

The Stereotype Threat Hypothesis: An Assessment from the Philosopher's Armchair, for the Philosopher's Classroom

GINA SCHOUTEN

According to Stereotype Threat Hypothesis (STH), fear of confirming gendered stereotypes causes women to experience anxiety in circumstances wherein their performance might potentially confirm those stereotypes, such as high-stakes testing scenarios in science, technology, engineering, and math (STEM) courses. This anxiety causes women to underperform, which in turn causes them to withdraw from math-intensive disciplines. STH is thought by many to account for the underrepresentation of women in STEM fields, and a growing body of evidence substantiates this hypothesis. In considering the plausibility of STH as an explanation for women's disproportionate attrition from undergraduate philosophy programs, one is struck by dissimilarities between STEM and philosophy that appear to undermine the applicability of STH to the latter. In this paper, I argue that these dissimilarities are either merely apparent or merely apparently relevant to the plausibility of STH as an explanation for gender disparities in philosophy. I argue further that, if research from STEM uncovers promising strategies for confronting stereotype threat, we should think about how to apply those strategies in our introductory philosophy classrooms.

New evidence suggests that the biggest drop in the proportion of women in philosophy occurs between enrollment in an introductory philosophy class and becoming a philosophy major (Paxton, Figdor, and Tiberius 2012).¹ Because introductory undergraduate philosophy courses constitute a particularly vulnerable joint in the leaky pipeline, we have good reason to think about how to make our introductory courses more hospitable—or less inhospitable—to women undergraduates.

Although the change may be occurring more slowly than some of us would like, my own impression is that philosophy instructors are increasingly embracing this goal. More and more of us recognize that the discipline of philosophy will be better for having more women practitioners, because women's contributions are valuable and

because the changes that would enable more of those contributions are independently good for the discipline. Some agree, further, that the undergraduate women who might have joined our ranks had we provided a more hospitable environment have missed out on the enriching experience of sustained engagement with philosophy. For the sake of the discipline, and for the sake of the women who are discouraged under the status quo, we should work to improve our retention of women. The introductory undergraduate philosophy course is a good place to start.

The gendered performance and retention gap in philosophy has received much less empirical attention than have similar gaps in science, technology, engineering, and math (STEM) disciplines. This should be remedied. The most effective practices for confronting women's attrition likely depend on the actual mechanisms that drive that attrition. We might try to discern those mechanisms to develop remedies for women's attrition; alternatively, we might simply test the efficacy of particular interventions at increasing the presence of women and bring effective measures to scale.

Philosophers are not well trained to do this kind of careful empirical research, but in the absence of the information it provides, we need not sit idly by. Distinctly philosophical skills can be put to good use in service of this goal, even as we await empirical data. First, philosophers can help set the agenda for empirical research. Given constraints on time and resources, we would do well to focus first on testing the most plausible hypotheses available. We can start by conceptually examining the various hypotheses that purport to explain women's attrition from STEM. The philosophy and STEM fields are different in important ways, however, and these differences might render some of the contending explanations of attrition in STEM implausible when applied to the discipline of philosophy. By reflecting on differences between philosophy and other disciplines with pipeline problems, and considering the possible relevance of those differences to diagnosing the pipeline problem in philosophy, philosophers can guide empirical attention toward the most *prima facie* plausible explanations and remedies for our discipline's difficulty with retaining women.

Second, philosophers can clarify and weigh the values at stake in making particular changes within our classrooms. Among remedies that are effective means of improving women's persistence, some might impose a high cost—whereas others might be *beneficial*—in terms of *other* pedagogical or disciplinary values. Some remedies might be good candidates to improve the retention of women in philosophy, *on any (prima facie) plausible explanation of what's causing their disproportionate attrition*. Depending on how these considerations play out, armchair reflection might be sufficient to justify making certain changes within our classrooms immediately, even in the absence of rigorous empirical data. Suppose we discern from the armchair that including more women philosophers on our syllabi is likely to enhance students' learning *whether or not* it increases the retention of women undergraduates, because the contributions of women philosophers are independently valuable and often overlooked. Suppose we discern, further, that including more women on our syllabi would probably be costless in terms of other pedagogical values, and that it would likely improve our classroom environments for women on *any (prima facie) plausible explanation of what's causing their disproportionate attrition*. If so, then there is good

reason to diversify our syllabi immediately. Armchair reflection should not *replace* rigorous empirical investigation, but we can make projections. On the basis of the projections supposed here, we could conclude that there is much to gain and little to lose by diversifying our syllabi. With respect to other classroom changes, the balance of reasons may point in a different direction.

In this paper, I do some armchair theorizing with regard to the Stereotype Threat Hypothesis (STH), which purports to explain the attrition of undergraduate women in STEM. STH has received a great deal of scholarly attention in STEM disciplines as more traditional explanations of gender disparities—including genetic explanations and cultural explanations invoking deeply internalized gender expectations—have come to be regarded as insufficient to account for the entirety of the disparities.² STH hazards a different explanation: Fear of confirming stereotypes about the low aptitude of women relative to men in mathematics causes women to experience anxiety in circumstances wherein their performance might potentially confirm those stereotypes, such as high-stakes testing scenarios in STEM courses. This anxiety causes underperformance, which in turn causes women to disidentify with and withdraw from math-intensive disciplines.³ Increasingly, STH is thought to account for some portion of the underrepresentation of women in high-level courses and professional positions in STEM fields, and a growing body of evidence substantiates this hypothesis.⁴

