Psychologically Inclusive Design

Cues Impact Women's Participation in STEM Education

René F. Kizilcec
Cornell University
Ithaca, NY
kizilcec@cornell.edu

Andrew J. Saltarelli
Stanford University
Stanford, CA
saltarel@stanford.edu

ABSTRACT

Visual and verbal cues can reinforce barriers to access for women in science, technology, engineering, and math (STEM) disciplines. Psychologically inclusive design is an evidence-based approach to reduce psychological barriers by strategically placing content and design cues in the environment. Two large field experiments provide estimates of the behavioral impact of psychologically inclusive cues on women's and men's enrollment behaviors in an online learning environment. First, a gender-inclusive photo and statement in an online advertisement for a STEM course increased the click-through rate among women but not men by 26% (N=209,000). Second, an inclusivity statement with a gender-inclusive course image to the enrollment page raised the proportion of women enrolling in a STEM course by up to 18% (N=63,000). These findings contribute evidence of the behavioral impact of psychologically inclusive design to the literature and yield practical implications for the presentation of STEM opportunities.

CCS CONCEPTS

• **Human-centered computing** → *Empirical studies in HCI*; • **Applied computing** → *Education*; • **Information systems** → *Online advertising*;

KEYWORDS

Education, Social Psychology, Design, Diversity, Inclusion, Equality

ACM Reference Format:
René F. Kizilcec and Andrew J. Saltarelli. 2019. Psychologically Inclusive Design: Cues Impact Women's Participation in STEM Education. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland, UK*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3290605.3300704

1 INTRODUCTION

Online learning has rapidly expanded in the last decade and today nearly one in three US college students takes at least one online course during their academic career and 15% are enrolled in entirely online programs [22]. Digital learning platforms have become ubiquitous on college campuses and many institutions see online learning as a strategy to expand enrollments and access while at the same time balancing shrinking budgets [18]. Massive open online courses have been cited as an effective way to affordably scale online learning while increasing access for traditionally marginalized students, especially in science, technology, engineering, and mathematics (STEM) disciplines [20, 64]. However, accumulating evidence suggests that online learning environments perpetuate enrollment and participation gaps in STEM courses [3, 23, 28, 34, 35]. Thus, it may be the case that, like face-to-face environments, online learning platforms contain contextual cues that activate psychological biases and make certain groups of learners (e.g., women in STEM courses) feel unwelcome. With rising demand for online higher education, especially in STEM, there is a pressing need to understand how to design online learning environments that are psychologically inclusive and welcoming for students with diverse backgrounds.

The presence of psychological barriers that contribute to educational inequalities is well-documented in the social-psychological literature [2, 59, 67]. Several studies have examined the experiences and coping strategies of women in STEM settings [9, 37, 55, 58, 69] and how attributes of STEM environments shape their perceptions about what it would be like in this environment. Cues in the environment, such as verbal, visual, or behavioral artifacts, can make individuals vulnerable to social identity threat—concerns about being seen as less capable because of one's social identity [60]—by signaling the relevance of particular social identities in a given context. These cues can communicate group stereotypes and cause individuals to consider disengaging from an academic domain [45].

How environmental cues raise identity-based concerns and influence anticipated belonging has been studied in a number of settings: gender-stereotypic TV commercials [16, 17], promotional videos for STEM events [44], physical and virtual-reality computer science (CS) classrooms [9, 10, 40], and the website for a CS course [42]. However, to what extent environmental cues affect real-world behavior and may be used strategically to increase STEM participation among women has not been studied systematically. In this paper, we make three main contributions to the literature:

(1) We introduce *psychologically inclusive design* with a taxonomy of cues for the design of digital learning environments, translating relevant theory and evidence from social psychology.
(2) We report large-scale experimental and quasi-experimental evidence on the behavioral consequences of cues strategically placed to increase participation of women in STEM education.
(3) We offer evidence-based recommendations for improving the participation of women in STEM education and an overview of open questions in the design of psychologically inclusive learning environments.

