syllabi for courses titled "Introduction to Logic" and "(Introduction to) Symbolic Logic" that have been used in the past ten years. Our data set consisted of 21 syllabi in total, 13 "Introduction to Logic" syllabi and 8 "Symbolic Logic" syllabi.[4] The 21 syllabi spanned 15 universities. Two were used at Liberal Arts Colleges (Hillsdale College and St Thomas More College), one at an R2 university (College of William and Mary), and the remaining 18 at R1 universities.[5]

Of the 21 syllabi considered, all but two (~90 percent) had exams. The two without exams were both "Introduction to Logic" courses, one of which was taught by a cognitive scientist. While some courses had lower stakes exams (worth 10 percent of the grade or completed as take-home exams), the vast majority featured midterm and final exams worth 25–50 percent of a student's grade. Regarding peer collaboration, one course featured a "Logic Lab," two had group assignments, and one used Discord to encourage classroom community. Four courses used online software for students to practice with (three used those corresponding to *Language, proof, and logic,* and one used Logic2010). Fourteen distinct textbooks were used. The texts were quite diverse, but *A Modern Formal Logic Primer*; *Language, Proof, and Logic*; and *forall x: Calgary: An Introduction to Formal Logic* were each used by three different courses.

In terms of content, six of the 20 courses whose content was listed in the syllabus considered applications to and extensions of the main course content. The typical applications were to arithmetic/set theory, and the typical extensions included probabilistic reasoning/inductive logic. Only two syllabi presented alternatives to classical logic: one mentioned "other logics" as a topic to be discussed at the end of the course, and another had a class devoted to "A brief summary of logic beyond [introduction to symbolic logic]."[6]

A few of the syllabi discuss how to approach the course. The most encouraging of these tells students that learning logic is like mastering a skill and that everyone must practice to acquire it. The others note that the course is rigorous and that the skills developed are cumulative. They advise students not to fall behind and warn that persistent self-discipline will be necessary for success.

Overall, we find that logic courses typically see logic as a skill that students need to practice. Students are asked to demonstrate their skill through exams, but the practice itself is not typically counted towards their grade. Students typically have little agency in choosing course content or assignments. Community is sometimes emphasized, especially as students are developing their skills. However, because so much emphasis is on exams written independently, communities are of limited utility. Finally, students are only sometimes asked to consider the applications of these formal systems and are rarely asked to evaluate their limitations.

Theoretical lenses

Two main theoretical lenses motivate the interventions proposed in this paper: feminist pedagogy and trauma-informed pedagogy. We discuss each below, highlighting the overlaps in their tenets and in the interventions they suggest.

Feminist pedagogy

Building on critical pedagogy, the aims of feminist pedagogy are often described as emancipatory: to democratize classrooms and challenge structures of domination (Light

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

et al. 2015, 14). To this end, feminists suggest critically examining existing instruction methods and adopting methods motivated by a set of feminist tenets, including empowerment, community, and leadership (Shrewsbury 1987). A feminist classroom is also seen as a connected classroom: the instructor and the students should all be connected to one another in a learning community, and the content of the classroom should be connected to students' other interests, classes, and activities. The goal here is a more "holistic" learning experience wherein the contents of a course are not isolated from one's everyday life but put in context with it instead (see Shrewsbury 1987; hooks 1994; Maher and Tetreault 2001; and Lintott and Skitolsky 2016 for arguments along these lines). Reflecting on the liberatory pedagogy of Paulo Freire (2018), Weiler recommends that feminist pedagogy should also "recogn[ize] the importance of personal experience as a source of knowledge" and should explore the perspectives of people of different races, classes, and cultures (1991, 449).

More recently, feminists have advocated for inclusive pedagogies more generally. Indeed, the inclusive pedagogy advocated in Jacquart et al.'s paper includes the following principles, many of which overlap with the tenets of feminist pedagogy: fostering a growth mindset, examining inclusive conceptions of authority, promoting transparency, encouraging flexibility, and continually promoting self-reflection for students and instructors (2019, 107). Though we, too, are motivated by inclusive pedagogy and adopt the principles advocated by Jacquart et al., we retain the terminology of feminist theory, mainly because we anticipate the greatest differences in our study to be along gender lines.

Additionally, we take our feminist lens to highlight the importance of two kinds of pluralism: in approach and in content.[7] By pluralism in approach, we mean whether students think there are multiple, equally viable ways of approaching a problem. Evidence from the mathematics pedagogy literature shows that a cohesive rather than fragmented view of mathematical knowledge results in better learning outcomes (Crawford et al. 1994). A cohesive view will acknowledge the interconnectedness of mathematical ideas and see them as less rigid. We believe a student with such an understanding of logic would also see multiple correct problem-solving approaches. In our survey questions and discussion, we refer to this as the "objectivity of logic" or the "objective nature of logic," and our goal is to decrease students' perceptions of objectivity.

We also aimed to highlight the importance of pluralism with respect to entire logical systems. Logical pluralism, the idea that there is more than correct or best logic, has been defended on many diverse grounds (see, e.g., Hjortland 2017; Blake-Turner and Russell 2021). Here, we draw on Saint-Croix and Cook (2024), who have recently highlighted the need for logical pluralism on feminist grounds. They use Longino's arguments for the importance of contextual values in science (Longino 1990), combined with an anti-exceptionalist framework that sees logic as continuous with science, to argue for logical pluralism.[8] Regardless of whether one fully embraces their view, they (and we) think it is important to teach students alternatives to classical logic. We describe the interventions performed towards this kind of pluralism in the last section of §3.5.

Finally, we acknowledge that pluralism with respect to one's logical system likely implies a kind of pluralism in approach (which we refer to as breaking down the "objectivity of logic"). However, we expect our survey questions to gauge objectivity, not pluralism of logical systems, because the mathematics questionnaire we modeled these questions on also aimed to assess students' perceptions of the rigidity of mathematical knowledge. Additionally, while we introduce alternative logical systems, we believe that

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

expecting students to engage deeply and substantially with logical pluralism is too advanced a goal for an introductory course in logic.

Trauma-informed pedagogy

The second theoretical lens we adopt is trauma-informed pedagogy. Trauma-informed pedagogy begins with the recognition that the majority (85 percent) of college students report having been exposed to a traumatic event in their past (Frazier et al., 2009). Given this fact, and in light of the COVID-19 pandemic, many argue that a trauma-informed approach is vital to creating inclusive, transformative classrooms that encourage collaboration and intellectual risk-taking (Carello and Thompson 2021).

In this literature, traumas are understood broadly. An individual is said to have experienced a traumatic event if they "experienced, witnessed, or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others" (Frazier et al. 2009, 450).[9] These kinds of events can have longstanding effects on students and may impact their classroom performance. The impacts may be seen on a student's physical well-being (body aches and pains or changes to sleep schedule), emotional signs (anxiety, depression, shame), and/or on their social well-being (CU Boulder Health and Wellness Services n.d.).

Trauma-informed pedagogy advocates for instructors to recognize the effects of trauma, respond with trauma-informed practices, and resist retraumatization. The framework is built on five main principles: safety, trustworthiness/transparency, empowerment/voice/choice, peer support, and collaboration (Thompson and Marsh 2022). These principles clearly overlap with one another and with those of feminist and inclusive pedagogy. This leads to a network of principles that mutually support one another and allows us to confidently employ these frameworks simultaneously in the present paper.

In addition to these general principles, we also recognize that many of our students, disproportionately the women, have had negative experiences in mathematics classrooms, leading them to develop "mathematics anxiety." Mathematics anxiety is defined as "feelings of apprehension and increased physiological reactivity when individuals deal with math, such as when they have to manipulate numbers, solve mathematical problems, or when they are exposed to an evaluative situation connected to math" (Luttenberger et al. 2018, 312). Students feel the effects of mathematics anxiety emotionally and cognitively: they suffer from feelings of nervousness when confronted with mathematics, and mathematics anxiety can compromise their working memory (312). In the context of mathematics exams, especially, Baloğlu and Koçak find that women experience higher levels of test anxiety than men (2006).

We propose to treat mathematics anxiety as a kind of trauma and use the framework of trauma-informed pedagogy to address it. Considering the similarities students may draw between mathematics and logic, we were worried that any mathematics anxiety and trauma students might feel would carry over to logic. Thus, the trauma-informed approach we take here is both general, in the sense that it recognizes that the majority of our students have dealt with hardships in the past and likely experience lingering effects, and specific, in the sense that it aims to recognize and ameliorate the trauma of mathematics anxiety.

Finally, it is worth noting that our study will not investigate the impacts of adopting these theoretical lenses on student learning outcomes. Rather, we would suggest the reader consult the literature cited above for evidence of the effectiveness of these lenses

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

for student learning. Here, we investigate whether the interventions designed based on these theoretical lenses lead to a change in students' perceptions of the inclusivity of logic. In particular, we investigate their perceptions of self-efficacy, of the objective nature of logic, and of the broader applicability of logic.

Methods

Data collection and instruments

Given the nature of our interventions, we used a quasi-experimental design common amongst education studies, a one-group pretest-posttest design. The dependent variable (i.e., student perceptions) is measured once before and once after the course (Jhangiani et al. 2019, §8; Garbacz and Kratochwill 2020). In this design, the pre-intervention students serve as a kind of control group for themselves. By comparing the impacts of our interventions on different demographic groups and looking for differential impact, we can still gather some information on the efficacy of our interventions.[10] In short, this design provides some evidence to identify the differential impact(s) our interventions had on various demographic groups. We begin by describing the surveys we used to measure student perceptions.

Two survey instruments were used to collect data across both universities. Both surveys collected data about students' demographics (regarding their gender, ethnicities, and racial identities), study majors, background with formal/technical subjects, perceptions of the transition from novice to expert in logic, self-efficacy with respect to the field of logic, and perceived value of collaboration in logic. The pre-course survey consisted of closed-response items only. The post-course survey additionally included open- and closed-response items asking about students' attendance and participation. Both surveys were distributed in the first and last three weeks of the respective quarter/semester.

Our survey questions were drawn from similar tools designed for mathematics courses, primarily because of the high perceived (and actual) similarities between logic and mathematics. In addition, these tools have already been tested and validated.