My project in this paper is twofold. First, I catalogue three apparently powerful conceptual reasons for thinking that STH is a *prima facie* implausible explanation of women's attrition *at the introductory undergraduate level of philosophy*, and argue that those reasons do not hold up upon further scrutiny. Second, I catalogue remedies that apparently improve women's performance and retention in STEM. I propose that we have some reasons to implement similar remedies in philosophy, even absent rigorous empirical data substantiating STH in philosophy. None of these reflections is intended to take the place of rigorous empirical research. But data specific to philosophy may not be available anytime soon.⁵ If improving our discipline for women is a priority, then we philosophers will have to do some trial and error based on evidence gathered in other disciplines. My project constitutes a modest contribution to the philosophical work that can be undertaken now. I conclude that STH should be among the hypotheses investigated in future research, and that there are good reasons immediately to undertake the kinds of remedies that STH calls for.

STH REJECTED?

In the absence of data on women's attrition in philosophy, we might undertake to apply the research on gender disparities in STEM, and the promising interventions that research has uncovered, to philosophy. But crucial differences between philosophy and STEM appear to undermine the applicability of STH to philosophy. Although STH probably accounts for some of the gender gap in undergraduate STEM retention, and perhaps for some of the gender gap in *philosophy at later stages*

of professional advancement (Saul 2012), there are powerful reasons to doubt that STH explains any significant portion of the gender disparity in undergraduate philosophy retention at the introductory level.

Three conditions are common in domains where STH has been investigated, and each appears crucial to its applicability (Steele 1997): First, the stereotype-threatened individuals are *aware* of negative stereotypes about their group's aptitude in the relevant domain. It is this awareness that causes them to experience anxiety. They perceive that others believe their group to be inferior in the relevant domain, and worry that their own performance might confirm that belief. Call this the "anxiety-inducer" condition. Second, there are *performance* disparities between the stereotyped group and the rest of the population, which mediate any further disparities that occur. In STEM, gendered disparities in *performance* mediate *retention* disparities: Anxiety causes underperformance, which in turn causes discouraged women to disidentify with and withdraw from math-intensive disciplines. Call this the "mediating-variable" condition. Finally, stereotype threat affects those who *strongly identify* with the domain in which they are stereotyped. Only because they care about and derive personal validation from that domain does the prospect of confirming negative stereotypes about their group's aptitude cause them the anxiety that leads to underperformance. Call this the "identification (with domain)" condition.

These three conditions are pervasive among the cases invoked to support STH, and are crucially implicated in the most basic explanations of it: Stereotyped individuals *fear that they will confirm stereotypes* about the low aptitude of members of a group to which they belong. Because *they identify with and derive self-worth from their achievement* in the domain in which they are stereotyped, they experience anxiety in circumstances wherein their performance might potentially confirm those stereotypes. Due to this anxiety, group members underperform in those circumstances, and *persistent underperformance in turn* causes them to disidentify with and withdraw from the domain (Steele 1997).

None of these conditions—each of which seems essential to STH—appear to obtain in introductory undergraduate philosophy courses. Consider the "anxiety-inducer" condition: Unlike explanations for gender disparities that invoke deeply internalized mechanisms to explain gender disparities, STH involves an *explicit* recognition by the stereotyped individual that a negative stereotype might be applied to her (Steele 1997, 617–18).⁶ Indeed, studies show that women underperform only in domains, like math, with known stereotypes about the relative aptitude of women and men (Steele 1997). Although gendered stereotypes about philosophical aptitude certainly exist, it is unlikely that introductory undergraduate students know of them.⁷ One common objective of introductory philosophy courses is simply to help students understand what philosophy is: what questions it asks, what capacities it exercises, what skills it develops. Students do not enter our classrooms believing that their gender predicts their success in philosophy, or that others believe that it does. And their limited knowledge of the discipline blocks the application to philosophy of more general stereotypes, such as the gendered stereotype regarding abstract analytical thinking. Students cannot fear confirming a stereotype of which they are unaware;

thus, undergraduate women in introductory philosophy courses appear to lack an apparently essential feature of STH: awareness of a negative perception about their group's aptitude in that domain.

The "mediating-variable" condition constitutes a second apparently essential feature that is apparently absent from introductory philosophy courses. In the STH literature, persistent underperformance mediates the relationship between anxiety and withdrawal from the discipline: Anxiety does not directly cause students to withdraw from the discipline; rather, anxiety leads to underperformance, *which in turn* leads to disidentification with (and subsequent withdrawal from) the discipline (Schmader and Johns 2003; Croizet et al. 2004; Kane et al. 2004; Inzlicht, McKay, and Aronson 2006; Good, Aronson, and Harder 2008). In STEM, there is ample evidence of gender disparities *both* in retention *and* in student performance. In philosophy, we have evidence of women's disproportionate attrition, but we do not (to the best of my knowledge) have compelling evidence that women systematically underperform.

The third important difference between undergraduate introductory philosophy and the disciplines in which STH has been documented concerns the "identification" condition. STH maintains that students experience anxiety at the prospect of confirming stereotypes because they identify with and derive validation from the domain in which they are stereotyped. The effects of stereotype threat, in other words, depend upon "a relationship between oneself and the domains of schooling such that one's self-regard significantly depends on achievement in those domains" (Steele 1997, 616). Only because one identifies with and derives self-worth from the discipline does fear of confirming negative stereotypes trigger the anxiety that drives the effects of STH (Steele 1997, 617; Huguet and Regner 2007; Miyake et al. 2010). Most introductory philosophy students have no prior experience in philosophy on which to base a judgment of their prospects for success, and few identify with the discipline as a source of personal worth or validation. In the United States, students rarely begin introductory courses with plans for extensive study in philosophy. Undergraduates in introductory philosophy courses thus appear to lack the kind of identification essential for STH.