CHI 2019, May 4–9, 2019, Glasgow, Scotland, UK

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300704

2 SOCIAL BELONGING & IDENTITY THREAT

Membership in social groups constitutes an essential part of people's identities (e.g., gender, race, ethnicity, nationality) and they strive to maintain a positive perception of their social identities [2, 62]. When people worry that they are seen as less competent or treated unfairly because of a particular social identity they hold, they are said to experience social identity threat [60]. This can impair cognitive and social processes, which hinders student learning [63], limits working memory [53], and reduces self-regulatory abilities [1]. Concerns about identity-based discrimination can also undermine feelings of belonging, which is a critical antecedent to motivation and success in academic contexts [48, 65, 67].

Negative group stereotypes are the primary reason for fearing social exclusion and devaluation by others; the prospect of confirming such stereotypes is psychologically threatening and can cause academic underperformance. Stereotype threat has been shown to contribute to major academic achievement gaps in the United States, including the underperformance of African American students [59], first-generation college students [24, 61], students from lower socioeconomic backgrounds [13], and women in historically male-dominated STEM programs [43, 57, 69]. It is plausible that psychological processes of social belonging and identity threat matter not only in face-to-face but also computer-mediated environments [32]. Appraisals of identity-safety and trustworthiness may therefore be equally important and consequential in online environments as they are in in-person settings. People's sense of trust is known to influence decisions to initially adopt and continue to use information technologies [52]. The extent to which this psychological insight translates to the context of online education is the focus of this research.

3 PSYCHOLOGICALLY INCLUSIVE DESIGN

How people experience an environment is subjective because of the processes by which individuals perceive and interpret information (construals) and generate causal explanations of events (attributions) in the environment. *Psychologically inclusive design* is an approach to creating digital environments that afford equal opportunities to diverse learners by strategically manipulating content and design cues that influence construals and attributions [32]. The digital format offers unique affordances for creating psychologically inclusive learning environments compared to physical classrooms. Changes to the design and content of online learning environments can be implemented swiftly and at much lower cost. In addition to the cues that can be changed in physical environments, such as the course content and classroom design, creators of online learning environments also have control over the interaction design, which has a strong influence on people's experience with digital technology. Cues can be categorized into four types (Fig. 1): verbal content (written and spoken word); visual content (images, animations, photos, videos, etc.); visual design (color scheme, organization of information, etc.); and interaction design (e.g., how users input information, make choices, navigate the environment, and interact socially).

Figure 1: Taxonomy of cues in psychologically inclusive design.

3.1 Verbal Content Cues

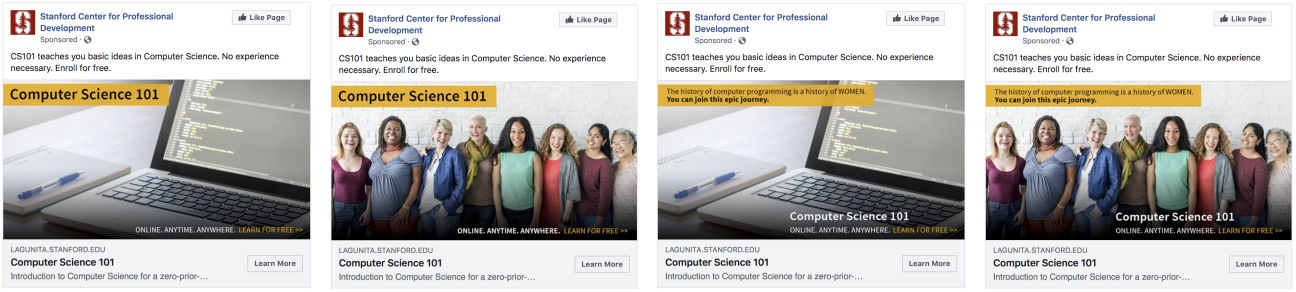
Online learning environments are filled with verbal content such as the course title and description, lecture content, and assessment prompts. Verbal cues can undermine women's sense of belonging or threaten their gender identity. Stating or merely insinuating the relevance of a stereotype and linking it to a person's social identity has been found to raise stereotype threat [e.g. 7, 30]. For example, Kray et al. [36] found that women experience stereotype threat when the importance of leadership skills is emphasized. Online learning environments make it quite easy to remove these types of threatening verbal cues.