The first diagnostic we drew from is the "Mathematics Attitudes and Perceptions Survey" (MAPS) (Code et al. 2016). MAPS was designed to assess student perceptions of the discipline of mathematics and, specifically, the transition from novice to expert perceptions of the discipline. The second is May's "Mathematics Self-Efficacy and Anxiety Questionnaire" (MSEAQ) (May 2009). The MSEAQ explores general mathematics self-efficacy and grade anxiety, amongst other factors.

Since we are interested in investigating how feminist and trauma-informed interventions intersect with students' perceptions of formal philosophy, we began by selecting various MAPS/MSEAQ questions and adapting them for our purposes. In some cases, this was straightforward (e.g., replacing "mathematics" with "logic"). In other cases, the questions had to be adapted in more complex ways. For example, we adapted the following item from MAPS: "School mathematics has little to do with what I experience in the real world." Suppose we replace "School mathematics" with "logic" as a first step. Even then, on the one hand, it is possible (for example) for a person to acknowledge that logic might have everything to do with real-life experiences, but without being able to articulate why. On the other hand, sense-making, or understanding, regarding women's lived social realities and experiences, is a core underlying principle of feminist research (Kiguwa et al. 2019), and we were specifically interested in designing our survey questions with feminist/trauma-informed frameworks in mind. Therefore, we decided to reframe this item to reflect better our focus on

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

sense-making concerning lived experiences. Our adaptation reads: "Logic helps me understand my experiences in the world." Other survey items followed a similar process. In total, ten of the survey items included in both the pre- and post-surveys were adapted from MAPS/MSEAQ in this way. Furthermore, we included three additional survey items that we felt were not adequately reflected by existing MAPS/MSEAQ counterparts. Again, these questions were designed with our feminist and/or trauma-informed lens in mind. For example, "Logic is the kind of field in which collaboration is important" draws on the trauma-informed principle of collaboration.

The closed-response items on both the pre- and post-surveys are listed here. Unless otherwise noted, all items were followed by a five-point Likert scale (Strongly Agree to Strongly Disagree). We have also indicated specific questions informed by the MAPS or MSEAQ tools.

A. How do you describe yourself? (Female, Male, Non-binary/gender non-conforming, Transgender, Prefer to self-describe (with write-in), Prefer not to say)
B. Are you of Spanish, Hispanic, or Latino origin? (Yes, No)[11]
C. Choose one or more races that you consider yourself to be. (White or Caucasian, Black or African American, American Indian/Native American or Alaska Native, Asian, Native Hawaiian or Other Pacific Islander, Other (with write-in), Prefer not to say; multiple selections allowed)
D. What is your major? (write-in)

1. How would you describe your background with formal or technical subjects (e.g., mathematics, computer science, logic)? (Five-point Likert scale, Very Strong to Very Weak)
2. I have been told that there are multiple correct approaches to solving a logic problem. (MAPS)
3. I have a perception that logic is the kind of field in which there is usually only one correct approach to solving a logic problem. (MAPS)
4. My logic ability is something I cannot change very much. (MAPS)
5. I feel comfortable approaching new logic problems. (MSEAQ)
6. When learning something new in logic, I relate it to what I already know rather than just memorizing it the way it is presented. (MAPS)
7. I feel confident enough to ask questions in my logic class. (MSEAQ)
8. I believe I can think like a logician. (MSEAQ)
9. Logic helps me understand my experiences in the world. (MAPS)
10. I feel like I use the skills used in logic in my everyday life. (MAPS)
11. Logic is more similar to math than philosophy.
12. Logic is the kind of field in which collaboration is important.
13. I feel confident that I can develop/have the skills needed to do well in this class. (MSEAQ)

Additionally, we asked the following open-response questions on the second survey:

1. Have any of the activities over the course of the class raised any new ideas/questions/issues for you about who would be welcome to join a community of logicians? (write-in)
2. How often did you attend lectures? (I attended all/most lectures, I attended some lectures, I did not attend lecture regularly)

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

3. How often did you attend sections? (I attended all/most sections, I attended some sections, I did not attend sections regularly)
4. How often do you normally participate in other classes? (I usually participate regularly, I usually participate sometimes, I do not usually participate)
5. How often did you participate in this class? (I participated regularly, I participated sometimes, I rarely participated)

University contexts

University of California, Irvine (UCI)

One of us, Helen Meskhidze, female-identifying, was the lead instructor for an Introduction to Symbolic Logic course at UCI in spring 2023. The course is cross-listed across two schools in University of California, Irvine (UCI) and counts toward one of two General Education undergraduate requirements (for all UCI students): Science and Technology, and Quantitative, Symbolic, as well as Computational Reasoning. The course was taught during the quarter system (ten weeks), and lectures occurred approximately 30 times (three times a week for 50 minutes each). Students also attended nine discussion sections over the course of the ten weeks (once per week for 50 minutes, except week 1). Discussion sections were led by two male-identifying teaching assistants. The instructional team (the lead instructor and two teaching assistants) met at the start of the quarter to discuss the course's framing and goals. We also met weekly to discuss how the course was going. The two teaching assistants had complete control over their own discussion sections. The course had 126 students enrolled.

The course introduced students to semantic validity (via truth tables and models) and deductive validity (via the Fitch natural deduction) for sentential and predicate logic. The course goals included for students to be able to articulate the strengths and limitations of formal systems, as well as to reflect on and appreciate their development as logicians and as active, reflective learners. While there were no formal prerequisites, some students had taken a Critical Reasoning course before Introduction to Symbolic Logic and so had some familiarity with truth tables and the Fitch system.

University of Michigan, Ann Arbor (UM)

Another of us, Francisco Calderón (FC), male-identifying, was the graduate student instructor (GSI) for an Introduction to Symbolic Logic course at University of Michigan, Ann Arbor (UM) in fall 2023. The course is listed in the Philosophy Department and is mandatory for Philosophy and Cognitive Science majors in the "Philosophy and Cognition" track. It is also taken by non-majors looking to satisfy the Quantitative Reasoning Requirement of the College of Literature, Science, and the Arts. While the course is typically targeted at upper-division students, it has no formal prerequisites and is the introductory formal logic course offered by the Philosophy Department.

The course was taught during the semester system (15 weeks), and lectures met approximately 30 times (twice per week for 90 minutes each). This study was conducted with the consent of the lead instructor, Gordon Belot, who led lectures.[12] FC led two weekly discussion sections (one hour each). Responsibilities also included grading course assignments and running weekly office hours. The course had 46 students enrolled.

The course goals included for students to be able to work with symbolic formal systems, develop the ability to think abstractly about formal systems, think about the

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

relation between formal systems and ordinary discourse, understand how one formal system can be studied from various perspectives, and develop a sense of the scope and limits of formal logic. During discussion sections, FC articulated goals at a similarly broad level rather than in terms of the "contents" or tools of logic itself.

Study populations

The overall study population consisted of undergraduate students at UCI and UM enrolled in each university's respective version of the course. All students were eligible to participate in the study. There were no exclusion criteria. Students were told that the surveys were about their perceptions of logic and their abilities (the surveys were discussed in lectures at both universities). Students were not told that the results would be analyzed by gender and other social locations to avoid triggering stereotype threats. All students were also sent reminders to complete the surveys via the learning management system (Canvas for both institutions). All students were allowed to opt out of/not complete the surveys. Of the total 192 students enrolled in the respective version of the course across both universities, 139 participated in the pre-course survey, and 79 participated in the post-survey.

We note that data on the number of students who dropped the course after taking the pre-survey were not available. As a result, we acknowledge the possibility of selection effects (whereby students who remained in the course were less likely to have had the issues targeted by the study) across unpaired pre- and post-survey populations. However, anecdotally, the number of course drops was likely low (no more than 10 percent), and we suggest the likelihood of course drops artificially inflating the overall efficacy of the study is also low. We also carry out paired comparisons; these data are unaffected by course drops. Next, we offer a breakdown of the study population by university.

UCI

Of the 126 students enrolled in the course at UCI, 104 participated in the pre-course survey, and 52 participated in the post-course survey. Figures 1, 2, and 3 indicate overall pre- and post-survey participant demographics.

We make a few observations on these statistics. First, as of the 2023 academic year, by gender, the overall UCI undergraduate population is 53 percent women, 44 percent men, and 3 percent genderqueer or non-binary.[13] Given that it is well-documented that women are more likely to respond to surveys than men, the statistics in our survey seem reasonably representative of the students enrolled in the course. Although there is a drop in overall participation, respondent gender is reasonably consistent across the two surveys.

Second, while not indicated in Figure 2, of the 37 respondents in the pre-survey who identified as Hispanic/Latino, 27 (73 percent) were women, while ten (27 percent) were men. In the post-survey, 17 were women (85 percent), while three were men (15 percent). Thus, we saw a larger proportion of Hispanic/Latino women responding to the second survey than the first.

Third, the proportion of women versus men in each race category was relatively consistent, except for those who identified as Asian. There were 30 (56 percent) female and 23 (43 percent) male Asian respondents in the pre-survey versus 11 (44 percent) female and 13 (52 percent) male Asian respondents in the post-survey. We do not know

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

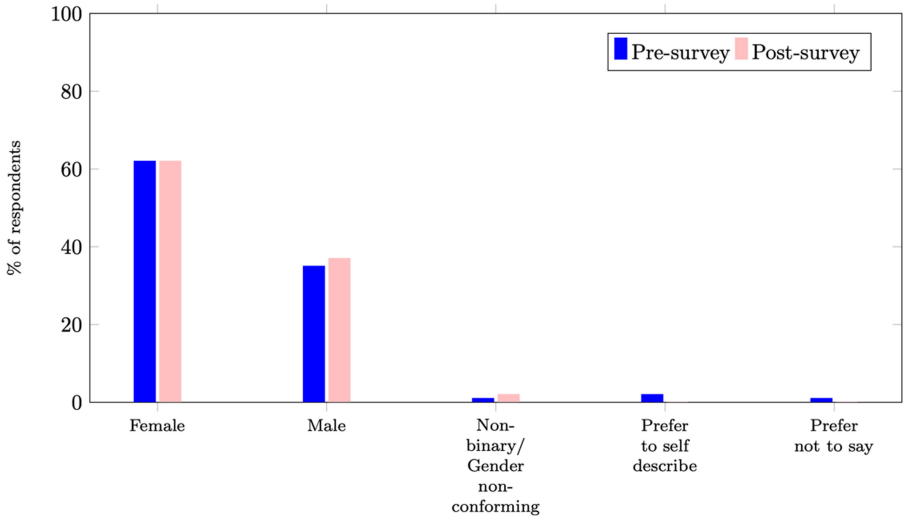


Figure 1. Pre- and post-survey gender demographics at UCI.