These differences between philosophy and the paradigmatic cases in which STH has been vindicated might lead us to conclude that STH is implausible as an explanation of women's disproportionate attrition after their first undergraduate philosophy course, and to infer that we ought not to prioritize investigation of STH in our agendas for combating women's attrition from philosophy. Given constraints on time and resources, we might conclude that we should focus empirical investigation and strategic intervention on more *prima facie* plausible explanations of philosophy's leaky pipeline.

That would, I think, be a mistake. The next section of this paper argues that the dissimilarities between philosophy and STEM are either *merely apparent* or *merely apparently relevant* to the plausibility of STH in introductory undergraduate philosophy courses. I argue that, despite appearances to the contrary, philosophy *does* meet the anxiety-inducer condition and the identification condition. And I argue that we have good reason not to regard the mediating-variable condition as necessary to the

occurrence of stereotype threat. Perhaps it is necessary, but we cannot know that it is from the mere fact that it has been present in the cases so far studied; moreover, there is some positive reason to suspect that it is *not* necessary. My conclusion is simple: If we want to enhance the educational experience of undergraduate women by sharing philosophy more fully with them, and if we want to improve the quality of philosophy by bringing women more fully into the philosophical community, then we should prioritize investigating STH in setting the empirical research agenda regarding philosophy's leaky pipeline.

STH RECONSIDERED

I think that the dissimilarities between philosophy and STEM do not provide good reason to discount the plausibility of STH, even though the dissimilarities apply to apparently definitional features of STH. Consider the anxiety-inducer condition first: The condition appears to be absent from introductory philosophy classes because students do not enter the classes aware of gendered stereotypes concerning philosophical performance. But whether or not they *enter* with the perception that women are less-promising philosophers than men (or with the perception that *others* have that perception), they very plausibly *develop* that perception within the first few weeks of class. Even if students do not show up on the first day aware of any gendered stereotypes regarding philosophy, we give them ample data that, cumulatively, supports the inference that such stereotypes exist.

First, many undergraduates will be exposed to only male experts during their first experience with philosophy. Especially in large classes with multiple teaching assistants, gender homogeneity may constitute a striking bit of data indeed. Students might explicitly infer from the prevalence of male experts that men tend to excel in philosophy, whereas women do not. Or they might reach this conclusion indirectly, noticing that the gender makeup of philosophy resembles that of math and science, and then applying gendered stereotypes from other disciplines to philosophy. Second, even the most well-meaning among us are likely to distribute syllabi on which men's contributions far exceed those of women. Even if students do not notice this immediately, they are likely during the first few weeks of the course to notice that all the articles they are reading were written by men, if only through our repeated use of male pronouns.⁸ Finally, many of us begin our courses with a brief introduction to logic. From this, students may infer that philosophy is similar to math, or that the two disciplines share important foundational components. They may infer that being good (bad) at math bodes well (ill) for one's prospects in philosophy. Students may then apply gendered math stereotypes to philosophy, or simply experience anxiety at the prospect of confirming math stereotypes in the philosophy classroom, much as they appear to experience anxiety at the prospect of confirming math stereotypes in physics classrooms (Saul 2012).

A few weeks of philosophy class is insufficient time for students to deeply internalize gendered stereotypes, but the *explicit perception* of stereotypes that drives the effects

of STH may plausibly develop in that time. If anything, the possibility that the courses themselves generate the stereotype gives us additional reason to entertain STH, as stereotype threat appears to have the greatest effect on those who experience the stereotype *merely* on a superficial level (Steele 1997, 617).⁹ During the first few weeks of introductory philosophy courses, our students receive ample data on which to form quick impressions about who succeeds as an expert in philosophy. These impressions constitute precisely the kind of mechanism that STH alleges to induce anxiety and lead to underperformance and attrition. The fact that students do not *show up* in our classrooms already aware of gendered stereotypes in philosophy—if indeed that is a fact—gives us no good reason to discount STH as a plausible explanation of gender disparities in our discipline. The anxiety-inducer condition may be necessary for the applicability of STH, but we have good reason to believe that introductory philosophy courses meet that condition, if only by *creating* the impression that men have greater philosophical aptitude than women.

We also have good reason to believe that the identification condition is met. First, philosophy may be very important to the self-concept of many of our students even though they lack a full and accurate understanding of what philosophy is. They may identify with their pre-theoretical notion of philosophy, as an endeavor of critical reflection in search of deep truth, or as an endeavor of winning arguments with aesthetically impressive rhetorical skills. Some may enter our courses already thinking of themselves as philosophers, on the grounds that they have deep thoughts, or question authority, or take mind-altering drugs. Whatever the merits or demerits of their conception of philosophy, these students may be just as susceptible to feeling anxious at the prospect of confirming newly-noticed stereotypes about their group's aptitude in the discipline. Moreover, identification can come in degrees. So even if underperformance in philosophy were *less* threatening to students than is underperformance in some other field of study, the prospect for failure in *any* domain of schooling is likely to cause anxiety among many of the intelligent, committed students who find their way into higher education.

