In Press, Journal of Experimental Social Psychology

Gender Representation Cues Labels of Hard and Soft Sciences

Alysson E. Light[1]

Tessa M. Benson-Greenwald[2]

Amanda B. Diekman[2]

[1] Corresponding author; Department of Behavioral and Social Sciences, University of the Sciences, Philadelphia, PA, 19104, a.light@usciences.edu

[2] Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St. Bloomington, IN 47405-7007

The authors thank Alexis Schirling her for assistance with this research, and Alyx McKiernan, Ilham Smaili, Jakob Shea, and Suhina Patel for assistance collecting pilot data.

Word Count: 10,589

Abstract

While women's representation in STEM fields has increased over the past several decades, some fields have seen a greater increase women's participation than others. In the present research, we explore how women's participation in STEM disciplines influences labeling of those disciplines as hard vs. soft sciences. Study 1 found that increasing perceived participation of women in a STEM discipline increased the likelihood that participants would label it a soft science. Study 2 found that among people who did not work in science, this tendency to associate women's participation with soft science was correlated with endorsement of stereotypes about women's STEM competency. And Studies 3A and 3B showed that labeling disciplines as soft sciences led to the fields being devalued, deemed less rigorous, and less worthy of federal funding. These studies show that stereotypes about women's STEM competency can impact perceptions of fields in which women participate, with consequences for how scientific disciplines are perceived.

Keywords: Gender stereotypes, STEM, Gender gaps, Hard and Soft Sciences

Gender Representation Cues Labels of Hard and Soft Sciences

Colloquially, people distinguish between "hard sciences" and "soft sciences". The so-called hard sciences include natural or physical sciences, including domains such as physics, chemistry, biology, and computer science, whereas the so-called "soft sciences" include psychology, sociology, and political science. The so-called hard sciences are imbued with perceptions of objectivity, whereas the so-called "soft sciences" are imbued with perceptions of subjectivity. For example, MRI evidence from the hard science discipline of neuroscience is perceived as more objective than cognitive test evidence from the "soft science" discipline of cognitive psychology (Munro & Munro, 2014). Yet the lines between these categories are at best blurry: Indeed, research on substrates of psychiatric illness and the physics of social neuroscience blend the "soft" science of psychology with the hard sciences of biochemistry and neuroscience. In this research, we examine how perceived gender representation within fields of study serve as heuristics to judge whether a field is a hard vs. soft science. Specifically, we argue that fields with greater representation of women are labeled as "soft sciences" and, consequently, devalued.

Stereotypes of Science as "Soft" or "Hard"

Consistent with social structural theories of stereotype content (Eagly & Koenig, 2021), beliefs about science, technology, engineering, and math (STEM) fields are associated with the groups observed to occupy these roles. This principle carries implications not only for overall perceptions of science, but for particular subcategories of science (i.e., "soft" or "hard" science).

Despite increases in women's engagement in STEM educational and career paths, women continue to be underrepresented in STEM (National Center for Science and Engineering Statistics, 2021). The link between STEM categories and maleness is strong: The traits ascribed to scientists overlap more with the traits ascribed to men than with those ascribed to women

(Carli et al., 2016), and concepts related to science and math are more easily associated with men than with women (Kessels, Rau, & Hannover, 2006; Nosek & Smyth, 2007). Additionally, students asked to draw a scientist are more likely to draw men than women (Miller et al., 2019; Rahm & Charbonneau, 1997). Such stereotypes that associate men or maleness with STEM can dampen women's interest in entering STEM fields (see Cheryan et al., 2017, for review).

Yet the reverse causal direction also operates —that is, the representation of women in a field can influence stereotypes about that field (Cejka & Eagly, 1999). That fields are judged in part by their gender ratios is consistent with research in sociology on the "feminization" of fields (England et al., 2007). Analyses of men and women's pursuit of doctoral degrees in different fields from 1971 to 2002 indicate that once women's representation reached a certain point (greater than 24%), men were subsequently less likely to pursue doctoral degrees in that field. This pattern is consistent with the devaluation of "feminized" fields, suggesting that perceptions of a field (and individuals' choices to enter that field) are influenced by gendered representation. Thus, the presence of women in scientific fields can change perceptions of specific STEM disciplines.

Labeling Sciences as "Soft" or "Hard." We posit that one manifestation of the gender stereotyping of science fields is the use of the labels soft science and hard science. The origin of these terms is somewhat unclear—Storer (1967) highlighted the two categories, but does not seem to have originated them. Empirical support for the distinction is largely credited to Biglan (1973), who asked faculty to sort academic disciplines (including fine arts and humanities) in similar clusters. Using multidimensional scaling, he identified three underlying dimensions that predicted clustering. These were labeled "hard/soft", "pure/applied", and "life/non-life." Biglan explained the hard/soft distinction by referencing Kuhn's (1962) concept of paradigm, arguing that the fields categorized as "hard" were paradigmatic, meaning they had "greater consensus about content and method" (Biglan, 1973, pp. 2020); the "soft" fields were classified as non-paradigmatic. This classification system has been persistently used in studies of higher education, despite concerns about flattening dimensions into categories (Simpson, 2015), and contradictions in how disciplines are categorized.

Perhaps unsurprisingly, women have historically participated at lower rates in fields Biglan classified as hard than in fields classified as soft (Malaney, 1986). But while universities tend to organize programs along the hard/soft dimension (Simpson, 2015), and teaching practices (Michel et al., 2018), career attitudes (Smart & Elton, 1982), and salaries differ among the categories (Smart & McLaughlin, 1978), whether the classifications represent clear distinctions in domains of study is less clear. There are apparent contradictions in how the system classifies fields—for example, nursing is categorized as "soft/applied/life" while dentistry is categorized as "hard/applied/non-life" (Stoecker, 1993). In addition, empirical analyses of the content of hard vs. soft sciences have failed to find clear evidence of objective differences along the continuum. For example, empirical investigations have failed to validate the claim that hard sciences have more consensus than soft sciences, despite concerted attempts (Cole, 1983). Other justifications for the hard/soft distinction describe hard sciences as incorporating advanced mathematics (Storer, 1967) and emphasizing formal models (Zuckerman & Merton, 1971). Yet exceptions to this typology are readily available—for example, advanced statistical modeling is prominent in sociology and psychology, and economics incorporates both mathematics and formal models, yet all three are typically categorized as a "soft science" (Mayer, 1980). Alternatively, others have suggested that the extent of scientific control distinguishes hard from soft sciences, with social sciences being considered soft because of the inability to control aspects of participants' experience outside the lab setting (e.g., Gutting 2012). Here too are many exceptions: The so-called hard sciences of astronomy and geoscience lack true experimental control, whereas the so-called soft science of experimental psychology relies heavily on establishing effective experimental control. The fuzziness of the hard vs. soft science categories is reflected in contradictions and inconsistencies in labeling. Biology, for example, is sometimes labeled a soft science (e.g., Previs, 2016) and other times labeled a hard science (Helmenstine, 2019). Indeed, Storer (1967) placed botany and zoology with economics in a middle-ground category of "medium-hard sciences."

While the academic literature reflects ambiguity regarding the basis for the hard/soft science distinction, the terms are widely used by members of the general public. News articles (Reinhold, 1981), opinion columns (Balko, 2018), profiles of entrepreneurs (Smith, 2020) and book reviews targeting general audiences (Maslin, 2010) have used the phrases without clarifying them to readers for decades. Google ngram, which charts the frequencies of particular word strings, indicates that the phrase "hard sciences" appeared in texts as frequently as common phrases like "science lab" or "scientific rigor." In contrast, the usage of "soft sciences" while somewhat less common, appeared about as frequently as terms such as "science news" or "science career". In other words, despite the lack of consensus from the academic literature on their meaning there is at least some evidence that these phrases are used and understood by members of the general public.

Gender representation guides labeling as "soft" or "hard." How, then, does the average person categorize scientific disciplines as hard vs. soft? We propose that colloquial use of the hard vs. soft science terms is guided by perceptions of gender representation across fields. Generally speaking, we predict that perceivers will categorize majority-men fields as hard sciences and majority-women fields as soft sciences. One reason for this alignment might be the perceiver's sense that a discipline matches the prototype of a hard science vs. a soft science. For many, hard science may closely align with prototypical views of science and of scientists. The tendency to view prototypical scientists as men (Carli et al., 2016; Rahm & Charbonneau, 1997) could lead people to perceive female-dominated STEM disciplines as less prototypic and therefore more likely to be labeled soft sciences.

Linguistic labels like soft and hard science may seem innocuous, but they potentially convey information about status. Indeed, sociological analysis has delineated a hierarchy among scientific disciplines. For example, Storer (1967) argued that soft sciences are easier to learn and study than hard sciences (see also Barnes et al., 2001; though cf. Laird et al., 2008), that findings in hard science are more objectively evaluated and thus subject to greater rigor (see also Fanelli, 2010), and that the currency with which science is rewarded (namely recognition) is thus justifiably greater and more consistent for hard science than for soft science. Such beliefs can have important downstream consequences if scientists or the public are less inclined to have respect for the field's expertise, to trust in the field's research findings, to invest resources in the field, or to pursue a career in the field. Whether colloquial use of the labels accord rigor and respect to hard sciences more than to soft sciences has yet to be tested. The labeling of female-dominated fields as soft sciences has the potential to result in decrements in respect, resources, and high-quality candidates.