Psychologically inclusive verbal cues can also be added to allay identity-based concerns. These cues assure people that they are respected for who they are as individuals and that their social identity does not pose any barriers in the environment for them. Adding an explicit "identity-safe sentence" refuting the presence of gender-based differences in performance can alleviate women's concerns [17, 46, 66]. Several studies document positive effects of reducing uncertainty by assuring women that everyone is welcome, respected, and treated fairly [9, 41]. Still, subtle verbal cues targeting individuals' construals of the situation, such as presenting tests as "puzzles" or "problem-solving exercises", can diffuse apprehension about evaluations [4, 56].

3.2 Visual Content Cues

Visual content cues can foster trust and belonging, but they are also a potential source of psychological threat by inadvertently conveying certain values, norms, and culture. A simple yet influential visual cue is numeric representation. How many in-group members are represented in an environment in absolute or relative terms can influence the performance of identity-threatened individuals [26, 27, 44]. A visual cue that conveys equal numeric representation, such as a photo depicting a gender-balanced group of learners, can signal diversity in an environment, or at least a commitment to diversity.

In addition to general numeric representation, the representation of in-group members as role models or in positions of power can foster trust, fortify belonging, and promote performance [14, 15, 39, 71]. In an online course on data science, Brooks et al. [6] found that visual gender cues in lecture videos (i.e. the gender of people in the background of the lecturer, and guest talks from either a man or a woman) influenced participation rates on the course discussion forum—women posted more in the female cue condition and vice versa.



(a) Default (b) Inclusive Image (c) Inclusive Text (d) Inclusive Image & Text

Figure 2: Four versions of an online course advertisement with a verbal cue and a visual cue manipulated to convey that the course also provides an inclusive environment for women. The visual cue shows eight women of diverse age and ethnicity. The verbal cue reads, "The history of computer programming is a history of WOMEN. You can join this epic journey."

3.3 Visual Design Cues

Design influences people's emotional and behavioral responses towards products, their usability, and their commercial success [47]. The appearance of an online learning environment can influence people's first impression and their willingness to seriously engage with it [50]. Product designers have long adapted the colors, shapes, and forms of everyday objects to appeal to women, as epitomized by the "shrink it and pink it" movement [12]. Yet psychological research on ambient belonging shows that it is not necessary to create overtly feminine designs; in fact, this approach may backfire. Instead, neutral visual designs that omit gender-stereotypic cues have been found to effectively raise women's anticipated sense of belonging [10, 42]. For example, Metaxa-Kakavouli et al. [42] improved gender inclusivity by manipulating the visual web design of a CS course page, including its color scheme, fonts, and background image. Identifying and questioning unconscious assumptions in design is particularly important when the designers are not representative of the (intended) user audience [54]. For example, gender bias is frequently encoded in the visual design of web interfaces and products created by male-dominated engineering teams [8, 70].

3.4 Interaction Design Cues

Interaction design can communicate the culture and values of those who created the environment and plays a critical role in enabling the authentic expression of identity in online environments. Recent work on culturally adaptive interfaces [51] and website morphing [25] has developed automated approaches to adapt the interaction and visual design of web applications, such as the website navigation structure, to accommodate users from different backgrounds. For instance, when the design of an interface requires members of a social group to give up or conceal aspects of their identity to gain access or to reach equal status, it can convey to these individuals that they are not valued and that they do not belong. Self-report forms frequently provide only two options to indicate gender, which can cause gender-identity-based concerns among individuals who identify as non-binary [29]. Moreover, algorithmic bias in artificial intelligence interfaces tends to manifest in interaction design, for example in gender-bias in search engine results for occupations [31] or lower face recognition accuracy for people who are not white males [38].