Finally, even if our students do *not* self-identify with philosophy, philosophy may nonetheless meet a plausibly refined version of the identification condition. Notice that, in order for a stereotype to be self-relevant in the requisite way, the *stereotyped* domain need not be identical to the *domain of identification*. Consider one of the paradigmatic cases for STH: an advanced college physics course for physics majors in which women underperform due to anxiety at the prospect of confirming negative stereotypes about women in math. The stereotyped domain is math, but the women's domain of identification is the discipline of physics. The two domains are nonidentical, but they are related in an important way: Poor performance *in physics*—the domain of identification—would confirm a stereotype about women's abilities *in math*—the stereotyped domain. Because providing that confirmation would undermine their standing in *physics*, undergraduate women physicists experience the stereotype as threatening.

We can deploy this insight—that the stereotyped domain and the domain of identification need not be identical—to develop a plausible hypothesis about what might

transpire in philosophy. Most college students are strongly identified *with school*, and with academic achievement. They have a favorable track record on which to base assessments of their prospects for success in academics, and their self-concept as “students” makes schooling self-definitional such that the prospect of diminishment in that domain might cause them great anxiety. Because poor performance in *any* academic domain threatens to diminish students’ academic success more generally, philosophy might plausibly be self-relevant in the manner apparently essential to STH, even if our students tend not to strongly identify with philosophy.

Despite appearances to the contrary, both the anxiety-inducer condition and the identification condition might plausibly obtain in undergraduate introductory philosophy courses. Thus, the noted dissimilarities between STEM and philosophy do not, on their own, undermine the plausibility of STH as an explanation for women’s attrition from philosophy after their first experience in the discipline. Now consider the mediating-variable condition: We lack evidence that women underperform in philosophy, and thus lack evidence of the condition alleged to mediate the effects of stereotype threat on attrition. Presently, I will argue that we have good reason to doubt that the mediating variable is essential. But it is worth noting in the meantime that we have, as yet, no evidence that women do *not* underperform in philosophy. Given the dearth of good research on persistence and performance disparities in philosophy, the absence of evidence *confirming* underperformance can quite plausibly be explained by our failure to search for that evidence. Moreover, the kinds of standardized tests that reveal performance disparities in STEM disciplines simply do not exist in philosophy. We would likely not know about any performance disparity in philosophy, so our unawareness is no evidence that the disparity does not exist. We must be clear, further, about just what *kind* of underperformance mediates the effects of STH. An underperforming group does not necessarily perform less well than non-group members; it just performs less well than other indicators would suggest it should. Women *might* underperform in philosophy—relative to, for example, their expectations for academic success based on past experience in other disciplines—without performing any less well than men. And they might notice this underperformance, experience it as failure, and on these grounds disidentify with and withdraw from philosophy.

Suppose, however, that women undergraduates do *not* underperform in introductory philosophy courses. Even if so, we should not discount STH. Rather, we should provisionally reject underperformance as a necessary condition for its application. Although underperformance is present in important ways in the paradigmatic cases of STH thus far studied, it is plausible to think that stereotype threat might affect persistence directly or through *other* mediating variables, bypassing the mechanism of underperformance. First, a lack of underperformance does not suffice to show that women feel no anxiety in philosophy courses. The high-stakes testing scenarios where effects of STH are manifest in STEM are less common in philosophy. And when philosophy students *do* take high-stakes tests, their performance is less dependent upon the cognitive mechanism thought to be undermined by the anxiety brought on by stereotype threat: working memory capacity (Schmader and Johns 2003; Croizet et al. 2004; Kane et al. 2004; Inzlicht, McKay, and Aronson 2006). Nor is students’

performance on the deliberative writing assignments characteristic of philosophy classes dependent on working memory capacity. So even if stereotypes did cause students to experience anxiety in philosophy, it may simply not manifest itself on the assessments typical of philosophy classes.

This anxiety may have other discernible effects, however, including withdrawal from the discipline. If women disproportionately associate feelings of anxiety with philosophy courses, we might expect fewer women than men to persist in the discipline. In this way, stereotype threat might undermine persistence directly; alternatively, there might be mediating mechanisms other than underperformance. Consider, for example, the possibility that anxiety causes women to experience diminished working memory capacity and thus discourages them from participating as freely in classroom discussions as they otherwise would. They might perceive this reticence as underperformance, or simply resent feeling (perhaps uncharacteristically) inhibited. If anxiety cannot directly lead to attrition, then *perceived* underperformance or *perceived* incompetence might plausibly mediate the relationship. Perhaps underperformance really is a necessary condition for the applicability of STH, but the mere fact that it obtains in STEM does not make it so, and other intuitively plausible causal stories are available.

Because of the strength of the evidence substantiating STH in STEM fields, we have some presumptive reason to consider it as an explanation for the leaky pipeline in philosophy. Though important dissimilarities threaten to override that presumption, the case for discounting STH does not hold up under scrutiny. It is plausible that the anxiety-inducer condition and the identification condition *do* obtain in philosophy; and it is plausible that the mediating-variable condition need not obtain in order for stereotype threat to operate. Given the evidence substantiating STH in STEM, and given that these apparent defeaters do not undermine the plausibility of applying STH to philosophy, we should make the investigation of STH a priority in setting the agenda for empirical research.

But what should philosophy teachers do in the meantime? In the next section, I discuss strategies that have been effective in improving the performance and retention of stereotype-threatened individuals in domains in which STH has been studied. I argue, rather modestly, that we have *some* reason to implement these remedies in our introductory philosophy courses, even in advance of data that substantiate STH in that context.