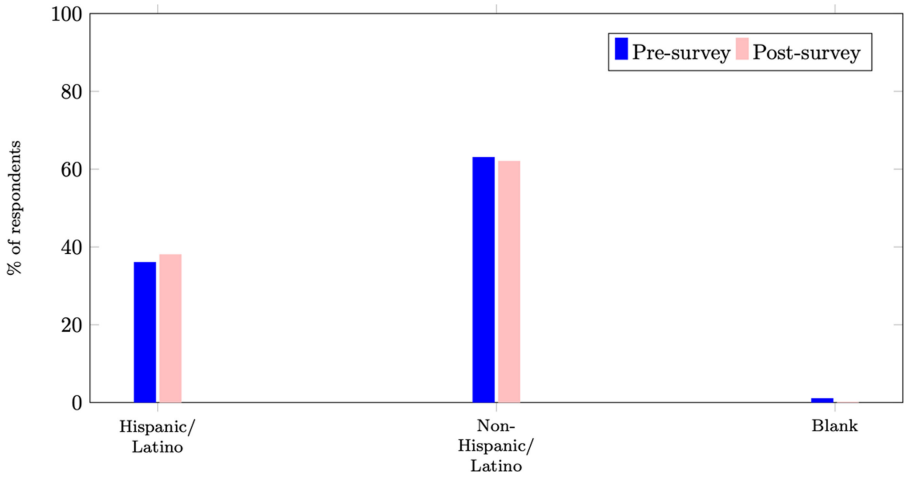


Figure 2. Pre- and post-survey ethnicity demographics at UCI.

whether this is a sampling bias or whether the student enrollment in the class changed as (Asian) students dropped the course. Interestingly, we did not see a similar effect—that the majority gender of the respondents in a racial group flipped—in any other racial demographic (although all other racial groups were much smaller).

UM

Of the 46 students enrolled in the course at UM, 35 participated in the pre-course survey, and 27 participated in the post-survey. Figures 4, 5, and 6 indicate overall pre- and post-survey participant demographics.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

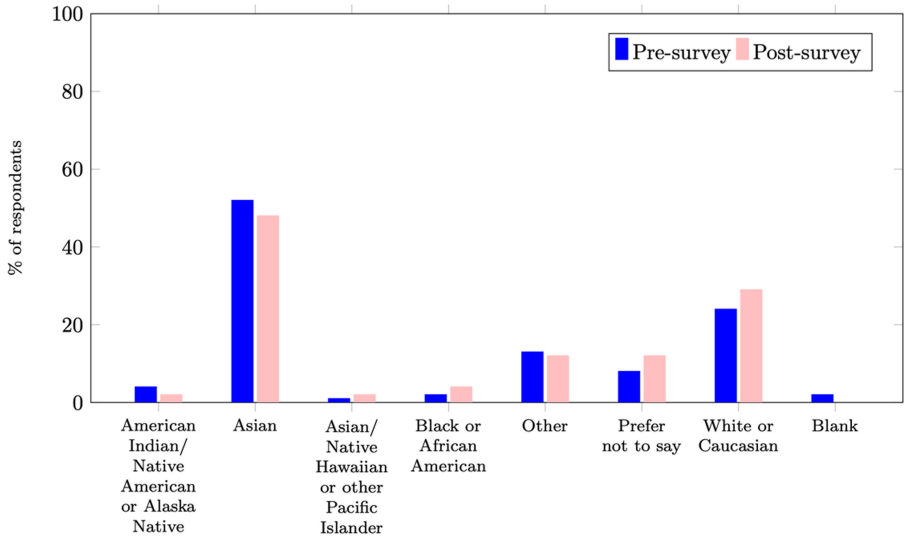


Figure 3. Pre- and post-survey racial demographics at UCI.

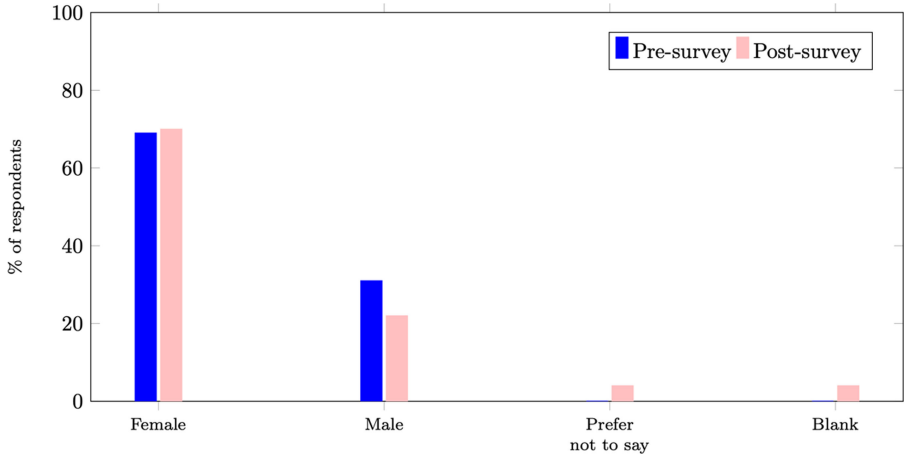


Figure 4. Pre- and post-survey gender demographics at UM.

Some observations about these statistics: first, in fall 2023, the overall gender distribution of undergraduate students at UM was 53 percent women and 47 percent men.[14] Since this course is mandatory for some Cognitive Science majors, it is also instructive to consider the demographics of that major. In the same semester, the enrollment of Cognitive Science majors was 66 percent women and 34 percent men, and 42 percent women and 58 percent men for Philosophy majors. Thus, the UM gender demographic data is relatively consistent with the demographics.

Second, the vast majority of respondents at UM identified as White.[15] This is somewhat less diverse than the demographics of the College of Literature, Science, and the Arts at UM. As we discuss in more detail below, this meant that most of the statistical

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

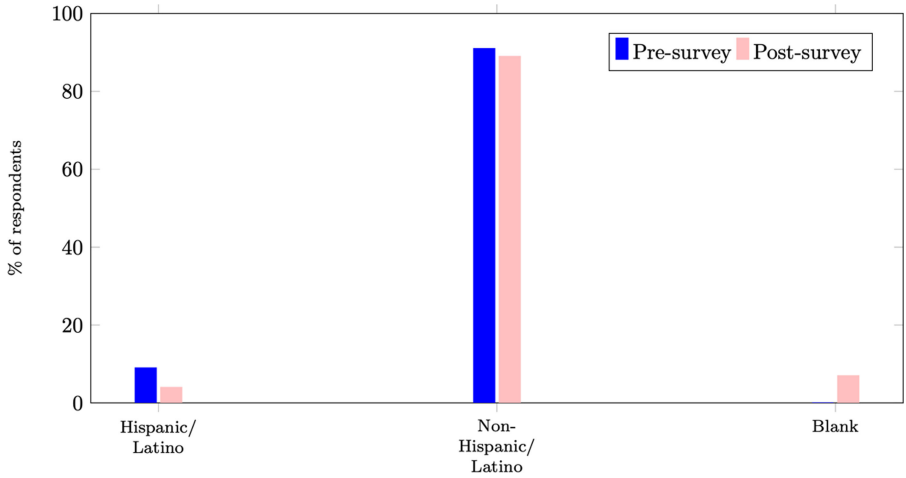


Figure 5. Pre- and post-survey ethnicity demographics at UM.

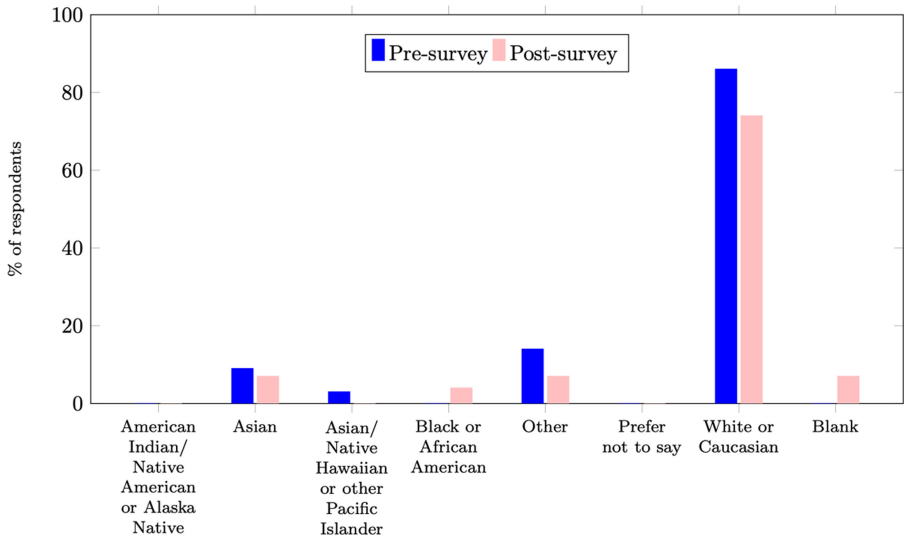


Figure 6. Pre- and post-survey racial demographics at UM.

weight of the intersectional analysis we were able to perform was derived from UCI. We also comment on the need for further work on similar intersectional studies below.

Ethical procedures

Survey responses were submitted anonymously to protect the confidentiality of all students. Strings of characters were assigned to responses across both surveys through the survey platform Qualtrics to ensure responses could be paired while retaining confidentiality. The study was determined exempt by UCI, IRB 2711, and UM, IRB HUM00235201.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Interventions

The interventions in our courses (at both universities) were informed by the theoretical lenses discussed in §2.2 and the literature on the sense of belonging in STEM. As a result, we sought to structure the course in ways that integrated feminist principles (i.e., empowerment, community, leadership, understanding regarding lived social realities, and pluralism) and trauma-informed principles (i.e., safety, trustworthiness, collaboration, choice, and empowerment) in responsive ways. Here, we present an overview and justification of our central course interventions and some of the central insights/principles from which they are derived. For more details on the interventions themselves, see the supplementary materials. Note that there is significant overlap in the (pedagogical) principles of each framework and in the harmful norms (e.g., political, epistemic) that each framework aims to destabilize (see, e.g., Kubala 2020). Thus, we elected not to organize this material according to the feminist and/or trauma-informed principles themselves. Moreover, we hope that the taxonomy below enables readers not just to track the differences between what is feasible for a lead instructor of a similar course versus a teaching assistant but also to make the interventions in the supplementary materials easier to identify.