4 RESEARCH QUESTIONS

Only a handful of studies have shown that digital cues can trigger concerns among women about their social belonging in STEM: Murphy et al. [44] showed that visual gender representation cues in a promotional video about a STEM conference influences women's anticipated sense of belonging at the event; Cheryan et al. [9] found that decorations in a virtual reality computer science classroom (a visual cue) influence women's anticipated belonging; Metaxa-Kakavouli et al. [42] used a gender-stereotypic web design (visual design cue) for a computer science course to influence anticipated belonging and enrollment intentions for the class; and Davies et al. [16, 17] showed that gender-stereotypic TV commercials (visual & verbal cues) raise stereotype threat in women and reduce interest in STEM-related educational/vocational options. These studies suggest that brief exposure to visual content, verbal content, and visual design cues can trigger concerns about social belonging that reduce women's reported inclination to enter a STEM environment. To what degree these attitudinal findings from the lab translate to behavioral consequences in the real-world is an important open question. We investigate the following research questions on how digital cues impact women's behavior towards STEM educational opportunities:

RQ1 How do psychologically inclusive design cues in social media advertising campaigns influence women's (and men's) decision to learn more about a STEM course?

RQ2 How do psychologically inclusive design cues on course enrollment platforms influence women's (and men's) decision to enroll in STEM courses?

We investigate these research questions in two field experiments instead of laboratory studies to strengthen the ecological validity of our findings and provide realistic effect sizes to guide policy and subsequent research. The first experiment investigates the impact of visual and verbal content cues on clicks on social media ads for a STEM course. We chose the advertising context because it is a point of entry into a STEM education; specifically, we use

Facebook as a major online advertising platform that is used to promote educational offerings. Our second study examines the impact of visual and verbal content cues on a course enrollment platform on enrollments in a STEM course. Although they are conceptually no less important than the other types of cues, we do not investigate visual and interaction design cues in this study. We test the following hypotheses which are grounded in theory and prior evidence:

H1 Psychologically inclusive verbal and visual content cues (a) increase women's engagement with advertisements for a STEM course, but (b) do not influence men's engagement.

H2 Psychologically inclusive verbal and visual content cues increase the proportion of women enrolling in a STEM course, but (b) do not affect enrollment rates among men.

5 STUDY 1: CUES IN ONLINE ADVERTISING

Advertising is a common method for disseminating information about educational opportunities and frequently the first source of information that shapes people's impressions of a product or service. Static online ads in search results or on social media platforms provide limited space to communicate general information about an online course or program. Yet the limited number of cues embedded in an ad can still (often unwittingly) affect anticipated belonging. We conducted a randomized trial on Facebook to address RQ1 about the effect of psychologically inclusive cues in advertising on women's and men's decisions to seek further information about an online computer science course. We independently manipulated visual and verbal content cues of the ad and examined aggregate ad engagement for men and women separately. Three criteria guided our choice of which course to advertise: (1) it had to be free to enroll in at any time, (2) have zero prerequisites, and (3) show a gender disparity in enrollments. In fact, analysis of prior enrollments indicated that women enrolled in the course at a much lower rate than men: only 27% of the most recent 40,000 enrollments were women.

5.1 Method

5.1.1 Study design. We conducted a field experiment with a 2 (visual cue) by 2 (verbal cue) factorial design on Facebook. Figure 2 shows the four versions of the ad we ran in cooperation with the Stanford Center for Professional Development. We ran two concurrent ad campaigns, one for women and one for men, each with a randomized content experiment using Facebook's advertising platform. Randomization was implemented using the "split test" functionality, which assigns people into ad content conditions on first exposure (i.e. a person is only exposed to one version of the ad in the experiment). The campaigns ran between June 9-23, 2018, each with a daily budget of $100, which was evenly split between the four ad versions: Default, Inclusive Image, Inclusive Text, and Inclusive Image & Text.