STH APPLIED

Research from STEM and other fields generates promising remedies to stereotype threat. But would it be premature to enact applications of those remedies within our own classrooms? I have not claimed that STH is *the* explanation for gender disparities in philosophy attrition at the introductory level, or even that it is *part of* the explanation. I have merely challenged apparently powerful conceptual reasons for ruling it out. I have argued that these apparently powerful reasons do not hold up under

scrutiny, and that STH should be part of the agenda for empirical research. The prima facie plausibility of STH generates some reason in favor of undertaking the remedies proven successful in other fields, but changing our courses is not costless. We might reasonably hesitate to overhaul our courses in response to a hypothesis that has not been empirically verified in our discipline. Our willingness to make changes on the basis of armchair reflection alone should reflect a weighting of various considerations: the plausibility of the armchair reflection, the depth of our commitment to remedying women's attrition, the likelihood that empirical verification of STH in philosophy will be forthcoming, the costliness of the changes STH prescribes, and the extrinsic benefits those changes are likely to yield.

I believe that the armchair reflection is plausible and that we should be fervently committed to retaining more women in the field. I suspect that the weight of incentives will likely incline empirical scholars to focus their attention on other disciplines that are larger or more highly esteemed than philosophy, and that we will accordingly have to engage in some trial-and-error investigations of our own if we are to confront the leaky-pipeline problem anytime soon.¹⁰ In this section, I make some rather modest remarks regarding the last two considerations we must weigh. First, I suggest that classroom interventions to counter stereotype threat are relatively low in cost compared with the remedies likely to be prescribed by alternative explanations of the leaky pipeline. In making this point, I catalogue the interventions that have been used successfully in other disciplines to counter stereotype threat.¹¹ In closing, I raise for consideration the possibility that these interventions might yield extrinsic benefits, in addition to whatever gains they produce in women's retention.

A growing body of empirical evidence suggests that low-cost, easily implemented interventions can be successful in mitigating the effects of stereotype threat (Miyake et al. 2010). Because STH posits a *situational* cause of gender disparities, it generates *situational* prescriptions for combating those disparities.¹² Countering genetic or socio-cultural sources of attrition would likely involve finding ways to change students, change society, or change the discipline of philosophy. Consider the plausible claim that some portion of women's disproportionate attrition is due to deeply internalized gender norms about who should be inclined to and excel at abstract theoretical thinking. To remedy this, we would need to change the norms themselves, or change philosophy to require less abstract thinking, or somehow interrupt the application of the norms to the discipline of philosophy. None of that sounds simple or straightforward. To counter the effects of stereotype threat, on the other hand, we need change nothing at all about our students or our discipline or society at large. We can remedy whatever portion of women's attrition is caused by stereotype threat simply by changing the circumstances of our classrooms to provide undergraduate women with less threatening educational environments.

The STH literature shows that this can successfully be accomplished in two main ways: first, by providing students with less data from which to derive gendered impressions about the relative prospects for success of men and women; second, by making evaluation processes less anxiety-inducing to stereotype-threatened students. By briefly considering just a few examples of possible interventions in each category,

I hope to show that the classroom changes STH calls for are fairly low in cost, at least relative to alternative explanations of the leaky pipeline.

The first family of promising interventions undertakes to “affirm domain belongingness” by removing stereotype-reinforcing messages and replacing them with stereotype-challenging messages that affirm the threatened group’s belongingness within the domain (Inzlicht and Ben-Zeev 2000; Sekaquaptewa and Thompson 2003; Huguet and Regner 2007). Because gender stereotypes in philosophy classrooms are likely generated and reinforced by gender imbalances in the set of experts to which students are exposed, we can send counter-stereotypic messages by exposing philosophy students to more women experts, both live in our classrooms and on our syllabi.¹³ Within the discipline, we can continue efforts to promote the careers of talented women philosophers and allocate teaching responsibilities so as to expose as many undergraduates as possible to at least one woman instructor.¹⁴ We can teach the work of our talented women colleagues, and invite those colleagues to our classes as guest discussion participants. We can prioritize inclusion of women when reflecting on possible opportunities for visiting experts, and diversify our syllabi to include contributions from women philosophers (Paxton, Figdor, and Tiberius 2012; Saul 2012; Lombrozo 2013).

Because studies suggest that the effects of stereotype threat are reduced when students think of their capacities as malleable, we might also affirm domain belongingness by emphasizing that intelligence—and philosophical aptitude in particular—is an expandable trait that can be improved over time with hard work (Aronson, Fried, and Good 2002; Good, Aronson, and Inzlicht 2003; and Cohen et al. 2006).¹⁵ We might explicitly discuss with students the skills that are relevant to successful study of philosophy, and the ways in which our courses will help them refine those skills. Throughout the course, we might provide affirmative feedback that specifically assesses those skills. We might also find ways to incorporate lessons about the expandability of intelligence into the content of our courses. For example, an ethics professor might discuss the environmental conditions conducive to intellectual development within the context of a unit on social justice and early childhood education.

The second family of successful interventions involves amending evaluation procedures to make them less anxiety-inducing to stereotype-threatened students. This has been executed in two ways: first, by reframing the assignments themselves so that stereotypes are not activated or are offset by other environmental cues; second, by incorporating values-affirming exercises into assignments. In STEM, reframing assignments has consisted in simply telling students, prior to administering exams, that the exams are “gender fair” (Steele 1997; Spencer et al. 1999; Quinn and Spencer 2001; Cadinu et al. 2006; Good, Aronson, and Harder 2008).¹⁶ In STEM, where women often underperform on exams, this tactic raises ethical questions regarding the permissibility of misleading students. But recall that gender disparities in philosophy may operate independently of underperformance. If telling students that women perform as well as men in philosophy reduced anxiety and attrition from the discipline, then we would have a defeasible reason to do so. If, additionally, women actually *do* perform as well as men in philosophy, then assurances to that effect seem warranted.