Low-stakes, frequent approach to assessments

Mathematics trauma is felt severely when students (disproportionately women) are asked to perform on mathematics exams (Luttenberger et al. 2018). Furthermore, test anxiety is relatively common at the undergraduate level (Gerwing et al. 2015). We therefore expected that test anxiety would pose a serious risk to students in the course. To avoid triggering anxiety of this kind (and because it has been shown to lead to better student outcomes), we opted for only having low-stakes, frequent assessments at UCI. This approach aligns with the idea that the trauma-informed principle of safety is a precursor to a conducive learning environment (Carello and Butler 2015, 264). For example, students were assigned weekly practice problems graded upon completion. At both universities, the instructor used the software Carnap.io to administer practice problems. To further reduce anxiety, the lowest grades of weekly practice problems were dropped (the lowest 2 of 8 at UCI and the lowest 3 of 11 at UM). We note also that assignments graded upon completion were used in tandem with assignments graded on correctness. For instance, at UCI, at the end of each of the course's four modules, the instructor also assigned a problem set (graded on correctness). To balance inherent safety concerns arising from traditional grading formats, we aimed at both schools to ensure that feedback on problem sets graded on correctness was returned consistently and quickly. This approach aligns with the trauma-informed principle of trustworthiness, whereby clarity and transparency in policies and expectations can help build trust between students and instructors (Carpenter et al. 2021).

Encouraging agency in learning

Underrepresented groups feel a lower sense of belonging in STEM courses (Rainey et al. 2018). To address this in the context of logic, we aimed to facilitate a sense of belonging in the classroom by emphasizing the importance of students taking responsibility for their learning, which is also a core principle of feminist pedagogical approaches (Light et al. 2015).

First, reflective exercises formed a significant component of the course. For instance, at UCI, the instructor asked students to submit reflections on which problems they felt

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

most/least prepared for, how they might change their study habits, and what lingering questions they might have. At UM, FC held student "wrapper" sessions after tests, prompting students to reflect frequently on their perceived abilities. These reflective exercises draw on several trauma-informed principles in ways that intersect. For example, many traumatic experiences involve vulnerability, but resilience can be developed through one's support network (Stephens 2020, 7). In our context, student–instructor relationships form part of this support network. By frequently providing space for students to reflect on their mistakes, we hoped to normalize a discourse of vulnerability, foster trust between students and the instructor/the TA, and subsequently foster resilience. The latter approach also goes hand in hand with the principle of empowerment, which we aimed to enact by helping students discover and develop their own capacities.

Second, we allowed students to choose which course topics they wanted to study. For instance, at UCI, students could choose two of four final supplemental modules to cover in the course. To augment students' choice, the instructor designed these supplemental modules with practical applications (fuzzy logic, logic gates, formal fallacies, and modal logic) since they learned from previous instructors that students in Symbolic Logic regularly question the applicability of the course content to their everyday lives. At UM, students were allowed to choose (e.g.) five out of seven problems (or similar) on tests, allowing students to perform according to their strengths and comfort levels. This approach draws on feminist ideas relating to choice and empowerment, specifically, empowerment *through* agency, where agency is understood as the process of overcoming oppressive social conditions to pursue one's own flourishing (Khader 2011, 176).

Logic Labs

We also responded to the insight that underrepresented groups feel a lower sense of belonging in STEM courses by aiming to foster community within the course. Theorists argue that the feminist classroom is an equitable and holistic social environment, and in nurturing this kind of environment, a sense of community (e.g., in which everyone's presence and participation are valued) is a contributing factor (Shrewsbury 1987; Maher and Tetreault, 2001; hooks 1994). This approach also aligned nicely with the trauma-informed principle of collaboration. At UCI, the instructor turned Friday course meetings into "Logic Labs," designing course-related activities for students to complete in groups assigned at the start of the quarter. The activities ranged from an evaluation of ChatGPT's ability to solve logic problems to games designed with questions similar to those found in students' homework. While active learning was also incorporated into other lectures, Fridays were the most active component of the course. Additionally, the instructor at UCI made a course Discord, allowing students to communicate informally with one another and, occasionally, the instructors themselves. At UM, the vast majority of the discussion section activities were devoted to problem-solving in groups (thus, discussion sections played the role of Logic Labs at UCI). These activities were structured heavily enough that students knew what was expected of them, but flexibly enough that there were no expectations of doing everything or doing everything "perfectly." The goal was to defuse rivalry mindsets and create accountability through teamwork and a sense of belonging. For instance, aligning with a pluralistic approach, groups would often compare their work to emphasize that there was no single winning strategy in solving logic problems.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Building intentional space for underrepresented voices and traditions

Finally, we responded to the insight that underrepresented groups feel a lower sense of belonging in STEM courses by intentionally incorporating space in both courses for underrepresented voices and non-classical traditions in logic. Various feminist ideas drive this approach. Feminist logicians have recently argued for logical pluralism on feminist grounds (Saint-Croix and Cook 2024), while feminist epistemologists argue that dominant knowledge practices disadvantage women and other underrepresented groups by excluding them from inquiry and/or denying them epistemic authority and by producing knowledge that reinforces gender and other social hierarchies (see Anderson 2024 and Dotson and Sertler forthcoming, for surveys of the relevant literature). We, therefore, aimed to include spaces in our courses for highlighting and elevating the voices of women in logic, to deconstruct the idea that the male voice is the loudest voice, and to highlight non-classical approaches to logic, to emphasize that there is no one "right" way of doing logic. For example, at UCI, where possible, the instructor highlighted underrepresented voices and their contributions (e.g., Ruth Barcan Marcus's work in modal logic). Most significantly, the course's final weeks were devoted to reasons for studying formal logic and its limitations. At both universities, students completed a "Logic Reflection" evaluating the limitations of classical logic. This reflection was guided by Eugenia Cheng's *The art of logic in an illogical world*.[16] (More details can be found in the supplementary materials.) At UM, the GSI made time during discussion sections for conversations about the limitations of classical logic and the motivations for the non-classical logics (e.g., multi-valued relevance logic) mentioned in the lecture or problem sets. The goal was to show how a logical system may be "broken"/extended/modified and to deconstruct the idea of logic as fixed and/or rigid.

Data analysis

Several factors guided our decision-making with respect to data analysis. First, many of the survey questions are interrelated. Given this and our relatively small sample size, we were more interested in the broad trends shown by our study than in tracking student responses to particular questions. Second, our study is preliminary. We expected certain question groupings to emerge in the data, but did not want to specify any such groupings beforehand. Thus, to measure the extent to which our feminist and trauma-informed interventions changed students' perceptions of formal philosophy, we carried out an exploratory factor analysis (EFA). EFA is appropriate when there are general structural trends giving rise to the data. In the case of surveys like ours, EFA groups various questions together under "factors." Additionally, EFA, as opposed to, for example, confirmatory factor analysis, is appropriate when the researcher does not yet know which or how many factors to use. More formally, EFA is a prerequisite for examining construct-relevant multidimensionality (Morin et al. 2016). Construct-relevant multidimensionality here refers to the idea that multiple questions ("dimensions") are grouped to measure some particular construct. The factors generated by EFA consist of highly correlated variables; therefore, an advantage of this method is reducing the number of variables by combining two or more variables into a single common factor.

We ran EFA on questions 1 to 13 of the pre- and post-survey (see §3.1).[17] The EFA consists of multiple stages: normality testing, factor extraction, and reliability testing. Then, using the extracted factors, we carried out hypothesis testing to test for significant differences between various populations. Below, we present a detailed description with justifications for each step of our data analysis to highlight the many choices one makes

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

in such an analysis. We also hope that our discussion can be used to guide others who may be interested in conducting similar surveys at their institutions. To the especially keen reader, we offer further details in the Appendix.[18]

We carried out significance testing for each factor identified in the EFA as follows. In both pre- and post-survey, we examined: gender differences (male versus female) at both UCI and UM; differences in students belonging to one or more underserved ethnic/racial populations versus students belonging to no underserved populations at UCI,[19] and students identifying as Asian versus students identifying as White/Caucasian at UCI. For paired responses (i.e., responses we could trace as belonging to the same student pre- and post-intervention), we examined pre-post differences in the samples at the corresponding universities and overall pre-post differences at both universities. We note that, due to the sample size, we could not carry out analysis by underserved population or Asian versus White/Caucasian racial identification at UM.

We highlight that care must be taken when drawing conclusions from these test results, since we carried out two different kinds of tests (one kind comparing two independent samples in both the pre- and post-survey, and the other kind comparing matched samples across both surveys) with the same potential relationship in mind (the impact of our interventions on perceptions of logic). However, carrying out both kinds of tests offers a more nuanced analysis of our research questions. For example, a lack of significant differences in matched samples by gender might mask significant gender differences that exist before the course but disappear afterwards. But both kinds of test are necessary to ascertain where these patterns exist. By carrying out both kinds of tests and offering a careful interpretation of the results, we hope to pay attention to these nuances and provide a more detailed and informative analysis and discussion.

Results

Exploratory factor analysis

Data screening

The first step of exploratory factor analysis (EFA) is data screening. No data samples contained missing data, and no outliers were detected (univariate, i.e., with respect to a single variable, or multivariate). Tables 1 and 2 show sample *N* sizes for the overall pre-survey (PRE), post-survey (POST), and paired (PAIRED) data samples at each university, along with sample *N* sizes for demographic populations used for testing.