Who Shows the "Women=Soft Science" Effect?

We examine three potential moderators of the tendency to equate female-dominated fields with the label of "soft science." We test (a) endorsement of gender stereotypes; (b) participant experience or intention to work in a science role; (c) participant gender.

First, we examine whether individual differences in stereotype endorsement will influence the use of soft and hard science labels. Individuals who believe that women and men differ strongly in their characteristics may be more likely to show other gender categorization effects (e.g., Ellemers, 2018). In this view, individuals who endorse gender stereotypes will associate men and women with different areas of science, showing a stronger "women=soft science" effect. We explore relationships between this association and specific dimensions of gender stereotypes: gender stereotypes about STEM interest, gender stereotypes about competence, and gender stereotypes about physical softness and fragility. Thus, personal endorsement of gender stereotypes might predict greater likelihood of using information about gender representation in applying soft vs. hard science labels.

Second, we test whether individuals within a science professional track differ in their use of gendered numeric representation for labeling hard and soft science fields. For example, more primary experience with science may lead scientists to rely less on gendered representation as a cue, focusing instead on the questions and methodologies a field employs. This prediction is consistent with evidence that stereotypes exert more impact on judgments under conditions of ambiguity or uncertainty (e.g., Kunda & Sherman-Williams, 1993; McConnell et al., 2008). Yet, scientists and non-scientists alike often hold similar stereotypes about science. For example, college students both in and outside of STEM majors perceive STEM careers as lacking opportunities for collaboration or prosociality (Diekman, Joshi, & Benson-Greenwald, 2020). Thus, scientists as well as nonscientists may use numerical representation of women to determine categorization of fields.

Third, we examine whether perceiver gender influences the extent to which gender representation is used as a signal for field categorization. Evidence suggests, though, that both men and women endorse gender stereotypes (e.g. Eagly et al., 2020). Similarly, we anticipated that both men and women might use gender representation as a heuristic for categorizing soft vs. hard sciences. In all reported studies, we tested moderation by gender. For simplicity, we report only analyses in which significant gender differences were identified.

Overview of Current Research

In the present research, we test the hypothesis that the numerical representation of women in a scientific field influences whether the field is categorized as a soft science vs. a hard science. Study 1 documents the core effect: Manipulating perceptions of a scientific field as majority-men vs. majority-women affects the assignment of hard vs. soft science labels. In Study 2, we show that this effect is associated with stereotypes about women's STEM competence. In Studies 3A and 3B, we provide evidence that the labels of soft vs. hard science have crucial consequences for evaluation of the field, with the label of soft leading to devaluation and lower funding priority. We also demonstrate that this association works in the reverse direction: Fields that are labeled as "soft sciences" are expected to have higher representation of women.

Ethical Oversight. All procedures of all studies were approved by the IRB at first author's institution.

Study 1

Study 1 had two primary aims. First, it served as an initial test of whether manipulating perceptions of women's representation in a given field influences categorization as a soft or hard science. Second, it explored whether this "women = soft sciences" effect depended on the actual numerical representation of women in a field. We hypothesized that, regardless of the true representation of women in a field, perceptions that a field is dominated by women would increase the likelihood of it being labeled as a "soft science."

Method

Participants. One hundred and sixty-seven residents of the United States, United Kingdom, and Canada were recruited from Amazon's Mechanical Turk. Participants received $0.50 for completing the study. Of these, 56 identified as women, 108 identified as men, and 3 declined to state their gender. Ages ranged from 20 to 74 (M = 34.5, SD = 10.94). Only 23 participants were currently students. A total of 60 participants reported that they worked in science, either previously, currently, or planned to in the future.

Study 1 followed a 2 (Numerical representation: majority-men vs. majority-women) × 2 (a priori categorization: social vs. engineering/physical science) within-subjects design. A pilot study with a similar design revealed initial support for our hypothesis, and obtained an effect size of f = .19 for the predicted effect of numerical representation (see Online Supplement for full details.) We conducted a sensitivity power analysis (n = 167, alpha = .05, power = .80) for a 2 (Numerical representation: majority-men vs. majority-women) × 2 (a priori categorization: social vs. engineering/physical science) repeated measures ANOVA with 4 measurements (1 for each level of each factor). Such a repeated measures ANOVA could detect could small effects of at least f = 0.09 or larger.

Procedure. Participants were asked to read descriptions of eight different scientific fields that included information about gender representation. Fields were selected to include a range of fields that varied in their actual gender representation (see Online Supplement for more details on selection process); the key experimental variable was the presented gender representation (majority-women or majority-men). To ensure a range of scientific fields, we included four social sciences (economics, psychology, political science, and sociology), and four engineering/physical sciences (computer science, biomedical science, earth science, and civil engineering). Similarly, we included social science and engineering/physical science fields that varied in their actual gender representation: Six fields were actually majority-men (economics, computer science, political science, earth science, and civil engineering), and three fields were actually majority-women (biomedical science, psychology, and sociology), determined by degrees conferred at the undergraduate and graduate level (National Center for Education Statistics, 2017).

Participants read brief descriptions of fields that included ostensible numbers of women and men in each field, without the discipline's name. The key manipulation of gender majority occurred through information about the relative numbers of women and men in each field. To understand the power of perceived representation independent of topic, for each field, half of participants saw numbers indicating that the majority of people in the field were women, and half of participants saw information indicating the majority of people in the field were men. For example, for the field of earth science, participants in the majority men information condition read as follows:

> The study of the earth or one or more of its parts, i.e. ocean, rock, atmosphere. There are 66,000 men in this field compared to 15,000 women.

In contrast, the majority women information condition read as follows:

> The study of the earth or one or more of its parts, i.e. ocean, rock, atmosphere. There are 15,000 men in this field compared to 66,000 women.

Which fields were described as majority men vs. majority women were counterbalanced across two versions of the questionnaire.

Labeling as soft or hard science. After reading each description, participants were asked to circle to indicate whether they thought the field was a hard science or a soft science.

Responses were coded such that 0 = soft science and 1 = hard science. For each participant, we calculated indices of perceived "hardness" of field by averaging within ostensibly majority-men fields and within majority-women fields. Higher numbers reflect a greater tendency to choose the hard science category.

Results

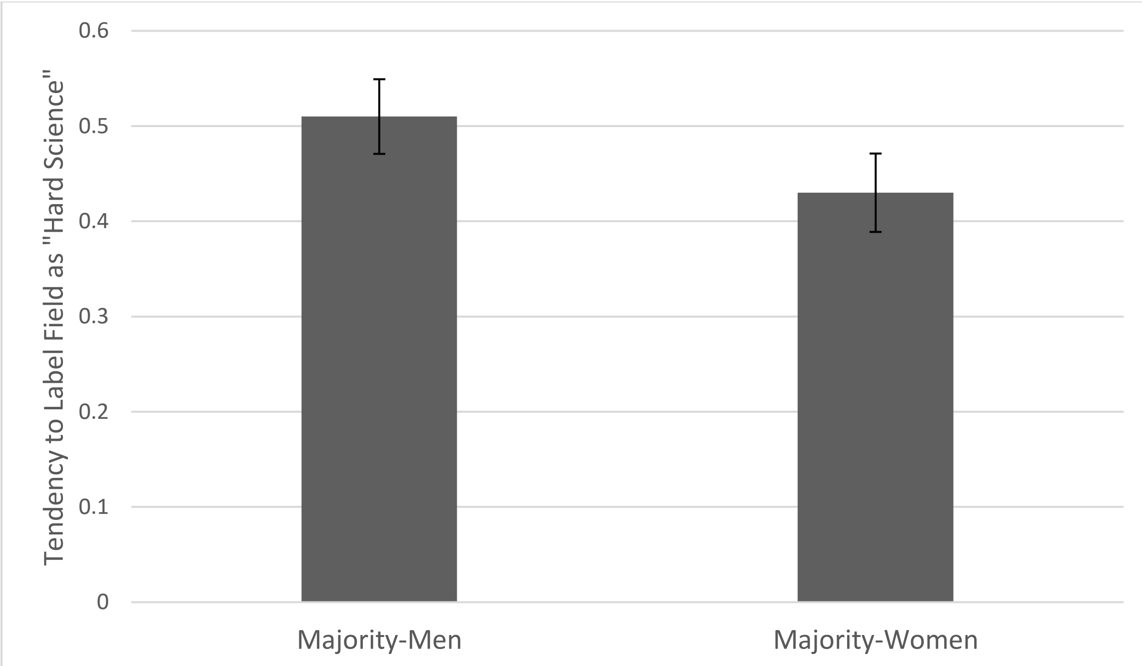
A 2 (numerical representation: majority-women vs. majority-men) ×2 (a priori categorization: social vs. engineering/physical science) repeated measures ANOVA examined effects on labeling of field as a soft vs. hard science. Engineering/physical sciences were more likely to be categorized as hard sciences (*M* = .61, *SE* = .022) than were social sciences (*M* = .33, *SE* = .021), $F(1, 165) = 77.94$, $p < .001$, $f = .69$. Key to our hypotheses, a significant main effect of numerical representation emerged, $F(1, 165) = 6.70$, $p = .011$, $f = .20$. Describing fields as majority-women led participants to categorize them as hard sciences less often (*M* = .44, *SE* = .021) than describing fields as majority-men (*M* = .51, *SE* = .020; see Figure 1). This "women = soft science" effect was not moderated by a priori categorization[1], $F(1,165) = .00$, $p = 1.00$, $f = .00$)[2].