Values affirmation involves students reflecting on self-defining values prior to performing the evaluated activity. This reflection appears to reinforce students' self-worth, inoculating them against the anxiety otherwise caused by stereotypes (Schiebinger 2001; Cohen et al. 2006; Cohen and Garcia 2008; Miyake et al. 2010). One study asked students in a science course to write reflectively about personally important values like family and friends. The topic of the reflections was unrelated to the course subject matter, but the exercise was conducted just before the students took an exam, and appears to have closed the gender performance gap on that exam (Miyake et al. 2010. See also Shih, Pittinsky, and Ambady 1999; Marx, Stapel, and Muller 2005; Martens et al. 2006; Elizaga and Markman 2008). Values-affirmation exercises could be introduced into philosophical evaluation scenarios as well. For example, an essay exam might invite reflection on the *personal* implications of a principle of justice, on any influence these personal implications have had on their endorsement of the principle, and on whether personal considerations constitute legitimate grounds for such endorsement.¹⁷

If women's withdrawal from philosophy is plausibly *unmediated* by underperformance, however, we might opt instead for a more general approach to values affirmation that is not limited to assessment scenarios. We can incorporate value-affirming practices into our curricula more broadly by choosing our examples so as to invite students to reflect on personally important values. In class discussion, we can resist the urge to subordinate personal reflection in favor of (allegedly) more objective argumentation.

None of these changes strikes me as particularly costly in terms of resources, know-how, or transition costs. And the changes seem all the more manageable when compared with those likely to be called for by alternative explanations of the gender gap in philosophy. If it turns out that STH accounts for *none* of the gender gap, then we might have made these changes for nothing. But the risk is small, and the *likely* reward—assuming STH *plausibly* accounts for some of the gap—is large in comparison.

Moreover, the payoff of these changes might not depend on whether or not STH explains the attrition of undergraduate women from philosophy. There might be independent pedagogical and disciplinary benefits to be reaped. Whether or not women are disproportionately affected by the anxiety students feel in philosophy courses, students would likely learn more if we could make them all feel more at ease. In the discipline more broadly, we might all do better philosophy if our disciplinary practices were more affirmative and less anxiety-inducing. The discipline might also benefit from a more general strategy of increased attentiveness to the work produced by members of underrepresented and marginalized groups, for example if circumstances obtain such that having a diverse set of voices in the conversation increases the likelihood of arriving at truth. Finally, implementing these interventions might provide additional evidence on which to evaluate the applicability of STH to philosophy. If more rigorous data developed by social scientists is not forthcoming, we may need to rely on the less formal data we can glean from trial and error. If interventions that undertake to undermine stereotypes or ease anxiety in philosophy classrooms prove effective when implemented in isolated classrooms, we will have more evidence to support bringing them to scale, and stronger reason to try more ambitious remedies to stereotype threat. If, on the other hand, those interventions turn out not

to be successful, we will have evidence that *challenges* the applicability of STH in philosophy. Either way, there is something to be learned about our discipline if we are thoughtful in implementing these changes and reflective in evaluating the outcomes. These conjectures are highly speculative. But if the changes we would need to make to counteract stereotype threat are changes we have independent reason to make, that weighs in favor of us making them immediately, rather than awaiting empirical confirmation that stereotype threat is, indeed, part of the story behind undergraduate women's attrition from philosophy.

The evidence substantiating STH in STEM provides some reason to think something similar is at work in philosophy. But there are apparently powerful reasons to discount STH as an explanation for gender disparities in undergraduate retention in philosophy. More study is clearly called for, but I have argued that these apparently powerful defeaters do not hold up under scrutiny. As we continue to await rigorous empirical data regarding the causes of and remedies for gender disparities in philosophy, the armchair plausibility of STH gives us some reason to think carefully about how to apply remedies that have successfully counteracted stereotype threat in other domains. The fact that these remedies are relatively low-cost in terms of resources, know-how, and transition costs provides further reason to do so, as does the likelihood that these remedies will yield general pedagogical and disciplinary benefits, apart from their effects in plugging the leaky pipeline. I don't pretend that armchair reflection is an adequate substitute for actual data in thinking about the leaky pipeline. But it's what we do, and it's not nothing.

NOTES

Thanks to Jeff Behrends, Harry Brighouse, and Josh DiPaolo for many helpful comments and suggestions on previous drafts. This paper owes much to their generosity and insights. Thanks also to anonymous reviewers at *Hypatia*, and to audiences at the APA Central Division's annual meeting and the Philosophical Society at Illinois State University for helpful comments and challenges.

1. For additional content excerpted from this and later references, see online supplemental information.

2. Steele 1997; Spencer, Steele, and Quinn 1999. On genetic explanations, see Benbow and Stanley 1980; 1983. On cultural explanations, see Eccles-Parsons et al. 1983; Eccles and Jacobs 1986; and Eccles 1987.

3. For pioneering work on STH, see Steele 1997. See also Schmader and Johns 2003; Croizet et al. 2004; Kane et al. 2004; Inzlicht, McKay, and Aronson 2006; and Good, Aronson, and Harder 2008.