Normality testing

The second step of EFA is normality testing. These tests are designed to detect whether the null hypothesis came from a normally distributed population and will ultimately determine the type of significance test required when we analyze statistical significance in our results. We used the Shapiro-Wilk's test (Shapiro and Wilk 1965) to examine assumptions of univariate normality in PRE, POST, and PAIRED (including across demographic subgroups). Shapiro-Wilk rejected the null hypothesis ($p<0.05$) for a vast majority of survey items. This suggests a violation of univariate normality for the data samples. This violation is consistent with the diversity of the demographics of our survey respondents and the relevance of this diversity for the questions we are asking. Put differently, because we expect the demographics of students to matter for the different survey items, we did not expect the data to be well-described as coming from a single,

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Table 1. Sample *N* at UCI

	PRE	POST	PAIRED
Overall	104	52	42
Female	64	32	26
Male	36	19	16
USP	41	23	18
Non-USP	62	29	24
Asian	51	22	19
White/Caucasian	19	11	6

Table 2. Sample *N* at UM

	PRE	POST	PAIRED
Overall	35	27	19
Female	24	19	14
Male	11	6	5

normally distributed population. As a result of likely non-normality, Wilcoxon signed rank tests were deemed more suitable for significance testing (below).

We also tested whether the data are suitable for EFA using two tests: the Kaiser-Olkin-Meyer Measure of Sampling Adequacy and Bartlett's Test of Sphericity (Dziuban and Shirkey 1974). The tests confirm the suitability of these tests (more details of these analyses are given in Appendix A.1).

Factor extraction

The next step in the analysis was to estimate an appropriate number of factors to characterize the data. Such an estimate simplifies the fit statistics analysis by narrowing the number of possibilities we need to consider. Recall that these factors should be understood as groupings of questions for our purposes. Several methods were used to roughly estimate an appropriate number of factors: the eigenvalue method (Kaiser 1960), a scree plot, and a parallel analysis (see Appendix A.2 for more). The eigenvalue method and parallel analysis suggested that the number of factors was four, while the scree plot suggested a number of factors somewhere below this.

Using these results, we decided to consider the fit statistics and suitability of a two-, three-, and four-factor model (details and fit statistics can be found in Appendix A.2). Table 3 describes the significance of each of our survey questions to the different factors (i.e., the factor loadings). Factor loadings of approximately 0.4 and above are considered stable (Guadagnoli and Velicer 1988).

Based on the EFA, we opted for the three-factor model consisting of factors comprised of the following variables: Q2, Q3 (factor 1 = F1); Q1, Q5, Q7, Q8, and Q13 (factor 2 = F2); and Q6, Q9, and Q10 (factor 3 = F3). Q4 was omitted from further analysis since its factor loading was so low. Factor 1 consists of variables (i.e., survey

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Table 3. Factor loadings and correlations for two-, three-, and four-factor models (* = significant at 1% level)

Survey Question	2-factor model		3-factor model			4-factor model			
	F1	F2	F1	F2	F3	F1	F2	F3	F4
1	0.581*			0.591*		0.801*			
2			0.391*				1.003*		
3			1.003*				0.436*		
4	−0.341			−0.372					0.471
5	0.478*			0.481*		0.303		0.470	
6		0.434			0.408*			0.714*	
7	0.536*			0.546*		0.423			
8	0.595*			0.596*		0.665*			
9		0.747*			0.709*			0.716	0.508*
10		0.759*			0.851*			0.739	0.638*
11									
12									
13	0.658*			0.700*		0.646*			
Factor correlations	F1	F2	F1	F2	F3	F1	F2	F3	F4
F1	1.00		1.00			1.00			
F2	0.46*	1.00	−0.212	1.00		−0.265	1.00		
F3			−0.118	0.525*	1.00	0.574	−0.279	1.00	
F4						−0.012	0.110	−0.346	1.00

questions) related to perceptions about the objective nature of logic. Factor 2 consists of variables related to perceptions about aptitude/self-efficacy with respect to logic. Factor 3 consists of variables related to perceptions about the broader applicability of logic.

We note that although fit statistics for the three-factor model did not quite meet all ideal thresholds, this factor composition seemed more appropriate than the composition suggested by the two- and four-factor models. The two-factor model consists of factors comprised of the following variables: Q1, Q4, Q5, Q7, Q8, and Q13 (factor 1), and Q6, Q9, and Q10 (factor 2). This does not seem to accord well with variables related to perceptions about the objective nature of logic. However, the perception of technical disciplines (such as mathematics) as objective runs counter to feminist pedagogical principles, for example, in understanding what constitutes knowing and how that knowing is achieved through didactic situations. Thus, variables related to perceptions about the objective nature of logic seem appropriate for identification as a separate (but related) dimensional construct for further investigation in the broader context of this study. The four-factor model exhibited significant cross-loading (i.e., individual survey questions were counted toward multiple factors) and did not accord well overall with component themes. Overall, the three-factor model seemed to fit the data best.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Table 4. Cronbach's alpha for the three-factor model

	Factor		
	F1	F2	F3
Cronbach's alpha	0.60	0.72	0.72

Reliability

Internal reliability for the selected three-factor model was first evaluated using Cronbach's alpha (Table 4). For factors F2 and F3, the values lay above 0.70, indicating generally acceptable reliability. We note that the value for factor F1 lies on the lowest acceptable threshold (0.6) (Cronbach 1951), possibly due to the small number of variables in factor F1. As a result, internal reliability for factor F1 was subsequently also evaluated using mean inter-item correlations, with a correlation coefficient of 0.44. This suggests overall reasonable homogeneity while retaining sufficiently unique variance so as not to be redundant (Briggs and Cheek 1986).

Significance testing

Using the three-factor model and informed by the results of our normality tests, unpaired Wilcoxon signed rank tests (McKnight and Najab 2010) on sum scores were carried out to analyze differences in responses by factor for our various samples on pre-survey responses and post-survey responses. In both the pre- and post-survey, we examined: gender differences (male versus female) at both UCI and UM; differences in students belonging to one or more underserved ethnic/racial populations versus students belonging to no underserved populations at UCI, and students identifying as Asian versus students identifying as White/Caucasian at UCI. Sample size dictated that we could not carry out analysis by underserved population or Asian versus White/Caucasian racial identification at UM. Furthermore, paired Wilcoxon signed rank tests were carried out on sum scores to analyze pre-post differences in demographic samples (at the corresponding universities) and overall pre-post differences at both universities. We report the z-scores and p values of the Wilcoxon signed rank tests below.[20]

UCI

At UCI, perceptions of aptitude/self-efficacy with respect to logic before the course differed significantly by gender, and Asian versus White/Caucasian racial identification, in PRE and POST data, respectively (see Table 5). No other differences were significant across PRE or POST data. Female median response was significantly lower than male median response in PRE (so female perceptions of self-efficacy were significantly lower pre-survey). Asian-identifying median response was significantly lower than White/Caucasian-identifying median response in POST (so Asian-identifying perceptions of self-efficacy were significantly lower post-survey).

At UCI, perceptions of objectivity of logic differed significantly from pre- to post-survey among: the overall population, male students, students not belonging to an underserved population, and Asian-identifying students (see Table 6). No other differences were significant across PAIRED data. Across each of those four populations, the pre-survey median response was significantly higher than the post-survey median response (so perceptions about the objectivity of logic decreased post-survey).

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Table 5. Wilcoxon *p* values and z-scores for differences in PRE and POST at UCI (* = significant at 5% level, ** = significant at 1% level)

		Factor F1 (Objectivity of logic)		Factor F2 (Self-efficacy)		Factor F3 (Applicability of logic)	
		p	z	p	z	p	z
PRE	Gender	0.127	−1.53	0.010**	−2.59	0.132	−1.51
	USP	0.668	−0.43	0.333	0.97	0.690	−0.40
	Racial identification	0.635	0.48	0.566	−0.58	0.354	0.93
POST	Gender	0.064	1.87	0.171	−1.38	0.937	0.09
	USP	0.483	0.71	0.956	−0.06	0.682	0.42
	Racial identification	0.535	0.64	0.033*	−2.15	0.602	−0.54

Table 6. Wilcoxon *p* values and z-scores for differences across PAIRED at UCI (* = significant at 5% level, ** = significant at 1% level, *** = significant at 0.1% level)

	Factor F1 (Objectivity of logic)		Factor F2 (Self-efficacy)		Factor F3 (Applicability of logic)	
	p	z	p	z	p	Z
Overall	<0.001***	3.45	0.268	−1.12	0.085	1.73
Female	0.133	1.53	0.272	−1.11	0.270	1.12
Male	0.002**	3.08	0.635	−0.51	0.112	1.63
USP	0.150	1.47	0.274	−1.12	0.428	0.83
Non-USP	0.002**	3.19	0.463	−0.75	0.112	1.61
Asian-identifying	0.014*	2.49	0.886	−0.17	0.086	1.75
White/Caucasian-identifying	0.170	1.51	0.073	−1.90	0.892	0.27

UM

At UM, perceptions of aptitude/self-efficacy with respect to logic before the course also differed significantly by gender, and no other differences were significant across PRE or POST data (see Table 7). Female median response was significantly lower than male median response in PRE (so female perceptions of self-efficacy were significantly lower pre-survey).

At UM, perceptions of objectivity of logic and perceptions of aptitude/self-efficacy differed significantly from pre- to post-survey among the overall population and female students (see Table 8). No other differences were significant across PAIRED data. For both populations, the pre-survey median response for factor 1 was significantly higher than the post-survey median response (so perceptions about the objectivity of logic decreased post-survey). Also for both populations, the pre-survey median response for factor 2 was significantly lower than the post-survey median response (so perceptions of aptitude/self-efficacy increased post-survey).