Figure 1. Tendency to Label Field as a Hard Science Depends on Presented Numerical Representation.

---

[1] For clarity, we chose to model our binomial dependent variable using linear probability models (Gomila, 2020; Hellevik, 2007). Means thus represent the probability of categorizing a field as a hard science. To ensure that these results were not due to improper modeling of the binomial dependent variable, we additionally tested a related generalized linear multi-level model with a logit link function, using specific field (e.g. psychology, earth science, etc.) and numerical representation as fixed factors, with a random intercept for participant. We additionally controlled for actual gender representation in the field, operationalized as the percent of doctoral degrees granted to women in the past year (Snyder, de Brey, & Dillow, 2019). In addition to expected differences between specific disciplines, there was a significant effect of depicted numerical representation such that fields described as majority women were less likely to be labeled as hard sciences ($B = -.35$, $SE = .12$, $Z = 2.97$, $p = .003$.) In addition, there was a marginally significant effect of actual gender representation ($B = .71$, $SE = .40$, $Z = 1.79$, $p = .073$) such that controlling for all other variables, greater participation of women predicted greater likelihood of the field being categorized as a hard science.

[2] Additional analyses exploring moderation of the key effects by participant science role and participant gender are reported in the Online Supplement.



Note. Error bars represent 95% Confidence Intervals.

Discussion

As predicted, Study 1 found that perceived gender representation affected the tendency to label fields as hard or soft sciences. In both studies, participants tended to label fields described as majority-women as "soft sciences" and tended to label fields described as majority-men as hard sciences. This effect did not differ depending on whether the field was a social science or engineering/physical science. In other words, the "women = soft science" bias was applied equally to fields regardless of the subject the field studied.

Study 2

Study 2 investigated three potential stereotypes underlying the "women = soft science" bias. Although some part of this association might align with a "kernel of truth" that women participate more in social sciences than in engineering/physical sciences, the findings from Study 1 show that the association of women with soft science occurs across different fields, and thus the nature of the work itself is not sufficient to explain the association. We thus turn to testing potential psychological explanations that might exacerbate or mitigate the “women=soft science” association. We test different forms of gender stereotyping that might underlie the “women=soft science” bias. One possibility is that specific dimensions of gender stereotyping associated with “hard” or “soft” meanings underlie the bias. First people may judge fields with more women in them as being “soft science” based on the belief that women are less capable at science and math and thus less engaged in “hard” or rigorous sciences. Both academics and non-academics believe that certain domains of science (e.g., physics) require greater brilliance or innate ability than others (e.g., psychology), and fields thought to require brilliance are more male-dominated in their actual composition (Storage et al., 2016; Meyer et al., 2015). People who believe that women and men differ in STEM competence may use the presence of women vs. men in a field as an indicator of the field’s rigor, and label male-dominated fields as “hard sciences.” A different stereotyping explanation focuses instead of gender stereotypes about physical characteristics and connotations of physical softness: Here, the label of “soft science” could be associated with fields dominated by women simply due to semantic associations between softness and femininity (Slepian, Weisbuch, Rule, & Ambady, 2011). Among people who strongly associate women with physical softness (vs. hardness), perceiving more women in a field might semantically prime the word “soft”, increasing the likelihood that fields dominated by women are labeled as “soft sciences.”

A different possibility is that the “women=soft science” bias is rooted in a general tendency to categorize by gender; if this is true, more general stereotyping dimensions would be associated with the bias. To investigate this, we assess stereotypes about the different motivations and characteristics of men and women. The demarcation of “hard” and “soft” sciences might then align with utilizing subtypes of science (i.e. “soft science” as a subtype of the category “science”) to retain (rather than adapt) endorsed stereotypes of fields (Maurer et al., 1995). We assessed broader gender stereotyping in line with the Stereotype Content Model (Fiske et al, 2002). Women are generally perceived as lower in competence than warmth , whereas men are generally perceived as higher in competence than warmth (Eckes, 2002). Consistent with evidence that people’s prototype image of a scientist is more masculine than feminine, stereotypes of scientists using the Stereotype Competence Model’s dimensions suggest scientists are viewed as more similar to the typical man—high in competence, and low-to-moderate in warmth (Fiske & Dupree, 2014). Given the centrality of competence-related traits to perceptions of science, we anticipated that just as gender stereotypes about STEM-related abilities might lead people to label fields with majority-men as hard sciences and fields with majority women as soft science, so too might gender stereotypes of competence predict “women = soft sciences” responding. Similarly, gender stereotypes that men more than women are interested in STEM might lead to greater tendency to associate men with hard sciences and women with soft sciences. Although gender differences in STEM interest are a focus of a great deal of research, less research documents gender stereotypes about interest compared to stereotypes about ability (see Master, 2021).

Further, we investigate whether any relationships between stereotyping endorsement and the “women=soft science” bias are moderated by perceivers holding science roles. Stereotypes are more likely to influence judgments in cases where individuals have less knowledge or information is ambiguous (e.g., Kunda & Sherman-Willaims, 1983; McConnell et al., 2008). We thus anticipate that endorsement of stereotypes will be associated with greater tendency to associate women with “soft science” among participants who are not in STEM fields.

In Study 2, we tested these hypotheses by exploring relationships between “women=soft science” bias and individual differences in different forms of gender stereotyping – (1) beliefs about gender differences in competence in STEM and competence generally, (2) associations between women and physical softness, and (3) beliefs about gender differences in interest in STEM. If indeed these forms of gender stereotyping leads to the “women = soft science” association, then participants who endorse those beliefs to a higher degree should show higher levels of the association than participants who do not endorse such gender stereotypes.

Method

Participants. Sample size was determined based on a power analysis, estimating the effect size of the “women = soft science” bias as $f = .21$ based on the results of Study 1. Using G*Power, we determined that a sample size of n = 177 would achieve power = .80 to detect the effects observed in Study 1. Following Simonsohn’s (2015) recommendations for identifying sample sizes for replication plus moderation designs, we estimated that we would require twice as many observations per cell to test the hypothesis that the effect of numeric representation would be moderated by individual differences in gender stereotypes. Given our 2 (Numerical representation: majority-men vs. majority-women) × 2 (a priori categorization: social vs. engineering/physical science) within-subjects design, this would be achieved with a target n = 354. A total of 361 MTurk participants from the United States, Canada, and the United Kingdom completed the study. Participants received $1.00 for completing the study. Sixteen participants failed an attention check (“Select disagree for this item,”) and were excluded from analysis, yielding a final sample size of n = 345. Of these participants, 146 were women (198 men). Ages ranged from 19 to 70 (M = 35.76, SD = 10.97). A total of 129 participants reported past, present, or future careers in science or engineering—of these, 48 were women, and 81 were men.

We conducted sensitivity power analyses using n = 345, alpha = .05, and power = .80. alpha = .05). The primary 2 (numerical representation: majority-men vs. majority-women) × 2 (a priori categorization: social vs. engineering/physical science) repeated measures ANOVA could detect a small effect of at least $f = .06$. A 2 (science pathway: scientist, nonscientist) × 2 (participant gender: women, men) between-subjects ANOVA could detect a small effect of $f = .15$. Finally, regressions with up to 3 predictors (stereotype endorsement, science role, and their interaction) could detect coefficients of at least $f = .02$.

Procedure. Participants first completed the same field categorization task as in Study 1, categorizing eight fields as soft sciences vs. hard sciences based on a brief description of each field and information about gender representation. The numbers of men vs. women in each field were manipulated identically to Study 1.

Science role. Participants reported whether or not they held or planned to hold a job or career in a science field.

Women = soft science index. Field categorization responses were coded such that 0 = soft science, 1 = hard science. An index of "women = soft science" responding was created by averaging responses to fields described as majority-women, averaging their responses to fields described as majority-men, and subtracting the former from the latter. Values potentially ranged from -1 to 1. A score of 1 would reflect labeling every majority-women field as soft and every majority-men field as hard. In contrast, negative values would reflect counterstereotypic responding, and a score of 0 would reflect no impact of gender information.