4. See Aronson, Quinn, and Spencer 1998; Steele, Spencer, and Aronson 2002; Maass and Cadinu 2003; Aronson and Steele 2005; and Schmader, Johns, and Forbes 2008 for reviews. STH has also been defended as an explanation for racial disparities. The race disparity in philosophy demands attention as well, and a pressing future project is to explore whether the considerations raised here may also be helpful in redressing the race disparity.

5. See online supplemental information.
6. For additional information about the conditions apparently essential to STH, see online supplemental information.
7. See Saul 2012. For a possible challenge to this claim, see Haslanger 2008; Calhoun 2009.
8. According to a study conducted at Georgia State University by Toni Adleberg, Morgan Thompson, and Eddy Nahmias, women comprise on average only six percent of authors in introductory philosophy textbooks. Moreover, the gender composition of course syllabi appears to mediate the relationship between gender and persistence in philosophy. See Lombrozo 2013.
9. Because stereotypes can be most threatening to those who would otherwise *favorably* assess their prospects within a domain, Steele conjectures that stereotype threat is likely to have “its greatest effect on . . . those who have not internalized the group stereotype to the point of doubting their own ability and have thus remained identified with the domain” (Steele 1997, 617).
10. See online supplemental information.
11. I limit my discussion here to variants of strategies that have been proven effective. But if STH is indeed a plausible explanation for women’s attrition in philosophy, we should also consider possible interventions that have *not* been attempted in other disciplines, but seem likely nonetheless to be promising remedies *in philosophy*. For example, if gendered stereotypes in philosophy classes are plausibly formed by students on the basis of the frequency or assertiveness of men’s discussion contributions relative to women’s, then we have reason to encourage more balanced participation, whether or not measures to do so have been proven effective in other disciplines. With variance in how stereotype threat operates in different disciplines come variations in the possible strategies for confronting it. I thank an anonymous reviewer for suggesting that I consider additional interventions.
12. For example, Steele suggests that “these problems might be more tractable through the situational design of schooling . . . that secures . . . students in the belief that they will not be held under the suspicion of negative stereotypes about their group” (Steele 1997, 624).
13. The effects of exposing students to successful role models from the stereotyped group are well documented in the STEM literature. See Marx and Roman 2002; Marx, Stapel, and Muller 2005; and Huguet and Regner 2007.
14. But we must carefully avoid over-taxing women faculty members.
15. These interventions can be quite simple: One study showed that underperformance can be reduced by showing students an instructional video about the ability of the brain to form new neural pathways (Good, Aronson, and Inzlicht 2003).
16. In a study of students enrolled in the most advanced calculus sequence at a large public university, women who were assured that a test was free of gender bias outperformed women in a control group; they also outperformed men in both the control and experimental groups (Good, Aronson, and Harder 2008). See also Steele and Aronson 1999.
17. On the other hand, we might worry that subjecting personal values to philosophical scrutiny would offset the anxiety-undermining effects of values affirmation. This possibility merits careful consideration.

REFERENCES

- Ambady, N., M. Shih, A. Kim, and T. L. Pittinsky. 2001. Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science* 12 (5): 385–390.
- Aronson, J., D. M. Quinn, and S. J. Spencer. 1998. Stereotype threat and academic underperformance of minorities and women. In *Prejudice: The target's perspective*, ed. J. K. Swim and C. Stangor. San Diego: Academic Press.
- Aronson, J., C. Fried, and C. Good. 2002. Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology* 38 (2): 113–25.
- Aronson, J., and M. Inzlicht. 2004. The ups and downs of attributional ambiguity: Stereotype vulnerability and the academic self-knowledge of African American college students. *Psychological Science* 15 (12): 829–836.
- Aronson, J., and C. M. Steele. 2005. Stereotypes and the fragility of human competence, motivation, and self-concept. In *Handbook of competence and motivation*, ed. C. Dweck and E. Elliot. New York: Guilford.
- Benbow, C. P., and J. C. Stanley. 1980. Sex differences in mathematical ability: Fact or artifact? *Science* 210 (4475): 1262–64.
- . 1983. Sex differences in mathematical reasoning ability: More facts. *Science* 222 (4627): 1029–31.
- Cadinu, M., A. Maass, M. Lombardo, and S. Frigerio. 2006. Stereotype threat: The moderating role of locus of control beliefs. *European Journal of Social Psychology* 36 (2): 183–97.
- Calhoun, Cheshire. 2009. The undergraduate pipeline problem. *Hypatia* 24 (2): 216–23.
- Cohen, G. L., J. Garcia, N. Apfel, and A. Master. 2006. Reducing the racial achievement gap: A social-psychological intervention. *Science* 313 (5791): 1307–10.
- Cohen, G. L., and J. Garcia. 2008. Identity, belonging, and achievement: A model, interventions, implications. *Current Directions in Psychological Science* 17 (6): 365–69.
- Croizet, J. C., G. Despres, M. E. Gauzins, P. Huguet, J. P. Leyens, and A. Meot. 2004. Stereotype threat undermines intellectual performance by triggering a disruptive mental load. *Personality and Social Psychology Bulletin* 30 (6): 721–31.
- Eccles, J. S. 1987. Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly* 11 (2): 135–72.
- Eccles, J. S., and J. E. Jacobs. 1986. Social forces shape math attitudes and performance. *Signs: Journal of Women in Culture and Society* 11 (2): 367–80.
- Eccles-Parsons, J. S., T. F. Adler, R. Futterman, S. B. Goff, C. M. Kaczala, J. L. Meece, and C. Midgley. 1983. Expectations, values, and academic behaviors. In *Achievement and achievement motivation*, ed. J. T. Spence. New York: Freeman.
- Elizaga, R., and K. Markman. 2008. Peers and performance: How in-group and out-group comparisons moderate stereotype threat effects. *Current Psychology* 27 (4): 290–300.
- Good, C., J. Aronson, and M. Inzlicht. 2003. Improving adolescents' standardized test performance: An intervention to reduce the effects of stereotype threat. *Journal of Applied Developmental Psychology* 24 (6): 645–62.