5.1.2 Sample. The ad campaigns were targeted to an audience expected to be interested in an online CS courses for professional development: 25-35 year old people with an Associate, Bachelor's, or Master's degree, located in the US and using the English (US) version of Facebook. To exclude individuals who may be overly skilled for an introductory class, we excluded those with an interest (inferred by Facebook) in "computer programming", "software engineering", and "information technology". The ads were displayed in the Facebook News Feed and were seen by over 200,000 people according to platform-provided statistics.

Table 1: Results for advertising experiment in terms of impressions (Imps.), total clicks, and estimated unique clicks.

	Men, $n \approx 104{,}896$		Women, $n \approx 103{,}584$	
Ad Version	Clicks (uniq.)	Imps.	Clicks (uniq.)	Imps.
Default	260 (242)	79,940	243 (227)	77,010
Inclusive Image	252 (233)	78,220	223 (206)	75,860
Inclusive Text	220 (204)	84,223	257 (233)	70,771
Inclusive Both	243 (227)	80,572	280 (258)	70,020

5.1.3 Measures. As outcomes, we obtained the following aggregate metrics for each ad version and campaign: the number of impressions (how often the ad was shown), the number of ad clicks including any type of interaction with the ad such as liking it or navigating to the page of the ad source, and the unique number of ad clicks (only counting the first time a person clicked the ad; a quantity that Facebook estimates). The click-through rate (CTR) is computed as percentage of total clicks for all impressions in each experimental group.

5.1.4 Power analysis. We conducted an *a priori* power analysis to determine the necessary number of impressions for a set of anticipated CTRs, informed by the outcomes of previous ad campaigns run by the same advertiser. The choice of the daily budget was based on these calculations. The actual CTR turned out to at the low end of what we had simulated. The power analysis, data set, and analysis script are available at https://osf.io/92pqx/.

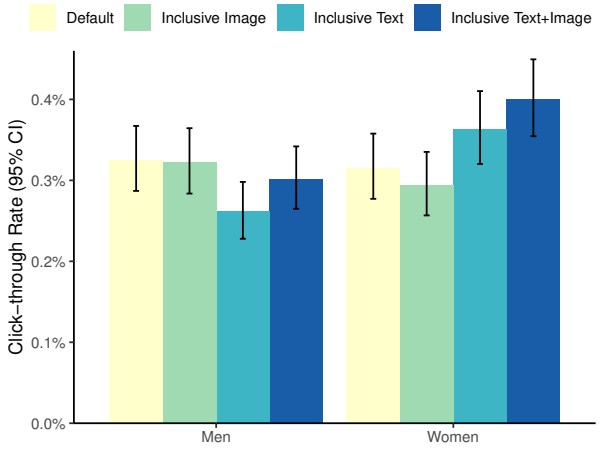


Figure 3: Engagement with four versions of an advertisement among men and women in terms of the click-through rate with 95% confidence intervals assuming independence.

5.2 Results

Overall, the ads were shown 616,616 times in the Facebook News Feed, reaching an estimated 104,896 men and 103,584 women. Table 1 provides descriptive statistics for each campaign. For the same ad campaign budget, the overall number of impressions was lower for the campaign targeting women than men (293,661 vs. 322,955), because it is more costly to show ads to those less likely to engage with them (a higher ad auction bid is required to win).

A direct comparison of click-through rates is shown in Figure 2. We analyzed CTR outcomes using two logistic regressions one for the campaign targeting women and one for the campaign targeting men. In the campaign targeting women, the odds of clicking were 26% higher ($z = 2.707, p = 0.0068$) for the inclusive ad (image & text) than the default. The ad version with the inclusive text showed a not statistically significant 15% increase ($z = 1.573, p = 0.1157$) over the default, while the inclusive image version showed no practical difference from the default ($z = -0.765, p = 0.4442$). The campaign targeting men showed a different pattern. The odds of clicking were 20% lower ($z = -2.397, p = 0.0165$) for the inclusive text version than the default. The CTR of the other two versions was comparable to the default ($|z| < 0.9, p > 0.39$).