STEM ability stereotypes. Participants responded to four items created for this study assessing endorsement of stereotypes that women lack competence in science and math ("On average, men tend to be better at math than women," "Women often struggle more than men in rigorous science and engineering courses," "Most scientific geniuses have been men," "In general, women have less of a talent for math and science than men"). Participants responded to each item on a fully-labeled, seven-point scale from *Strongly Disagree* to *Strongly Agree.* Responses were coded such that higher numbers reflected stronger endorsement of gender-STEM ability stereotypes. We averaged these to create indices for stereotypes of STEM ability (α = .90, *M* = 3.55, *SD* = 1.54)

STEM interest stereotypes. Participants responded to three items created for this study assessing endorsement of stereotypes about women's interest in science careers ("Women are less interested in careers in science and technology than men," "Women are more interested in people, while men are more interested in objects," "Men are typically more focused on achievement and careers, while women are typically more focused on their families."). Participants responded to each item on a fully-labeled, seven-point scale from *Strongly Disagree* to *Strongly Agree.* We averaged these to create an index for interest stereotypes (α = .84, *M* = 3.95, *SD* = 1.59), with higher numbers indicating stronger endorsement of gender-STEM interest stereotypes.

Trait stereotypes. Participants next rated men, women, scientists, and several filler groups (teachers, managers, engineers, lawyers, and social workers) on 13 semantic differentials (adapted from Kervyn et al., 2013). Ratings assessed perceived competence (e.g., competent/incompetent, brilliant/stupid), warmth (e.g., warm/cold, agreeable/disagreeable), and physical characteristics (e.g., soft/hard, sturdy/fragile). Participants responded on seven-point bipolar scales. For example, perceived competence included ratings from 1 (*competent*) – 7 (*incompetent*) and 1 (*brilliant*) to 7 (*stupid*) scales. Ratings were reverse scored so that higher numbers reflect higher levels of competence, warmth, or physical softness. We averaged ratings within each of these three domains to create measures of perceived competence, warmth, and physical characteristics for men, women, and scientists. We created gender stereotyping indices by subtracting ratings of women from ratings of men, with higher numbers reflecting a tendency to ascribe a given characteristic more to women than men (Competence: α = .78, *M* = -.05, *SD* = .97; Warmth: α = .72, *M* = .82, *SD* = 1.05; Physical softness/fragility: α = .79. *M* = 1.02, *SD* = 1.57;)

Results

Replication of women=soft science association. As in Study 1, we conducted a 2 (numerical representation: majority-women vs. majority-men) × 2 (a priori categorization: social vs. engineering/physical science) repeated measures ANOVA predicting categorization of the field as a soft vs. hard science[3]. Replicating Study 1, describing fields as majority-women led to less categorization as hard sciences (*M* = .488, *SE* = .013) than when describing fields as majority-men (*M* = .52, *SE* = .013; $F(1, 344) = 4.47$, $p = .035$, $f = .11$. This effect was not moderated by a priori categorization as engineering/physical sciences or. social sciences, $F(1, 344) = .041$, $p = .84$, $f = .00$.

To simplify subsequent analyses, we combined categorization of fields described as majority-women vs. fields described as majority-men into a single index of "women = soft science" bias, as previously detailed. Positive numbers represent a tendency to label fields described as majority-men as hard sciences, and fields described as majority-women as soft sciences, whereas negative numbers represent counterstereotypic labeling. Comparing this index to a value zero with a one-sample t-test is computationally equivalent to testing the effect of

[3] Including participant gender and science role as additional factors did not change the pattern or significance of the effects of interest.

numerical representation in the previously described repeated measures ANOVA (M = .039, SE = .019, t(344) = 2.11, p = .035, Cohen's d = .11)

Stereotype endorsement as predictors of women = soft science association. We first investigated whether the index of "women = soft science" bias was correlated with any of the three forms of gender stereotyping [i.e., competence (general competence or STEM ability), interest, or physical softness]. Zero-order correlations failed to show significant relationships between the women=soft science index and (1) stereotypes of competence [r(345) = .05, p = .32] or STEM ability stereotypes [r(345) = .06, p = .28], or (2) stereotypes of STEM interest [r(345) = .05, p = .37], or (3) physical softness [r(345) = .09, p = .10]. Thus, endorsing any of these stereotypes does not appear to relate to stronger labelling of majority-female fields as soft sciences. We then turned to examine whether stereotype endorsement effects occur differentially among respondents who are within a science pathway or not.

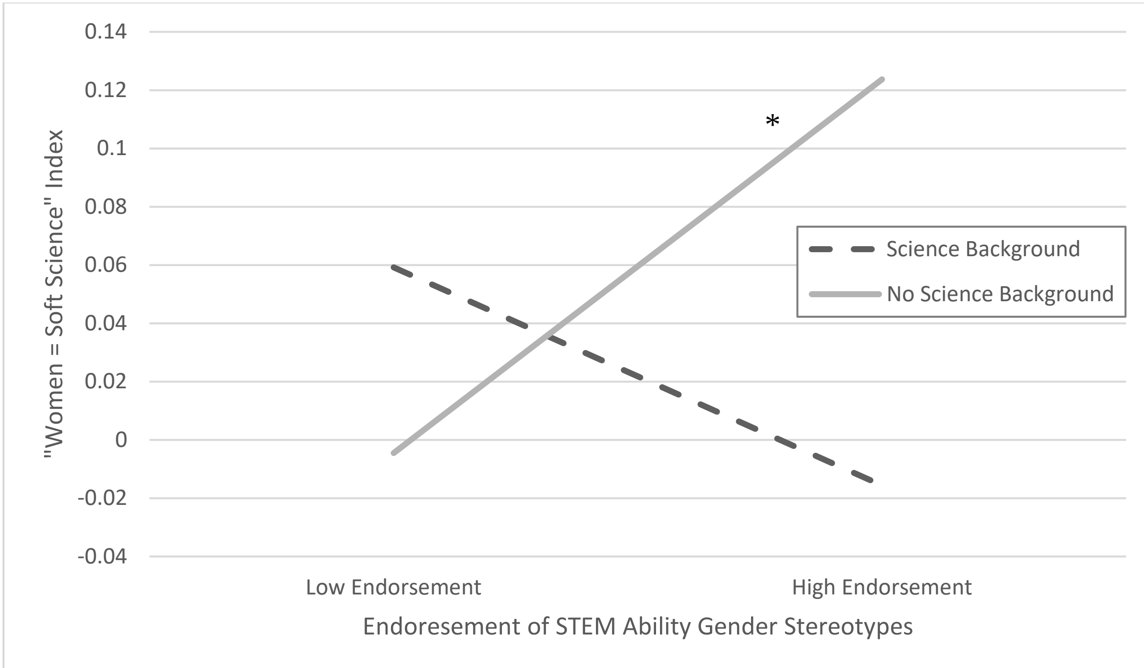
Stereotype endorsement and science roles moderate women = soft science association. We next explored whether science career experience/intention moderated the relationship between gender stereotypes and the "women = soft science" bias. Using Hayes (2017) Process Model 1, we investigated whether the relationship between each stereotyping dimension and the "women = soft science" bias was moderated by science pathway. The stereotype-to-"women = soft science" bias relationship was moderated by science pathway for competence stereotyping, both in terms of general competence ($B_{Competence \times Science\ Role}$ = -.088, SE = .040, p = .028, f = .12) and STEM-related ability ($B_{STEM\ Ability \times Science\ Role}$ = -.057, SE = .025, p = .022, f = .14).

We proceeded to probe the interaction between competence/ability stereotypes and participant science roles (see Figure 2). Among nonscientists, believing that women were less competent than men (B = .050, SE = .024, p = .035, f = .14) and believing that women had less STEM-related ability (B = .037, SE = .015, p = .018, f = .15) both predicted higher levels of "women = soft science" responding. Among scientists, this relationship did not emerge (competence stereotypes: B = -.037, SE = .032, p = .24, f = .12; STEM ability stereotypes: B = -.020, SE = .019, p = .30, f = .11)[4]. Thus, stereotypes about women's lesser competence and STEM ability predicted the tendency to label fields described as majority-women as soft sciences among non-scientists but not among scientists.

By contrast, stereotypes about STEM interest ($B_{\text{STEM interest} \times \text{Science Role}}$ = -.038, SE = .024, p = .12, f = .08), physical softness ($B_{\text{Physical softness} \times \text{Science Role}}$ = -.017, SE = .025, p = .68, f = .03), or warmth ($B_{\text{Physical softness} \times \text{Science Role}}$ = .02, SE = .036, p = .60, f = .03) did not interact with science role.

Figure 2. Gender-based STEM ability stereotypes positively predict "women = soft science" responses among non-scientists, but not among scientists.

[4] As in Study 1, we additionally modeled these results using a generalized linear mixed model with a logit link function, predicting labeling of the field as a soft science, with gender numeric information (majority men vs. majority women), participant endorsement of stereotypes about women's STEM ability, participant's STEM experience, and all 2-way and 3-way interactions, controlling for a priori categorization and actual gender representation (percent of PhDs awarded to women in the prior year), all entered as fixed factors. This analysis replicated the findings using the index—there was a numeric information × stereotype endorsement × participant science role interaction (B = -.31, SE = .011, Z = 2.77, p = .006), such that stereotype endorsement was positively associated with labeling fields described as dominated by women as soft sciences (and conversely labeling fields described as dominated by men as hard sciences) among non-scientists, while among participants with a science role, there was no relationship between stereotype endorsement and soft vs. hard science labeling. As in Study 1, a priori categorization significantly predicted categorization as hard vs. soft science (B = 2.19, SE = .10, Z = 21.28, p < .001). In addition, fields that actually had higher representation of women were more likely to be categorized as hard sciences (B = .62, SE = .29, Z = 2.13, p = .03.) We note that this surprising result only emerges when controlling for a priori categorization of the fields—when not controlling for natural/engineering vs. social science category, higher actual representation of women negatively predicted participants categorizing the field as a hard science (B = -2.27, SE = .23, Z = -9.71, p < .001.)