- Good, C., J. Aronson, and J. A. Harder. 2008. Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology* 29 (1): 17–28.
- Haslanger, Sally. 2008. Changing the ideology and culture of philosophy: Not by reason (alone). *Hypatia* 23 (2): 210–23.
- Healy, Kieran. 2011. Percentage of Ph.Ds awarded in the U.S. to women in 2009. <http://kieranhealy.org/blog/archives/2011/02/04/gender-divides-in-philosophy-and-other-disciplines/> (accessed September 22, 2014).
- Huguet, Pascal, and Isabelle Regner. 2007. Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of Educational Psychology* 99 (3): 545–60.
- Inzlicht, M., and T. Ben-Zeev. 2000. A threatening intellectual environment: Why females are susceptible to experiencing problem-solving deficits in the presence of males. *Psychological Science* 11 (5): 365–71.
- Inzlicht, M., L. McKay, and J. Aronson. 2006. Stigma as ego depletion: How being the target of prejudice affects self-control. *Psychological Science* 17 (3): 262–69.
- Kane, M. J., D. Z. Hambrick, S. W. Tuholski, O. Wilhelm, T. W. Payne, and R. W. Engle. 2004. The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General* 133 (2): 189–217.
- Keller, J. 2002. Blatant stereotype threat and women's math performance: Self-handicapping as a strategic means to cope with obtrusive negative performance expectations. *Sex Roles* 47 (3-4): 193–98.
- Keller, J., and D. Dauenhimer. 2003. Stereotype threat in the classroom: Dejection mediates the disrupting threat effect on women's math performance. *Personality and Social Psychology Bulletin* 29 (3): 371–81.
- Lockwood, P. 2006. Someone like me can be successful: Do college students need same-gender role models? *Psychology of Women Quarterly* 30 (1): 36–46.
- Lombrozo, Tania. 2013. Name five women in philosophy. Bet you can't. <http://www.npr.org/blogs/13.7/2013/06/17/192523112/name-ten-women-in-philosophy-bet-you-can-t> National Public Radio blog (accessed on July 31, 2013).
- Maass, A., and M. Cadinu. 2003. Stereotype threat: When minority members underperform. *European Review of Social Psychology* 14 (1): 243–75.
- Martens, A., M. Johns, J. Greenberg, and J. Schimel. 2006. Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology* 42 (2): 236–43.
- Marx, D. M., and J. S. Roman. 2002. Female role models: Protecting female students' math test performance. *Personality and Social Psychology Bulletin* 28 (9): 1185–97.
- Marx, D. M., D. A. Stapel, and D. Muller. 2005. We can do it: The interplay of construal orientation and social comparisons under threat. *Journal of Personality and Social Psychology* 88 (3): 432–46.
- Miyake, Akira, Lauren E. Kost-Smith, Noah D. Finkelstein, Steven J. Pollock, Geoffrey L. Cohen, and Tiffany A. Ito. 2010. Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science* 330 (6008): 1234–37.

- 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.
- Quinn, D. M., and S. J. Spencer. 2001. The interference of stereotype threat with women's generation of mathematical problem-solving strategies. *Journal of Social Issues* 57 (1): 55–71.
- Saul, Jennifer. 2012. Women in philosophy. *The Philosophers' Magazine*. <http://philosophypress.co.uk/?p=1079> (accessed on July 31, 2013).
- Schiebinger, L. L. 2001. *Has feminism changed science?* Cambridge, Mass.: Harvard University Press.
- Schmader, T., and M. Johns. 2003. Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology* 85 (3): 440–52.
- Schmader, T., M. Johns, and M. Barquissau. 2004. The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles* 50 (11-12): 835–50.
- Schmader, T., M. Johns, and C. Forbes. 2008. An integrated process model of stereotype threat effects on performance. *Psychological Review* 115 (2): 336–56.
- Sekaquaptewa, D., and M. Thompson. 2003. Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal of Experimental Social Psychology* 39 (1): 68–74.
- Shih, M., T. Pittinsky, and N. Ambady. 1999. Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science* 10 (1): 80–83.
- Spencer, S. J., C. M. Steele, and D. Quinn. 1999. Stereotype threat and women's math performance. *Journal of Experimental Social Psychology* 35 (1): 4–28.
- Steele, C. M. 1997. A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist* 52 (6): 613–29.
- Steele, C. M., and J. Aronson. 1999. Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* 69 (5): 797–811.
- Steele, C. M., S. Spencer, and J. Aronson. 2002. Contending with images of one's group: The psychology of stereotype and social identity threat. In *Advances in experimental social psychology* 34, ed. M. Zanna. San Diego: Academic Press.
- Stout, Jane G., Nilanjana Dasgupta, Matthew Hunsinger, and Melissa A. McManus. 2010. STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology* 100 (2): 255–70.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Online Supplemental Information

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the author for the article.