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Table 7. Wilcoxon *p* values and z-scores for differences in PRE and POST at UM (** = significant at 1% level)

		Factor F1 (Objectivity of logic)		Factor F2 (Self-efficacy)		Factor F3 (Applicability of logic)	
		p	z	p	z	p	z
PRE	Gender	0.771	0.31	0.009**	−2.65	0.914	0.13
POST	Gender	0.741	0.36	0.317	−1.03	0.897	0.16

Table 8. Wilcoxon *p* values and z-scores for differences across PAIRED at UM (* = significant at 5% level, ** = significant at 1% level)

	Factor F1 (Objectivity of logic)		Factor F2 (Self-efficacy)		Factor F3 (Applicability of logic)	
	p	z	p	z	p	z
Overall	0.014*	2.48	0.004**	−2.91	0.076	−1.81
Female	0.029*	2.23	0.006**	−2.79	0.237	−1.24
Male	0.462	0.92	0.773	−0.58	0.198	−1.47

Discussion

There are several takeaways from the above discussion. First and most significant is the only trend we saw across both universities: Female students entered the course with lower perceptions of their aptitude/self-efficacy than male students. In other words, at the start of the course, women felt less comfortable approaching new logic problems, less confident asking questions, and less sure they could develop the skills necessary to succeed in the course. This aligns with our expectations from the literature on mathematics anxiety and trauma. By the end of the course, however, we had closed the gap between men's and women's perceptions of self-efficacy in both university contexts. In other words, we have some evidence that *our interventions successfully ameliorated gender-based differences in perceptions of self-efficacy*. This conclusion is carefully stated; notice that we cannot conclude (e.g.) from our matched samples that female perceptions of self-efficacy significantly increased during our course (they did at UM, but not at UCI). However, we can offer a little more. Inspecting Tables 5 and 6, female perceptions of self-efficacy increased during our course, *but increased more than male perceptions of self-efficacy increased*. Thus, we offer a more nuanced picture of how our interventions successfully ameliorated gender-based differences in perceptions of self-efficacy. Our interventions are particularly effective for female students with respect to perceived self-efficacy.

When survey responses were paired, at UCI, we saw a change in student perceptions of the objectivity of logic.[21] For men, students not belonging to an underserved population, and Asian-identifying students, their perceptions of the objectivity of logic decreased. We hypothesize that women and those belonging to underserved populations entered the course with lower perceptions of the objectivity of logic. Thus, the course did not have a statistically significant impact on their perception. However, this hypothesis requires further investigation, especially considering the results at UM. There, we saw women demonstrate a change in their perception of the objectivity of logic, but not men.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

(The student population at UM did not have sufficient demographic diversity to comment on this trend for Asian-identifying students or those belonging to an underserved population.) It is unclear whether this is due to the differences in study populations, the different course interventions, or something else.

From the UCI data, we also saw a differential impact of our interventions on Asian students' perceptions of self-efficacy. Asian students' perceptions of self-efficacy were not significantly different from those of White/Caucasian-identifying students at the start of the course, but by the end, they were significantly lower. However, this is not to say that Asian-identifying students' perceptions of self-efficacy decreased. Rather, by inspecting Table 6, both Asian-identifying and White/Caucasian-identifying students' perceptions of self-efficacy increased during our course. In particular, White/Caucasian-identifying students' perceptions of self-efficacy increased more than Asian-identifying students' perceptions, and this seems to underlie the observed post-survey significant difference between these two groups. These results highlight the importance of being attuned to the ways in which social identity and context intersect and interact. We are not confident that our interventions contributed substantially to the increase in Asian-identifying students' perceptions of self-efficacy. Rather, in the context of logic, we suggest our study shows that Asian students' experiences do not track those of other underserved populations. Finally, it is worth noting that we did not find any changes in students' perceptions of the broader applicability of logic at either university. We hypothesize that one course is insufficient to change their attitudes on this topic.

Finally, we again note a significant consequence of our study design: while we are confident in our claims about the differential impact of our interventions, we are less certain about the absolute impact. For example, by comparing how different groups respond to our interventions, we can confidently say that our interventions had a larger impact on women than men. However, we cannot distinguish between the impact of simply taking a logic course and the impact of our interventions on women's versus men's perceptions. Nonetheless, since the frameworks of feminist and trauma-informed pedagogy each have independent evidence of their efficacy, we firmly believe in the beneficial impact of our interventions.

We hope to have outlined the benefits of the types of interventions presented above. We also hope to empower readers to survey their own courses. These two goals are independent and independently justified. What we show in our results, though, is that they are thankfully overlapping (i.e., the interventions we propose *do* make the course more inclusive according to our surveys, especially for women). Nonetheless, for those who do not want to overhaul their courses, we argue that you should still consider surveying your students. For those who do not want to survey your students, we still urge you to consider some modifications to your courses to make them more inclusive along the lines discussed above.

Future work

An advantage of our study is that it highlights how interventions can be made at different levels. We recognize that some interventions can be especially difficult to implement as a teaching assistant. This is not only because some require more control over the course structure than others, but also because different lead instructors and institutions will have different norms and expectations about how much ownership a teaching assistant has over course activities. Still, we hope some teaching assistants are encouraged to try whatever interventions seem available to them.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Even in contexts where the lead instructor supports a teaching assistant drawing from inclusive or other non-traditional pedagogies, there are limitations to the effectiveness of the interventions that a teaching assistant can carry out. Indeed, one limitation of this study, particularly in the UM context, was that the interventions run in the discussion sections were part of the activities used to assign participation grades (10 percent of students' overall grades). On the one hand, having a participation grade correlated with high attendance rates. On the other hand, a grade could have become an incentive for continuous, active, and enthusiastic participation in section activities, including our interventions. We cannot estimate the degree to which this affects our results. We hope that future studies can shed more light on the effectiveness of such interventions at various levels.

Another advantage of our study is the differences in the study populations. The demographics of students at the two universities are quite different. This is a strength of our study since it means that our interventions were carried out on a greater diversity of students. However, it also means that the data are difficult to compare and generalize. Thus, we hope future work will conduct such interventions in other university contexts with different student demographics. Additionally, both universities where we conducted the study are classified as R1. We hope future work will also consider other types of universities and include universities outside the US context. We also hope a deeper intersectional analysis can be conducted with more data and more students to survey. In particular, our study suggests that female Asian students would be an interesting intersectional identity to investigate.

Though the study populations at the universities differed, the course content and level were the same. We hope that future work can demonstrate the effectiveness of the kinds of interventions described in this paper in different levels of logic courses (i.e., in higher-level or even graduate courses). Such studies would also be able to use the theoretical lenses outlined here, but investigate different kinds of interventions. We would be especially interested in the results of studies that intervene more directly in the core course content.

Finally, though we focused on logic courses, our initial motivation was to analyze student perceptions of formal philosophy more broadly. Thus, we hope future work can generalize beyond logic, intervening in and surveying other kinds of "formal philosophies." Many of the interventions described here can be readily adopted into philosophy of physics, philosophy of biology, etc., course contexts.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/hyp.2025.10025

Acknowledgements. Thanks to the audiences at the 2024 LogIn Project Workshop, the First Feminist Philosophy of Physics Workshop, and the 2024 American Association of Philosophy Teachers Workshop-Conference for helpful conversations, comments, and encouragement. Thanks to Gordon Belot for his support in the UM branch of our study, Jeffrey DeVries for statistical guidance, Melissa Jacquart for feedback on earlier stages of the project and especially for pointing us toward useful literature from the STEM disciplines on similar issues, and Nicole Winter for helpful recommendations on study design. FC would also like to thank Maegan Fairchild, Becca Pickus, Charlotte Probst, and Margot Witte for helpful conversations at the early stages of the project.

Notes

1 Jacquart et al. (2019) still think diversity is important and necessary, but just not sufficient. We think it is also worth asking whether recruitment efforts can be made in earnest before making our (sub-)discipline more just towards the women already here.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

2 We see this discussion as especially useful for readers unfamiliar with the traditional (and still widespread) approach and hope it can provide some context for the interventions we have designed that are discussed later in the paper (§3.5). For readers familiar with the traditional approach used when teaching logic, we suggest skipping to §2.2.

3 It is important to keep in mind that we are not claiming that the interventions we describe are more effective for student learning than the traditional methods. Indeed, a different kind of study than that described here—perhaps one that tracks progress toward specific learning outcomes or one that tracks retention rates—would be required to provide evidence for such a claim. Such studies have, of course, been conducted, and independent research supports the idea that our interventions are effective for student learning. For instance, MacPhee et al. (2013) show that interventions for students from underrepresented groups with low perceptions of self-efficacy can improve performance in the context of STEM. Here, we aim to show that our interventions are quite novel to logic pedagogy (and, we suspect, formal philosophy pedagogy).

4 We did not include courses like "Critical Reasoning" as comparisons because we wanted the content to be closer to the content we covered in the courses from our study.

5 The sample has many syllabi from R1 universities. We are not sure whether this is representative of where such courses are taught or is a consequence of other factors (e.g., instructors at R1 universities being more likely to make syllabi public). For our purposes, the selection biases are not so important, especially because the two courses in which we conducted our study were taught at R1 universities.

6 The limitations of using syllabi to gauge course content are especially pronounced: instructors may be covering this material in lecture but not explicitly including it in the syllabus. Some may, for instance, discuss alternative approaches for formalizing the material conditional in their initial presentation of the truth table or inference rules (thanks to an anonymous reviewer for this example). However, we suggest that these kinds of skills (understanding the applications and limitations of classical logic and reasons for turning to alternative systems) should be considered part of the main course objectives in teaching symbolic logic. Thus, we would argue that they *should* be included in the syllabus.

7 We thank our anonymous reviewers for pressing for more clarity on our understanding of these issues.

8 Saint-Croix and Cook (2024) note that this argument for pluralism, based on Longino's contextual values, is novel to the logical pluralism literature. The extant literature has defended arguments for pluralism based on the grounds of different subject matters requiring different logics and different logical consequence relations. They see the fact that several arguments converge to logical pluralism as further evidence of pluralism itself.

9 Some examples of traumatic events include unexpected deaths, witnessing family violence, stalking, partner violence, natural disasters, etc.

10 The method we use here—one-group pretest-posttest design—is commonly used amongst educators who cannot have a strict control group because they do not wish to subject a subset of their students to worse instruction. It is even more appropriate for us because worse instruction, in this case, risks retraumatizing an already vulnerable student population.

11 Participants were asked to disclose Hispanic/Latino heritage separately from their racial background in acknowledgment of the fact that the (ethnic) category of Hispanic/Latino is not itself a race.

12 Some of the interventions below were course policies independently designed by the lead instructor. We have included those aligned with our principles since we would have aimed to implement similar strategies anyway.

13 These statistics are from https://www.universityofcalifornia.edu/about-us/information-center/fall-enrollment-glance.

14 These statistics are from https://ro.umich.edu/reports/ethnicity. The report does not include other genders.

15 Note that while we allowed selecting multiple races in our survey, the analysis here only considers one race at a time. In other words, we do not conduct specific analyses for multi-racial identities.