Note. * $p < .05$. Depicted results are of STEM ability stereotyping; a similar pattern emerged for general competence stereotyping.

Science role and gender moderate women = soft science association. A 2 (Science role: Non-scientist, Scientist) × 2 (Participant gender: women, men) between-subjects ANOVA revealed only a significant Science Role × Participant Gender interaction, $F(1, 341) = 10.46$, $p = .001$, $f = .18$. Among non-scientists, men showed higher levels of the "women = soft science" association ($M = .09$, $SE = .031$) than women ($M = .00$, $SE = .034$; $F(1, 341) = 3.72$, $p = .054$, $f = .11$). Among scientists, however, women showed higher levels of this association ($M = .12$, SE = .049) than did men ($M = -.037$, $SE = .038$; $F(1, 341) = 6.73$, $p = .01$, $f = .14$. The main effects of science role and participant gender were nonsignificant: science role: $F(1, 341) = .003$, $p = .96$, $f = .00$; participant gender: $F(1, 341) = .82$, $p = .36$, $f = .04$. This pattern of moderation was unexpected, and not observed in other studies, but may reflect that women in science are sensitized to the association between women and "soft science."

Discussion

Replicating Study 1, Study 2 provides additional evidence that labeling of soft sciences vs. hard sciences is influenced by the perceived gender representation in the field. Study 2 also demonstrates that for nonscientists, negative stereotypes about women's relative competence and ability in STEM predict the "women = soft science" bias. This pattern is consistent with the hypothesis that stereotyping women as less capable, both in and outside of STEM, is associated with labeling ostensibly majority-women fields as soft sciences. Yet the tendency to affix a particular label matters because there are consequences to that labeling. We investigate the consequences of the soft science label in the next two studies.

Studies 3A and 3B

In these studies, we test the downstream consequences of labeling fields as "soft science." In Study 3A, we measured participants' subjective evaluation of hard sciences vs. soft sciences. In Study 3B, we tested the hypothesis that describing a field as a soft science leads to devaluation. Finally, to evaluate the robustness of the "women = soft science" association, we tested whether labeling a field as a soft science led people to assume the field had a higher representation of women.

Study 3A

Method.

All pre-registration materials are available at *https://osf.io/tsj8f*

Participants. We aimed to recruit 220 residents of the United States from Amazon's Mechanical Turk to participate in a study on perceptions of science. Participants received $0.20 for completing the study. Although 220 responses were requested, a total of 222 respondents answered questions relevant to the present analyses (115 women). The study consisted of a 2 condition (hard sciences vs. soft sciences) within-subjects design. Ages ranged from 20 to 69 (M = 36.28, SD = 11.20). A total of 76 participants reported that they currently worked, had previously worked, or planned to work in science, whereas 144 reported that they had never worked in science and did not plan to in the future.

We conducted sensitivity power analyses for all effects of interest (n = 222, alpha = .05, power = .80). Our core analysis (a two-tailed paired samples t-test; hard sciences vs. soft sciences) was powered to detect a small effect size of at least d = 0.19. Moderation analyses using mixed model ANOVAs with 2 groups and 2 measures [2 (Participant gender: men, women) × 2 (Field: Hard sciences, soft sciences) and 2 (Science role: scientists, non-scientists) × 2 (Field: Hard sciences, soft sciences)] were powered to detect a small effect of at least $f = .09$. Finally, regressions with up to 2 predictors (hard science evaluations, soft science evaluations) could detect coefficients of at least $f = .04$.

Procedure. Participants were asked to provide their general impressions of hard and soft sciences. No other information about the fields was provided; participants were simply instructed: “Now we'd like you to think about fields that are known as hard [soft] sciences. Please answer the following questions while thinking about hard [soft] sciences”. All participants rated the hard and soft science categories on the same items, in counterbalanced order. Participants were asked to answer the following questions: (1) how important it is for the federal government to fund research this field (*not at all important* to *extremely important)*, (2) how difficult coursework in this field is (*extremely easy* to *extremely difficult)*, (3) the extent to which they would listen to the advice of experts in this field (*never* to *always)*, (4) how reliable they thought research in this field was (*extremely unreliable* to *extremely reliable)*, (5) and how prestigious the field was (*not at all prestigious* to *highly prestigious)*. All responses were made on five-point, fully-labeled scales.

Finally, participants were given information about allocation of the National Science Foundation's budget. Based on data from 2016, we estimated that around 12% of the NSF's funding for extramural research is allocated to social, behavioral, and economics sciences. We informed participants that 12% of the budget was typically allocated to soft sciences, and 88% to hard sciences. We then asked participants how they thought funding should be allocated. Participants responded by dragging sliders to indicate the percentage to be allocated to each general area, with the requirement that the two percentages had to add to 100%.

Table 2. Evaluations of Soft and Hard Sciences

Dimension of Evaluation	**Hard Sciences**	**Soft Sciences**	***t***	***df***	***d***
	M (SD)	*M (SD)*			
Importance of federal funding	4.00 (.94)	3.44 (1.05)	8.17*	221	.55
Difficulty of coursework	4.12 (.98)	3.59 (.91)	7.58*	220	.51
Taking advice of experts	3.98 (.83)	3.58 (.90)	6.77*	218	.46
Reliability of research	4.14 (.68)	3.71 (.88)	7.91*	218	.53
Prestige	4.09 (.91)	3.45 (.97)	10.04*	221	.67
Percentage of funding	64.18 (16.26)	35.82 (16.26)	13.00*	221	.87

Note. Ratings were made on scales ranging from 1 to 7; * $p < .05$

Results.

Are Soft Sciences Devalued? Paired-samples *t*-tests indicated that across all five dimensions, soft sciences were devalued relative to hard sciences (see Table 2). Additionally, a significant Field Type × Participant Gender interaction emerged, $F(1, 209) = 10.04$, $p = .002$,

$d = .49$. Specifically, both women and men devalued soft sciences significantly more than hard sciences, but this effect was larger among men [*women*: $F(1, 209) = 41.19$, $p < .001$, $d = .91$; *men*: $F(1, 209) = 109.84$, $p < .001$, $d = 1.46$]. Women evaluated soft sciences more positively ($M = 3.62$, $SE = .07$) than did men ($M = 3.38$, $SD = .07$; $F(1, 209) = 5.88$, $d = .33$), even as women and men did not differ in their evaluation of hard sciences (*women*: $M = 4.02$, $SE = .06$; *men*: $M = 4.06$, $SE = .078$; $F(1, 209) = .22$, $p = .64$, $d = .03$).

Finally, the devaluation of soft sciences was not moderated by science role, $F(1, 218) = .67$, $p = .41 d = .11$. The only effect of science role was an unexpected main effect: Scientists evaluated both hard and soft sciences more negatively ($M = 3.65$, $SE = .07$) than did non-scientists ($M = 3.90$, $SE = .05$), $F(1, 218) = 9.19$, $p = .003$, $d = .41$.

Are Soft Sciences Allocated Less Funding? Given that responses to the funding allocation questions were not independent (i.e., they were required to add to 100%), we used Friedman's test to identify whether participants were more likely to allocate greater funding to hard sciences than soft sciences. As predicted, participants allocated less funding to soft sciences ($M = 35.82$, SD = 16.26) than to hard sciences ($M = 64.18$, SD = 16.26; $\chi^2(1) = 103.91$, $p < .001$, $d = .87$).

It is possible, however, that participants perceive hard science research as more expensive, due to needs for equipment, personnel, and other materials, and that these expense expectations explain discrepancies in funding allocation for hard vs. soft sciences. If so, individual differences in evaluation of hard vs. soft sciences would not predict funding allocation above and beyond the hard/soft evaluation. Yet if funding allocations are related to attitudes, we would expect this relationship to emerge. Thus, in exploratory analyses, we further probed this result by testing whether evaluations of hard vs. soft sciences are associated with differences in funding allocation. We ran a regression using average evaluations of hard sciences and average evaluations of soft sciences to predict the discrepancy in funding allocation to hard vs. soft science. Both were significant predictors of funding allocation—participants' favorable evaluation of hard sciences positively predicted their funding allocation to hard vs. soft science (B = 22.76, SE =3.82, t(219) = 5.96, p < .001), and participants' favorable evaluation of soft sciences negatively predicted their allocation to hard vs. soft sciences (B = -20.56, SE = 3.49, t(219) = 5.89, p < .001). The two predictors combined explained 17% of the variance in funding allocation. This analysis provides evidence that differential evaluation enters into funding allocation.

Discussion. Consistent with our hypothesis, participants asked to consider the category of "soft sciences" reported more negative impressions than when they considered the category of "hard sciences." This method has the advantage of assessing beliefs based solely on the labels of hard or soft science; however, this method also leaves open the question of which scientific disciplines participants thought of when rating each category. Our contention is that labeling a field as soft science contributes to devaluation, not simply that some members of the public value physics more highly than political science. In Study 3B, we sought to demonstrate that the mere label of hard vs. soft science causally influences judgment of specific fields.