To summarize, the ad combining inclusive visual and verbal cues substantially increased engagement among women without reducing engagement among men. While the inclusive text ad which highlighted the role of women in CS tended to be appealing to women, it actually reduced the CTR among men. These results answer our first research question by providing evidence in support of hypothesis H1a that visual and verbal content cues (in combination) increase women's engagement; for men (H1b), the cues manipulation made no significant difference except for the verbal content cue, which inadvertently reduced engagement.

6 STUDY 2: CUES AT COURSE ENROLLMENT

Study 1 focused on the first step in the "pipeline" into STEM education: using psychologically inclusive design to increase women's information-seeking behavior about a learning opportunity by influencing anticipated belonging. Study 2 focuses on the very next step: the decision to enroll in a STEM learning opportunity. To address RQ2 about women's and men's decisions to enroll in a STEM course, we conducted a quasi-experiment to test the effect of adding psychologically inclusive visual and verbal content cues on the enrollment page of a statistics course. The course was selected according to the same criteria as in the prior study: free availability, no prerequisites, and evidence of a gender gap (only 23% of the most recent 26,000 enrollments were women).

6.1 Method

6.1.1 Study design. We designed a quasi-experimental study to quantify the behavioral consequences of psychologically inclusive cues on online course enrollment. The study combines an ABAB interrupted time-series design with a matched difference-in-difference (DID) design. Participants enrolled in online courses on the Stanford Lagunita platform, which offers free open-scale online courses. The ABAB design was implemented by successively changing elements of the course landing page that people visit before enrolling in the target course, an introductory course on Probability and Statistics. The matched DID design was implemented by manipulating a single course on the platform and comparing outcomes with similar courses that were open for enrollment during the same time.

Table 2: Timeline of quasi-experimental study.

	Condition	Start Date	End Date	Enrolled	Women
Pre	Default	1-2-2016	30-3-2017	26,154	22.6%
A	Default	30-3-2017	27-4-2017	3,697	21.9%
B	Inclusive	27-4-2017	16-9-2017	17,911	26.9%
A	Default	16-9-2017	10-6-2018	16,728	26.8%
B	Inclusive	10-6-2018	8-9-2018	3,510	21.5%

(a) Default

(b) Inclusive image

An Inclusive and Supportive Place to Learn
This course provides an accessible, inclusive, and supportive space to learn. Anyone can enroll from anywhere in the world and everyone, no matter what their gender, ethnicity, or socio-economic status, can be successful. Our goal is to help you learn interesting topics in probability and statistics. Welcome!

(c) Inclusivity statement

Now anyone can develop an understanding of probability and statistics with this self- paced, interactive course. Creative problem solving activities and simulations walk you through the basics of exploratory data analysis, how to sample data, and how to draw inferences. The course provides a fair and inclusive learning environment to support learners of all backgrounds. Free Enrollment.

(d) Course description in marketing email

Figure 4: Landing page and marketing manipulation.

6.1.2 Sample. Table 2 provides an overview of the study timeline, sample size, and the percentage of enrollments from women in the target course. For comparison, we also show historical data of a previous offering of the target course indicated as "Pre". Over the course of the study, the target course enrolled a total of 41,846 learners, 44% located in the United States, and based on self-report 26% women, mean age 33 (SD=11), 39% with a college degree, 31% with a master's degree, and 6% with a PhD. While the target course was running, 108 other courses on a variety of subjects were available and enrolled 338,878 unique learners (508,937 total enrollments).

6.1.3 Manipulation. The treatment was a manipulation of two cues on the landing page of an online statistics course. The course image was changed from a "cold" graphic with graphs and equations in the background (Fig. 4a) to a "warm" picture featuring a woman creating a pie chart (Fig. 4b). Moreover, an inclusivity statement (Fig.

4c) was added to the page in a prominent location. The statement explicitly says that individuals of all genders and from anywhere in the world are welcome. We timed the first change of the landing page (but not subsequent ones) with a wide email announcement of course offerings which incorporated a short course description with psychologically inclusive language (Fig. 4d). The announcement was sent to 1.9 million email addresses and it was opened 700,000 times according to email tracking data. As detailed in Table 2, after observing course enrollment rates for two months with the default landing page ("A"), we manipulated the course landing page by adding psychologically inclusive verbal and visual content cues ("B") and tracked enrollment rates for nearly five months. Then, we switched the enrollment page back to the default ("A") for nine months before adding the cues back in again ("B") for three months.