16 Students watched Cheng discuss her book through a portion of https://www.youtube.com/watch?v=yhzkx0h6cue&t=215s.

17 We note that questions 1 to 5, which were included only in the post-survey (see §3.1), are omitted from the analysis: student response rate was very low for the write-in question, and the remaining questions are tangential to our purposes for this paper.

18 All analyses were performed using R statistical software (R Core Team 2021, v4.4.1). EFA was carried out using the lavaan R package (Rosseel 2012, v0.6.18).

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

19 Aligning with the findings in (Schwitzgebel and Jennings 2017), an "underserved population" was defined as one of the following: Hispanic (any race), non-Hispanic Native American, Alaska Native, or Black.

20 Z-scores indicate the extent to which groups are evenly distributed, while *p* values describe the probability of getting the statistics given the null hypothesis (here, that the groups are evenly distributed). If the groups are evenly distributed, the z-score will be closer to 0. The greater the absolute value of the z-score, the higher the statistical significance and the higher the *p* value (e.g., a *p* value of 0.05 is approximately equal to a z-score of 2.5).

21 As discussed earlier, a high measure of objectivity would indicate that students see logic as very rigid and believe there to be only one correct approach to solving logic problems; our interventions were aimed to reduce students' perceptions of the objectivity of logic.

References

Anderson, Elizabeth. 2024. Feminist epistemology and philosophy of science. In *The Stanford Encyclopedia of Philosophy*, Fall 2024, ed. Edward N. Zalta and Uri Nodelman. Palo Alto, CA: Metaphysics Research Lab, Stanford University.

Baloğlu, Mustafa, and Recep Koçak. 2006. A multivariate investigation of the differences in mathematics anxiety. *Personality and Individual Differences* 40 (7): 1325–35. https://doi.org/10.1016/j.paid.2005.10.009

Baron, Sam, Tom Dougherty, and Kristie Miller. 2015. Why is there female under-representation among philosophy majors? Evidence of a pre-university effect. *Ergo* 2 (14): 330–65.

Beebee, Helen, and Jenniffer Saul. 2011. *Women in philosophy in the UK: A report by the British Philosophical Association and the Society for Women in Philosophy in the UK*. Joint BPA/SWIP Committee for Women in Philosophy (September, 2011). https://bpa.ac.uk/wp-content/uploads/2018/11/BPA_Report_Women_In_Philosophy.pdf (accessed July 28, 2025).

Bentler, Peter M., and Ke-Hai Yuan. 1999. Structural equation modeling with small samples: Test of global self-esteem. *Journal of Personality and Social Psychology* 89: 623–42.

Blake-Turner, Christopher, and Gillian Russell. 2021. Logical pluralism without the normativity. *Synthese* 198 (20): 4859–77.

Briggs, Stephen R., and Jonathan M. Cheek. 1986. The role of factor analysis in the development and evaluation of personality scales. *Journal of Personality* 54 (1): 106–48.

Brown, Timothy A. 2015. *Confirmatory factor analysis for applied research*. New York: Guilford Publications.

Carello, Janice, and Lisa D. Butler. 2015. Practicing what we teach: Trauma-informed educational practice. *Journal of Teaching in Social Work* 35 (3): 262–78. https://doi.org/10.1080/08841233.2015.1030059.

Carello, Janice, and Phyllis Thompson, eds. 2021. *Lessons from the pandemic: Trauma-informed approaches to college, crisis, change*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-83849-2

Carpenter, Russell, Shirley O'Brien, Travis Martin, Heather Fox, Clint Pinion, Susan Skees Hermes, Camille Skubik-Peplaski, and Casey Humphrey. 2021. Faculty development for transparent learning and teaching: Perspectives from teacher-scholars. *Journal of Faculty Development* 35 (2): 58–64.

Code, Warren, Sandra Merchant, Wes Maciejewski, Matthew Thomas, and Joseph Lo. 2016. The Mathematics Attitudes and Perceptions Survey: an instrument to assess expert-like views and dispositions among undergraduate mathematics students. *International Journal of Mathematical Education in Science and Technology* 47 (6): 917–37.

Conklin, Sherri Lynn, Michael Nekrasov, and Jevin West. 2023. Where are the women: The ethnic representation of women authors in philosophy journals by regional affiliation and specialization. *European Journal of Analytic Philosophy* 19 (1): S4–48.

Crawford, K., S. Gordon, J. Nicholas, and M. Prosser. 1994. Conceptions of mathematics and how it is learned: The perspectives of students entering university. *Learning and Instruction* 4 (4), 331–45. https://doi.org/10.1016/0959-4752(94)90005-1

Cronbach, Lee J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16 (3): 297–334. https://doi.org/10.1007/bf02310555

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

CU Boulder Health and Wellness Services. N.d. Five things everyone should know about trauma. https://www.colorado.edu/health/blog/trauma (accessed August 30, 2024).

Dotson, Kristie, and Ezgi Sertler. Forthcoming. Political epistemologies. In *A companion to epistemology*, ed. Kurt Sylvan, Jonathan Dancy, Ernest Sosa, and Matthias Steup. 3rd ed. Maldon, MA: Wiley-Blackwell.

Dziuban, Charles D., and Edwin C. Shirkey. 1974. When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin* 81 (6): 358–61. https://doi.org/10.1037/h0036316

Frazier, Patricia, Samantha Anders, Sulani Perera, Patricia Tomich, Howard Tennen, Crystal Park, and Ty Tashiro. 2009. Traumatic events among undergraduate students: Prevalence and associated symptoms. *Journal of Counseling Psychology* 56 (3): 450.

Freire, Paulo. 2018. *Pedagogy of the oppressed*. 50th anniversary ed., trans. Myra Bergman Ramos. New York: Bloomsbury Academic.

Garbacz, S. Andrew, and Thomas R. Kratochwill. 2020. Single-case experimental designs. Oxford Bibliographies Online in Psychology. https://www.oxfordbibliographies.com/view/document/obo-9780199828340/obo-9780199828340-0265.xml

Gerwing, Travis G., Joshua A. Rash, Alyssa M. Allen Gerwing, Bev Bramble, and Jeff Landine. 2015. Perceptions and incidence of test anxiety. *Canadian Journal for the Scholarship of Teaching and Learning* 6 (3). https://doi.org/10.5206/cjsotl-rcacea.2015.3.3

Guadagnoli, Edward, and Wayne F. Velicer. 1988. Relation of sample size to the stability of component patterns. *Psychological Bulletin* 103 (2): 265–75. https://doi.org/10.1037//0033-2909.103.2.265

Herfeld, Catherine, Jan Müller, and Kathrin Von Allmen. 2022. Why do women philosophy students drop out of philosophy? Some evidence from the classroom at the bachelor's level. *Ergo: An Open Access Journal of Philosophy* 8 (51): 741–86. https://doi.org/10.3998/ergo.2252

Hjortland, Ole Thomassen. 2017. Anti-exceptionalism about logic. *Philosophical Studies* 174 (3): 631–58.

hooks, bell. 1994. *Teaching to transgress: Education as the practice of freedom*. New York: Routledge.

Jacquart, Melissa, Rebecca Scott, Kevin Hermberg, and Stephen Bloch-Schulman. 2019. Diversity is not enough: The importance of inclusive pedagogy. *Teaching Philosophy* 42 (2): 107–39.

Jhangiani, Rajiv S., I-Chant A. Chiang, Carrie Cuttler, and Dana C. Leighton. 2019. *Research methods in psychology*. Surrey, BC: Kwantlen Polytechnic University.

Kaiser, Henry F. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20 (1): 141–51. https://doi.org/10.1177/001316446002000116

Khader, Serene J. 2011. *Adaptive preferences and women's empowerment*. Oxford: Oxford University Press.

Kiguwa, Peace. 2019. Feminist approaches: An exploration of women's gendered experiences. In *Transforming Research Methods in the Social Sciences: Case Studies from South Africa* (pp. 220–35), eds, Sumaya Laher, Angelo Frynn, and Sherianne Kramer. Wits University Press. http://www.jstor.org/stable/10.18772/22019032750.19

Kings, Amy E. 2019. Philosophy's diversity problem: Understanding the underrepresentation of women and minorities in philosophy. *Metaphilosophy* 50 (3): 212–30. https://doi.org/10.1111/meta.12358

Kline, Rex B. 2023. *Principles and practice of structural equation modeling*. New York: Guilford Publications.

Kubala, Julie. 2020. Of trauma and triggers: Pedagogy and affective circulations in feminist classrooms. *Feminist Formations* 32 (2): 183–206. https://doi.org/10.1353/ff.2020.0030

Light, Tracy Penny, Jane Nicholas, and Renée Bondy. 2015. *Feminist pedagogy in higher education: Critical theory and practice*. Waterloo, ON: Wilfrid Laurier University Press.

Lintott, Sheila, and Lissa Skitolsky. 2016. Inclusive pedagogy: Beyond simple content. *Hypatia* 31 (2): 447–59. https://doi.org/10.1111/hypa.12247

Lockard, Claire A., Helen Meskhidze, Sean Wilson, Nim Batchelor, Stephen Bloch-Schulman, and Ann J. Cahill. 2017. Using focus groups to explore the underrepresentation of female-identified undergraduate students in philosophy. *Feminist Philosophy Quarterly* 3 (4): 1–29.

Longino, Helen E. 1990. *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton: Princeton University Press.

Luttenberger, Silke, Sigrid Wimmer, and Manuela Paechter. 2018. Spotlight on math anxiety. *Psychology Research and Behavior Management* 11: 311–22.

MacPhee, David, Samantha Farro, and Silvia Sara Canetto. 2013. Academic self-efficacy and performance of underrepresented STEM majors: Gender, ethnic, and social class patterns. *Analyses of Social Issues and Public Policy* 13 (1): 347–69. https://doi.org/ 10.1111/asap.12033

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Maher, Frances A., and Mary Kay Thompson Tetreault. 2001. *The feminist classroom: Dynamics of gender, race, and privilege*. Lanham, MD: Rowman & Littlefield.

May, Diana Kathleen. 2009. Mathematics self-efficacy and anxiety questionnaire. PhD diss., University of Georgia, Athens, GA, USA.