Study 3B

All pre-registered materials available at https://osf.io/evdny

Method.

***Participants*.** Study 3A identified an average effect size of d = .60 for the effect of soft vs. hard science label on evaluation of the field; however, we anticipated that the effect of applying the label to a *specific* field would be smaller. Thus, we conducted a power analysis to determine the sample size needed to detect an effect of d = .30 with power of .95 with a 2 (target field: computer science vs. psychology)× 2 (label: hard science vs. soft science) between-subjects design. Using G*Power, we determined the necessary sample size was n = 580. We recruited 580 residents of the United States using MTurk. Participants received $0.40 for completing the study. A total of n = 585 participants completed the relevant measures. Of these, 253 identified as women, 330 as men, and one as non-binary. Age ranged from 19 to 82 (M = 36.43, SD = 11.36). A total of 229 reported holding or intending to hold a science role in the future, while 353 reported no science role or intention.

We conducted sensitivity power analyses for all effects of interest (n = 585, alpha = .05, power = .80). The core 2 (target field: computer science vs. psychology) × 2 (label: hard science vs. soft science) between-subjects ANOVA was powered to detect an effect size of at least $f =$ .12. A 2 (Science role: scientist, non-scientist) × 2 (target field) × 2 (label) between-subjects ANOVA was powered to detect an effect size of at least f = 0.19.

Procedure. The key variables were manipulated in a 2 (target field: Computer Science, Psychology) × 2 (label: hard science, soft science) between-subjects design. Participants were randomly assigned to read about and evaluate a single scientific field (Computer Science vs. Psychology) that was labeled as either a hard or soft science. The full text of these manipulations is given in Table 3.

Table 3. Experimental Stimuli: Manipulations of Field and Hard/Soft Label

	Label	
Field	***Soft***	***Hard***
Computer Science	While many people think of computer science as a hard science, it is actually more accurately understood as a soft	Many people classify Computer Science as a hard science, and that's definitely true. Much of the work in computer science is

	science. Much of the work in computer science is focused on humans' interactions with computers, and designing computer systems around the ways that people think and act. Think about all the ways that computers impact your daily life, from programs you use for work, to tools for communication, to the ways that you just use the to relax and have fun. Computer scientists need to understand the way that people interact with their machines in order to make computers work for you.	focused on electrical engineering, and designing computer systems to apply our knowledge of physics and chemistry. Think about all the ways that computers impact your daily life, from programs you use for work, to tools for communication, to the ways that you just use the computer to relax and have fun. Computer scientists need to think beyond to the mechanics of how machines work in order to make computers work for you.
Psychology	Many people classify Psychology as a soft science, and that's definitely true. Much of the work in psychology requires an intuitive understanding of human behavior and emotions, and using that understanding to explain behavior. Think about all the ways that people around you act in your daily life, from the people you work with, to your neighbors, to the people you relax and have fun with. Psychologists rely on intuitions about emotions and the mind to explain how people think, act, and feel.	While many people classify Psychology as a soft science, it's actually more accurately understood as a hard science. Much of the work in psychology requires use of the scientific method to conduct carefully controlled studies, and complex mathematical analysis to understand the results. Think about all the ways that people around you act in your daily life, from the people you work with, to your neighbors, to the people you relax and have fun with. Rather than simply relying on intuitions, psychologists conduct systematic research and focus on quantifying the mind and emotions to explain how people think, act, and feel.

Measures.

Evaluation. Participants evaluated the field on the same items as Study 3 (importance of funding, difficulty of coursework, taking advice of experts, reliability of research, and prestige). These items were averaged to create an overall index of evaluation (α = .73).

Estimate of Gender Representation. Participants estimated what percentage of people in the field were men vs. women, using sliding scales varying from 0-100, with the total of the two summing to 100%.

Results.

Field Evaluations. Evaluations of target field were submitted to a 2 (field: computer science vs. psychology) × 2 (label: hard vs. soft science) between-subjects ANOVA. As predicted, a significant main effect of hard vs. soft science label emerged, such that fields framed as soft sciences were evaluated more negatively (*M* = 3.66, *SE* = .038) than when they were framed as hard sciences (*M* = 3.82, *SE* = .037; $F(1, 581) = 9.72$, $p = .002$, $f = .13$). Additionally, psychology (*M* = 3.61, *SE* = .037) was evaluated more negatively than computer science (*M* = 3.87, SE = .03; $F(1, 581) = 23.26$, $p < .001$, $f = .20$.). Finally, an unexpected, albeit marginal, Field × Label interaction emerged, $F(1, 581) = 3.34$, $p = .068$, $f = .08$, such that the effect of hard vs. soft science label more strongly influenced evaluations of computer science than psychology.

***Estimate of Gender Representation*.** Perceptions of gender representation were submitted to a 2 (field: computer science vs. psychology) × 2 (label: hard vs. soft science) between-subjects ANOVA. As predicted, participants perceived greater representation of women in fields labeled as soft sciences (*M* = 43.28%, *SE* = .81) than hard sciences (*M* = 39.74%, *SE* = .79; $F(1, 581) = 9.83$, $p = .002$, $f = .13$). Moreover, participants believed women to represent a higher percentage of psychologists (*M* = 49.93%, *SE* = .79) than computer scientists (*M* = 33.09, *SE* = .81; $F(1, 581) = 221.97$, $p < .001$, $f = .62$). The impact of labeling the field as a hard science did not significantly differ by field, $F(1, 581) = .81$, $p = .37$, $f = .03$.

Participant Science Role. Scientists evaluated fields more negatively (*M* = 3.67, *SE* = .042) than nonscientists [*M* = 3.78, *SE* = .034; $F(1, 574) = 4.12$, $p = .043$ $f = .08$], but science role did not moderate either the effect of soft vs. hard science framing, $F(1, 574) = 1.66$, $p = .20$, $f = .05$, nor the effect of field, $F(1, 574) = .54$, $p = .46$, $f = .03$. Moreover science role did not influence estimates of gender representation, $F(1, 574) = 1.52$, $p = .22$, $f = .05$, nor did it moderate the effect of soft vs. hard science framing, $F(1, 574) = .04$, $p = .84$, $f = .00$, nor the effect of field, $F(1, 574) = .74$, $p = .39$, $f = .03$. When controlling for participants' science role, the effect of hard vs. soft science label on both evaluations and perceptions of gender representation remained significant.

Discussion

Consistent with the results of Study 3A, the label of soft science led to subjective devaluation of scientific fields. Moreover, Study 3B provides evidence that the association between women and soft science also operates in the opposite direction—labeling fields as soft sciences led participants to believe that there were more women in the field.

General Discussion

Collectively, these studies demonstrate a link between women's numerical representation and the label of soft science. Studies 1 and 2 show that experimentally manipulating perceptions of gender representation influences whether both lay people and science students/professionals will categorize a field as a soft science. Study 2 demonstrated that among non-scientists, this association between women and soft science is stronger among people who perceive women to be less competent and to have less STEM-related ability. Studies 3A and 3B show that these labels have consequences—that the general category of soft sciences is perceived as less rigorous, trustworthy, and worthy of funding than hard sciences, and that describing a field as a soft science leads it to be devalued compared to describing it as a hard science. Finally, Study 3B demonstrated that the strength of the association between soft sciences and women's numerical representation by demonstrating that the cues also operate in reverse: describing a field as a soft science leads to the perception that a greater percentage of its professionals are women.

The present work seeks to understand the perceptions and consequences of these labels of scientific fields, rather than to address genuine differences in fields' subject matter, methods, or general practices that might cleave them into the hard and soft categories, as explored by historians and sociologists of science (e.g., Storer, 1967). However, it is interesting to note that many of the fields that are inconsistently categorized in previous research are often fields that do not cleanly fit into a the "women = soft science" category. For example, biological sciences have gained relative gender parity in recent years, and are categorized as hard sciences in some formulations (e.g., Biglan, 1973), and soft sciences in others (Camp et al., 2009). Similarly, economics is a social science, which would typically lead it to be categorized as a soft science but has fewer women that most other social sciences (Goldin, 2013). Interestingly, Storer (1967) chose to classify economics as a "medium hard" science, along with botany and zoology. While there likely are measurable, objective differences among fields categorized as hard sciences vs. soft sciences that are unrelated to gender representation, the fact that these edge cases are often fields whose gender representation runs counter to their typical category suggests that academicians who construct these classification systems (or the faculty they survey to create them, in the case of the Biglan classifications; 1973) may, too, be influenced by the "women = soft science" association. An implication may be that labeling fields with high participation of women as soft sciences could lead stereotyping of women's capacity and interest in STEM if these fields are devalued relative to "hard" sciences.