McKnight, Patrick E., and Julius Najab. 2010. Mann-Whitney U Test. In *The Corsini encyclopedia of psychology*, eds. Irving B. Wiener and W. Edward Craighead. John Wiley & Sons, Inc. https://doi.org/10.1002/9780470479216.corpsy0524

Morin, Alexandre J. S., A. Katrin Arens, and Herbert W. Marsh. 2016. A bifactor exploratory structural equation modeling framework for the identification of distinct sources of construct-relevant psychometric multidimensionality. *Structural Equation Modeling* 23: 116–39. https://doi.org/10.1080/10705511.2014.961800

Paxton, Molly. 2015. Potential sources of gender disparities in philosophy: An examination of philosophical subfields and their relation to gender. *American Philosophical Association*. https://cdn.ymaws.com/www.apaonline.org/resource/group/bf785b0d-eb59-41f8-9436-1c9c26f50f8e/Paxton,_Gender_Disparities_i.pdf

Paxton, Molly, Carrie Figdor, and Valerie Tiberius. 2012. Quantifying the gender gap: An empirical study of the underrepresentation of women in philosophy." *Hypatia* 27 (4): 949–57. https://doi.org/10.1111/j.1527-2001.2012.01306.x

Rainey, Katherine, Melissa Dancy, Roslyn Mickelson, Elizabeth Stearns, and Stephanie Moller. 2018. Race and gender differences in how sense of belonging influences decisions to major in STEM. *International Journal of STEM Education* 5: 1–14.

R Core Team. 2021. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. https://www.R-project.org/

Rosseel, Yves. 2012. Lavaan: An R package for structural equation modeling. *Journal of statistical software* 48 (2): 1–36. https://doi.org/10.18637/jss.v048.i02

Ruetsche, Laura. 2020. What is it like to be a woman in philosophy of physics? In *The Routledge handbook of feminist philosophy of science* (pp. 397–408), eds. Sharon Crasnow and Kristen Intemann. Routledge.

Saint-Croix, Catharine, and Roy T. Cook. 2024. (What) is feminist logic? (What) do we want it to be? *History and Philosophy of Logic* 45 (1): 20–45.

Schwitzgebel, Eric, and Carolyn Dicey Jennings. 2017. Women in philosophy: Quantitative analyses of specialization, prevalence, visibility, and generational change. *Public Affairs Quarterly* 31: 83–105.

Schwitzgebel, Eric, Liam Kofi Bright, Carolyn Dicey Jennings, Morgan Thompson, and Eric Winsberg. 2021. The diversity of philosophy students and faculty. *The Philosophers' Magazine* 93: 71–90.

Shapiro, Samuel Sanford, and Martin B. Wilk. 1965. An analysis of variance test for normality (complete samples). *Biometrika* 52 (3/4): 591–611.

Shrewsbury, Carolyn M. 1987. What is feminist pedagogy? *Women's Studies Quarterly* 15 (3/4): 6–14.

Stephens, Darryl W. 2020. Trauma-informed pedagogy for the religious and theological higher education classroom. *Religions* 11 (9): 449. https://doi.org/10.3390/rel11090449

Thompson, Morgan. 2021. Underrepresentation in philosophy of science and what we can do about it. Newsletter of the Philosophy of Science Association's Women's Caucus, *Science Visions* 3 (1).

Thompson, Morgan, Toni Adleberg, Sam Sims, and Eddy Nahmias. 2016. Why do women leave philosophy? Surveying students at the introductory level. *Philosophers' Imprint* 16 (6): 1–36.

Thompson, Phyllis, and Heidi Marsh. 2022. Centering equity: Trauma-informed principles and feminist practice. In *Trauma-informed pedagogies: A guide for responding to crisis and inequality in higher education*, ed. Phyllis Thompson and Janice Carello. Cham: Springer.

Weiler, Kathleen. 1991. Freire and a feminist pedagogy of difference. *Harvard Educational Review* 61 (4): 449–75.

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Appendix: Data analysis

A.1 Suitability of EFA

The Kaiser-Olkin-Meyer Measure of Sampling Adequacy and Bartlett's Test of Sphericity were used to examine assumptions required to carry out EFA (Dziuban and Shirkey 1974). The higher the value of KMO, and the small values of the significance level of Bartlett's test, indicate that factor analysis is feasible. Results are shown in Table A1.

Table A1. KMO and Bartlett's Test results

Kaiser-Meyer-Olkin Measure of Sampling Adequacy	0.700	
Bartlett's Test of Sphericity	Approx. Chi-Square	383.387
	Degrees of Freedom	78
	Significance	0.000

The KMO value of 0.7 and small values of the significance level of Bartlett's test indicate that factor analysis was feasible.

A.2 Factor extraction

The Eigenvalue method, scree plot, and parallel analysis were all used to provide initial guidance on the number of factors.

For the eigenvalue method, the components are the number of factors. The number of appropriate factors is the number for which the "Total" column is greater than 1.0 (here, up to four; see Table A2).

A scree plot is a line plot of the eigenvalues of the factors. The appropriate number of components is found when either the eigenvalues drop below 1 (the horizontal line, also known as the "Kaiser criterion") or there are no longer any drops observed in the magnitude of the eigenvalues (see Figure A1).

Table A2. Eigenvalues

Component	Total	Proportion of variance	Cumulative proportion
1	3.459	0.261	0.261
2	1.482	0.117	0.378
3	1.327	0.109	0.487
4	1.179	0.089	0.576
5	0.967	0.083	0.659
6	0.913	0.075	0.733
7	0.819	0.063	0.797
8	0.639	0.047	0.844
9	0.555	0.044	0.888
10	0.530	0.038	0.926
11	0.469	0.028	0.954
12	0.341	0.024	0.978
13	0.320	0.022	1.000

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Scree plot

Factor Number

A scree plot is also used for a parallel analysis, but here, one compares the real data to the data from a Monte-Carlo-based simulation of uncorrelated (random) normal variables. Again, the appropriate number of components is after the line levels off (see Figure A2).

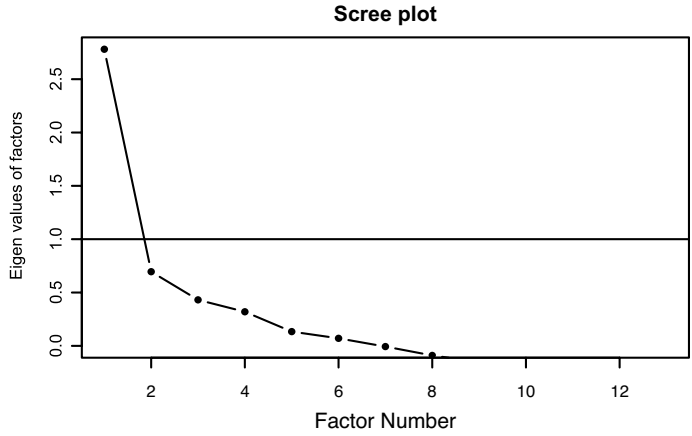


Figure A1. Scree plot.

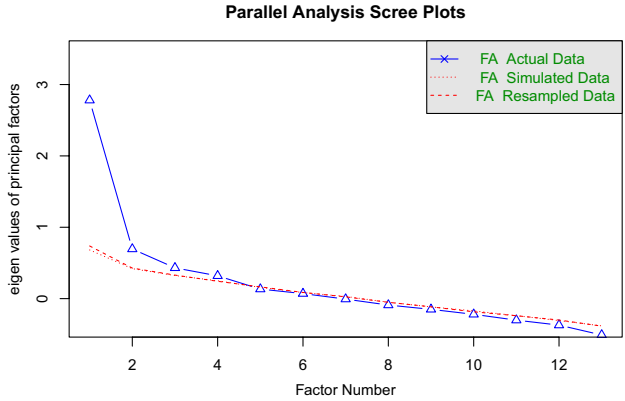


Figure A2. Parallel analysis.

Parallel Analysis Scree Plots

Factor Number

After these initial tests, factor extraction was performed using a robust rescaling-based estimator via MLR (maximum likelihood estimation with Huber-White standard errors and a scaled test statistic). MLR was chosen since it is appropriate for estimating standard errors and chi-square statistics and remains

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press

Table A3. Fit statistics for two-, three-, and four-factor models

Model ($N = 139$)	Chi-square value	Chi-square df	Chi-square/df	CFI	TLI	RMSEA	SRMR
2-factor	105.590	53	1.992	0.837	0.761	0.084	0.067
3-factor	67.833	42	1.615	0.920	0.852	0.067	0.050
4-factor	48.847	32	1.526	0.948	0.873	0.062	0.039

appropriate for small to medium-sized samples (Bentler and Yuan 1999). The MLR was used for parameter estimates with Geomin factor rotation, an oblique method that allows factors to have some degree of correlation. To evaluate EFA model fit, we used the following criteria (Brown 2015; Kline 2023): chi-square/df ratio less than 3, CFI (≥0.95), TLI (≥0.95), RMSEA (≤0.06), and SRMR (≤0.08). We also took into account relatedness between variables for each of the models. Table A3 shows fit statistics for each model tested.

Francisco Calderón is a PhD candidate in Philosophy at the University of Michigan in Ann Arbor, where he also completed a certificate in Science, Technology, and Society. Most of his work focuses on the philosophy and history of physics, specifically on the quantum field theories that make up the standard model of particle physics.

Thomas Colclough is a Postdoctoral Scholar in the Center for Knowledge, Technology, and Society at the University of California, Irvine. He completed his PhD in the Department of Logic and Philosophy of Science at the University of California, Irvine. He works in the philosophy of mathematics, educational development, and STEM education.

Helen Meskhidze is an Assistant Professor in the Departments of Philosophy and Physics at the University of Cincinnati. Previously, she was a post-doctoral fellow at the Black Hole Initiative. She completed her PhD in the Department of Logic and Philosophy of Science at the University of California, Irvine. Her work focuses on foundational issues in philosophy of physics as well as epistemological issues in philosophy of astrophysics. She has also worked on several projects in the scholarship of teaching and learning.

Cite this article: Calderón F, Colclough TM, Meskhidze H (2025). "Feminist and Trauma-Informed Approaches to Teaching Formal Philosophy." *Hypatia* 1–31. https://doi.org/10.1017/hyp.2025.10025

https://doi.org/10.1017/hyp.2025.10025 Published online by Cambridge University Press