Crucially, the present studies show that whether a field is labeled as a hard science or a soft science has important consequences. Studies 3A and 3B showed that soft sciences are perceived as less difficult, less reliable, less prestigious, and less worthy of funding than hard sciences, and that people are less likely to listen to experts in soft science compared to hard science. Study 3B went further by showing that describing a specific field as a soft science (vs. a hard science) led to similar devaluation of the field. This pattern suggests the disturbing possibility that contrary to the claim that increasing women's representation in STEM will diminish negative gender stereotypes, perceivers will instead devalue fields that women enter at high rates (e.g., see England et al., 2007). Unless gender stereotypes about women's competence are challenged, scientists working in fields with increasing representation of women may face negative consequences, such as less respect and resources in the eyes of the public, funding agencies, or trainees considering the field. As disturbing as this possibility may be, understanding the psychological processes at play is essential in challenging these biases.

In Study 3A, we found that participants thought less federal funding should be allocated to soft sciences than to hard sciences, which we argue reflects a devaluation of fields labeled as soft sciences. However, it is plausible that participants responses did not reflect their valuation of hard vs. soft sciences, but rather their assumptions about the relative expense of hard vs. soft sciences. It may be true that engineering/physical sciences are more resource intensive, which may factor into funding allocations in this study – yet our analyses also show that the differential evaluation of hard vs. soft sciences is related to recommended funding allocations. What a field is perceived to require may in part be driven by cultural norms and expectations, and association of maleness with certain scientific fields can imbue those fields with greater status and in turn greater resources.

Overall, these findings are consistent with broader work on the "feminization" of academic disciplines, which similarly suggests that women's numerical representation in a field can impact evaluations of and desirability of the discipline (England et al., 2007). Collectively, this work presents a troubling complication of the assumption that stereotypes about women's competence in STEM can be eradicated by increasing women's representation in STEM. Rather, this research suggests that as women enter specific STEM fields in higher numbers, the public's impressions of those fields may shift to accommodate gender stereotypes.

Limitations and Future Directions

Although these findings clearly demonstrate a link between gender representation and the labeling of fields as hard or soft sciences, as well as the consequences of those labels, additional questions remain, and future research can address the limitations of the current work.

Science role. These findings detected differences between how scientists and non-scientists use the hard and soft science labels. Yet we acknowledge that individuals with experience or intent to work in science were a minority, and thus cautious interpretation is warranted. For example, these participants are not in positions where they make funding recommendations, and so the funding allocation is best regarded as indexing public opinion about funding allocation than as a proxy for actual funding decisions. Yet the patterns that emerged cohere with existing research and suggest avenues for future research. For example, Study 2 identified differences between groups in the predictors of "women = soft science" responding. For nonscientists, the women=soft science response correlated with beliefs that women are less competent in general and in STEM specifically; for scientists, these relationships did not emerge. Similarly, while nonscientist women exhibited less "women = soft science" bias than nonscientist men, the pattern was reversed among scientists.

These results may suggest that different mechanisms may guide the application of soft and hard sciences labels among people with more or less experience with the scientist role. Stereotyping effects on judgment tend to be stronger when information is ambiguous (e.g., McConnell et al., 2008), and thus future research could directly manipulate experience with a particular scientific field to see if the effects of competence stereotypes on “women=soft science” bias are mitigated. Another route would be to explore the different implications of the science role by gender: It may be that women in science particularly show the “women=soft science” association because they are sensitized to the gender disparities that emerge across different fields of science. Women in science might be exhibiting something akin to the “Queen Bee” response to being a gender minority in a male-dominated field (Ellemers et al., 2004): Perhaps this effect is driven by women in science fields with more men differentiating their fields from those with more women by labeling such fields as “soft sciences.” To investigate this idea, systematic study of the beliefs of women in fields with different levels of female representation is warranted.

Subtyping mechanism and consequences? We speculate that the process by which gender representation influences hard and soft science labels may be akin to subtyping in response to stereotype disconfirming individuals. In this view, “soft sciences” may be a subtype of the larger category of science. Subtyping in social perception occurs when individuals who disconfirm a group stereotype are mentally clustered together and excluded from the stereotyped group (Richards & Hewstone, 2001). Individuals with higher prejudice against a particular group are more likely to subtype positive group members (Riek, Mania, & Gaertner, 2013), and subtyping appears to maintain stereotypes (Richards & Hewstone, 2001). Just as perceivers may subtype counter-stereotypic individuals, so too may they subtype counter-stereotypic scientific disciplines by clustering them together and mentally excluding them from the category of “science”. People who believe that women lack competence for math and science would tend to perceive scientific fields with high rates of participation by women as counterstereotypic. Subtyping these fields—by categorizing them as “not real sciences”, or perhaps “soft sciences”, could allow the individual to maintain stereotypes about STEM fields in the face of such stereotype disconfirming evidence.

Consistent with this idea, Study 2 demonstrated that the “women = soft science” association was highest among participants who stereotyped women as being less competent generally and in STEM, in a similar fashion as subtyping for racial outgroup members is more common among participants with higher levels of prejudice (Riek et al., 2013). Yet, the present studies do not offer a direct investigation of subtyping effects. Future research could test whether the use of hard and soft science labels maintains beliefs about the broader requirements and attributes of STEM fields.

Manipulation checks. Finally, we chose not to include manipulation checks in Studies 1 and 2 to assess whether participants believed the numerical information about gender representation provided, nor did we include a manipulation check in Study 3b to assess whether participants accepted the framing of psychology/computer science as soft vs. hard sciences. We elected not to use manipulation checks across studies because we believed they might undermine the efficacy of manipulations. Specifically, we were concerned that asking participants whether they believed the description of the field in question (e.g., economics has more women than men, or computer science as a soft science) might encourage them to be skeptical and generate counterarguments. Indeed, the use of manipulation checks has been noted as a complicated issue given that they have the potential to influence responses to the dependent variable itself (Hauser et al., 2008). Although we felt that this design choice was important, it does have its drawbacks as we are unable to definitively demonstrate whether or not participants rejected the framing. In open-ended questions at the end of all studies, virtually no participants spontaneously mentioned suspicion regarding manipulations. That the framing impacted participants' responses as predicted provides some evidence that they believed and accepted the framing provided to them.

Conclusion

Many STEM fields have made great strides toward gender parity over the past several decades. For example, as of 2017, women make up about 49% of the life sciences workforce, up from 40% in 2003 (National Science Board, 2019). Yet in other areas, women's advancement in STEM has been stubbornly stalled, such as in the physical sciences where women's representation has held steady at around 30% since 2003. The present research highlights how stereotypes about women's STEM competency may continue to impact science even as women's participation in STEM increases. STEM fields with a high percentage of women may be evaluated through the lens of gender stereotypes, with consequences for the general public's perception of those fields. Moreover, the language used to communicate about science may be covertly carrying these gender stereotypes.

Open Science Practices

Across all studies, we report all measures, manipulations, and exclusions. In addition, we provide pre-registration materials for Studies 3A (https://osf.io/exsta) and 3B (https://osf.io/evdny). No changes were made to the pre-registered analysis plans for the primary confirmatory analyses for these studies, and all analyses described in the registered plans are reported in the article.

References

Balko, R. (2018, August 9). Opinion | Rod Rosenstein still doesn't get the problem with forensics. *Washington Post*. https://www.washingtonpost.com/news/the-watch/wp/2018/08/09/rod-rosenstein-still-doesnt-get-the-problem-with-forensics/

Barnes, L.L.B, Bull, K.S., Campbell, N.J., & Perry, K.M. (2001). Effects of academic discipline and teaching goals in predicting grading beliefs among undergraduate teaching faculty. *Research in Higher Education, 42,* 455-467. https://doi.org/10.1023/A:1011006909774

Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology, 57(3),* 195-203. https://doi.org/10.1037/h0034701

Camp, A. G., Gilleland, D. S., Pearson, C., & Putten, J. V. (2009). Differentiating between Women in Hard and Soft Science and Engineering Disciplines. *Journal of College Student Retention: Research, Theory & Practice, 11(3*), 363–384. https://doi.org/10.2190/CS.11.3.d

Carli, L. L., Alawa, L., Lee, Y., Zhao, B., & Kim, E. (2016). Stereotypes about gender and science: Women ≠ scientists. *Psychology of Women Quarterly, 40(2),* 244–260. https://journals.sagepub.com/doi/10.1177/0361684315622645

Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin, 25(4*), 413–423. https://doi.org/10.1177/0146167299025004002

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143(1),* 1–35. https://doi.org/10.1037/bul0000052

Cole, S. (1983). The hierarchy of the sciences? *American Journal of Sociology, 89(1)*, 11-139, https://doi.org/10.1086/227835

Diekman, A. B., Joshi, M. P., & Benson-Greenwald, T. M. (2020). Goal congruity theory: Navigating the social structure to fulfill goals. In B. Gawronski (Ed.), *Advances in Experimental Social Psychology* (Vol. 62, pp. 189–244). Academic Press. https://doi.org/10.1016/bs.aesp.2020.04.003

Eagly, A. H., & Koenig, A. M. (2021). The Vicious Cycle Linking Stereotypes and Social Roles. *Current Directions in Psychological Science*, *30*(4), 343–350. https://doi.org/10.1177/09637214211013775

Eagly, A. H., Nater, C., Miller, D. I., Kaufmann, M., & Sczesny, S. (2020). Gender stereotypes have changed: A cross-temporal meta-analysis of US public opinion polls from 1946 to 2018. *American Psychologist, 75(3*), 301–315. https://doi.org/10.1037/amp0000494

Eckes, T. (2002). Paternalistic and Envious Gender Stereotypes: Testing Predictions from the Stereotype Content Model. Sex Roles, 16.

Ellemers, N. (2018). Gender stereotypes. *Annual Review of Psychology*, *69*, 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Ellemers, N., van den Heuvel, H., de Gilder, D., Maass, A., & Bonvini, A. (2004). The underrepresentation of women in science: Differential commitment or the queen bee syndrome? *British Journal of Social Psychology, 43*, 1-24. doi:10.1348/0144666042037999

England, P., Allison, P., Li, S., Mark, N., Thompson, J., Budig, M.J., & Sun, H. (2007). Why are some academic fields tipping toward female? The sex composition of U.S. fields of doctoral degree receipt, 1971-2002. *Sociology of Education, 80(1)*, 23-42. https://doi.org/10.1177/003804070708000102

Fanelli, D. (2010). “Positive” results increase down the hierarchy of the sciences. *PLOS One*, https://doi.org/10.1371/journal.pone.0010068

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. https://doi.org/10.3758/BF03193146

Faul, F., Erdfelder, E., Buchner, A, & Lang, A.G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41(4)*, 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Goldin C. (2013) Notes on Women and the Undergraduate Economics Major. *CSWEP Newsletter. 2013;(Summer)* :4-6.

Gomila, R. (2020). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000920

Gutting, G. (2012). How reliable are the social sciences? *The New York Times.* https://opinionator.blogs.nytimes.com/2012/05/17/how-reliable-are-the-social-sciences/

Hauser, D. J., Ellsworth, P. C., & Gonzalez, R. (2018). Are manipulation checks necessary? *Frontiers in Psychology*, *9,* 998. https://doi.org/10.3389/fpsyg.2018.00998

Hayes, A.F. (2017) *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach (2nd Ed.)* Guilford Press.

Hellevik, O. (2007) Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity, 43(1),* 59–74. http://doi.org/10.1007/s11135-007-9077-3

Helmenstine, A.M. (2019). What is the difference between hard and soft science? *ThoughtCo.* https://www.thoughtco.com/hard-vs-soft-science-3975989

Kervyn, N., Fiske, S. T., & Yzerbyt, V. Y. (2013). Integrating the stereotype content model (warmth and competence) and the Osgood semantic differential (evaluation, potency, and activity). *European Journal of Social Psychology*, *43*(7), 673–681. https://doi.org/10.1002/ejsp.1978

Kessels, U., Rau, M., & Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology, 76(4*), 761–780. doi:10.1348/000709905X59961

Kunda, Z., & Sherman-Williams, B. (1993). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin*, *19*(1), 90–99. https://doi.org/10.1177/0146167293191010

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago.

Laird TFN, Shoup R, Kuh GD, Schwarz MJ (2008) The effects of discipline on deep approaches to student learning and college outcomes. *Research in Higher Education 49*, 469–494. https://doi.org/10.1007/s11162-008-9088-5

Malaney, G.D. (1986). Characteristics of graduate students in Biglan areas of study. *Research in Higher Education, 25*, 328-341. https://doi.org/10.1007/BF00992129

Maslin, J. (2010, August 4). Hard Science, Softened With Stories. *The New York Times*. https://www.nytimes.com/2010/08/05/books/05book.html

Master, A. (2021). Gender Stereotypes Influence Children's STEM Motivation. *Child Development Perspectives*, *15*(3), 203–210. https://doi.org/10.1111/cdep.12424

Maurer, K. L., Park, B., & Rothbart, M. (1995). Subtyping versus subgrouping processes in stereotype representation. *Journal of Personality and Social Psychology*, *69*(5), 812–824. https://doi.org/10.1037/0022-3514.69.5.812

Mayer, T. (1980). Economics as a hard science: Realistic goal or wishful thinking? *Economic Inquiry, 18(2),* 165-178. https://doi.org/10.1111/j.1465-7295.1980.tb00568.x

McConnell, A. R., Rydell, R. J., Strain, L. M., & Mackie, D. M. (2008). Forming implicit and explicit attitudes toward individuals: Social group association cues. *Journal of Personality and Social Psychology*, *94*(5), 792–807. https://doi.org/10.1037/0022-3514.94.5.792

Meyer, M., Cimpian, A., & Leslie, S.-J. (2015). Women are underrepresented in fields where success is believed to require brilliance. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00235

Michel, J.O., Campbell, C.M., & Dilsizian, K. (2018). Is STEM too hard? Using Biglan to understand academic rigor and teaching practices across disciplines. *The Journal of the Professoriate, 9(2),* 28-56.

Miller, D. I., Nolla, K. M., Eagly, A. H., & Uttal, D. H. (2018). The development of children's gender-science stereotypes: A meta-analysis of 5 decades of US Draw-a-Scientist studies. *Child Development, 89(6)*, 1943–1955. https://doi.org/10.1111/cdev.13039

Munro, G. D., & Munro, C. A. (2014). “Soft” Versus “Hard” Psychological Science: Biased Evaluations of Scientific Evidence That Threatens or Supports a Strongly Held Political Identity. *Basic and Applied Social Psychology*, *36*(6), 533–543. https://doi.org/10.1080/01973533.2014.960080

National Academy of Engineering. (2015). NAE Grand Challenges for Engineering. Retrieved from http://engineeringchallenges.org.National Science Foundation, National Center for Science and Engineering Statistics. (2019). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2019. Special Report NSF 19-304.* https://www.nsf.gov/statistics/wmpd.

National Center for Science and Engineering Statistics. 2021. *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2021.* Special Report NSF 21-321. Alexandria, VA: National Science Foundation. Available at https://ncses.nsf.gov/wmpd.

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology, 54(1),* 14–29. https://doi.org/10.1027/1618-3169.54.1.14

Previs, K. K. (2016). Gender and Race Representations of Scientists in Highlights for Children: A Content Analysis. *Science Communication, 38(3),* 303–327. https://doi.org/10.1177/1075547016642248

Rahm, J., & Charbonneau, P. (1997). Probing stereotypes through students' drawings of scientists. *American Journal of Physics, 65*(8), 774-778. https://doi.org/10.1119/1.18647

Reinhold, R. (1981, February 9). Reagan's plans on research cuts are said to aim at "soft" sciences. *The New York Times*. https://www.nytimes.com/1981/02/09/us/reagan-s-plans-on-research-cuts-are-said-to-aim-at-soft-sciences.html

Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, *5(1*), 52–73. https://doi.org/10.1207/S15327957PSPR0501_4

Riek, B. M., Mania, E. W., & Gaertner, S. L. (2013). Reverse subtyping: The effects of prejudice level on the subtyping of counterstereotypic outgroup members. *Basic and Applied Social Psychology, 35(5),* 409–417. https://doi.org/10.1080/01973533.2013.823616Simonsohn, U. (2015) No-way interactions. DOI: 10.15200/winn.142559.90552

Simpson, A. (2015). The surprising persistence of Biglan's classification scheme. *Studies in Higher Education, 42(8)* 1520-1531. https://doi.org/10.1080/03075079.2015.1111323

Slepian, M. L., Weisbuch, M., Rule, N. O., & Ambady, N. (2011). Tough and tender: Embodied categorization of gender. *Psychological Science, 22(1)*, 26–28. https://doi.org/10.1177/0956797610390388

Smart, J. C., & Elton, C. F. (1982). Validation of the Biglan model. *Research in Higher Education*, *17(3),* 213–229. https://doi.org/10.1007/BF00976699

Smart, J. C., & McLaughlin, G. W. (1978). Reward structures of academic disciplines. *Research in Higher Education, 8(1*), 39–55. https://doi.org/10.1007/BF00985855

Smith, R. A. (2020, November 15). How Chess Sparked a Career Switch to Software Development; Spurred by a daughter's question, a father gets serious about chess and dares to focus on a long-ignored love of hard science. *Wall Street Journal (Online)*. https://www.proquest.com/docview/2460600648/citation/F485577178E34CFCPQ/2

Snyder, T.D., de Brey, C., and Dillow, S.A. (2019). *Digest of Education Statistics 2017* (NCES 2018-070). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Stoecker, J.L. (1993). The Biglan classification revisited. *Research in Higher Education, 34(4),* 451-464. https://doi.org/10.1007/BF00991854

Storage, D., Horne, Z., Cimpian, A., & Leslie, S.-J. (2016). The frequency of “brilliant” and “genius” in teaching evaluations predicts the representation of women and African Americans across fields. *PLoS ONE*, *11*(3). https://doi.org/10.1371/journal.pone.0150194

Storer, N.W. (1967). The hard sciences and the soft: Some sociological observations. *Bulletin of the Medical Library Association, 55(1),* 75-84.

Zuckerman, H.A. & Merton, R.K. (1971.) Patterns of evaluation in science: institutionalization, structure and functions of the referee system. *Minerva, 9*, 66-100. https://doi.org/10.1007/BF01553188