6.1.4 Measures. We collected individual course enrollments with self-reported gender. The percentage of women is computed relative to all enrollments (not just those that self-reported a gender), thus assessing the decision to self-report gender as "woman". As described next, the analysis compares the proportion of women enrolling in the treated course and in matched control courses during three two-month study periods, each centered around the time inclusivity cues were manipulated. We consider only one enrollment for each unique user in each of these three periods independent of their course engagement and subsequent decision to unenroll (re-enroll).

6.1.5 Analysis. The causal effect estimate is least biased close to the time of manipulation because whether an individual sees the page in the default or inclusive state is plausibly random and contemporaneous changes unrelated to the manipulation are less likely to occur. We therefore consider three two-month periods, one month before and one month after each interface change: A to B (period 1), B to A (period 2), and A to B (period 3). Figure 5 illustrates how enrollment rates changed in each period for the target course and the matched control courses. For each period, we estimate the difference-in-difference estimator which quantifies the effect attributable to the manipulation. We identified matching control courses based on the proportion of enrolled women. In each period, we retain control courses with at least 30 enrollments and a proportion of enrolled women within 10 percentage points of the target course in the month before the manipulation. We ran regression analyses of individual enrollments as a binary outcome (female 0/1) using a linear probability model with cluster-robust standard errors (clustering on course IDs). Weekday was added as a fixed-effect covariate to mitigate variance due to cyclical trends.

6.2 Results

We tested to what extent the manipulation of inclusivity cues affected the enrollment of women in the course (H2). In periods 1 and 3, psychologically inclusive cues were added; in period 2, they were removed. Figure 5 illustrates the change in the proportion of women enrolling in the target course and matched control courses during the three two-month periods around the change of the landing page. The difference-in-difference estimate in the first period indicated that the proportion of women increased by 3.88 percentage points (an 18% increase) when cues were added and it dropped in period 2 by 2.35pp (an 8% drop) when cues were removed. Surprisingly, women enrollment also dropped in period 3 by 2.1pp (a 9% drop) despite cues being added back in. Pooling across the three study periods, the combined DID effect estimate of inclusivity cues is positive and statistically significant. Estimates and test statistics are provided in Table 3.

Furthermore, we examined how the enrollment rate changed for men and women separately, as the observed change in the percentage of women could also be due to enrollment changes among men. Figure 6 shows the normalized rate of enrollment in

Table 3: Difference-in-difference estimates of the effect of showing psychological inclusivity cues on the proportion of women enrolling in three study periods.

Period	N_{target}	N_{matched}	DID estimate
1	8,644	22,456; 30 courses	3.88pp, $t_{30} = 4.64$, $p < 0.001$
2	3,611	12,387; 22 courses	-2.35pp, $t_{22} = -2.19$, $p = 0.039$
3	1,390	12,387; 26 courses	-2.09pp, $t_{26} = -2.66$, $p = 0.013$
Pooled	13,645	49,210; 48 courses	2.71pp, $t_{48} = 4.75$, $p < 0.001$

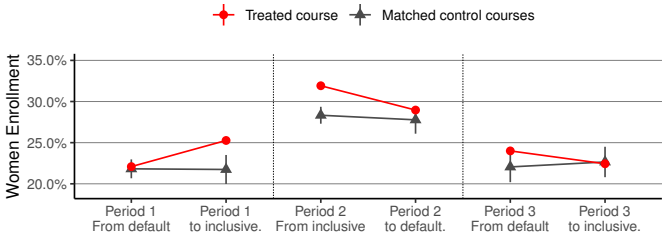


Figure 5: The proportion of course enrollments from women in the month before and after each manipulation of the landing page of the treated course is shown in red, and for comparison, the corresponding proportion in matched control courses is shown in gray. Error bars show robust standard errors clustered on course.

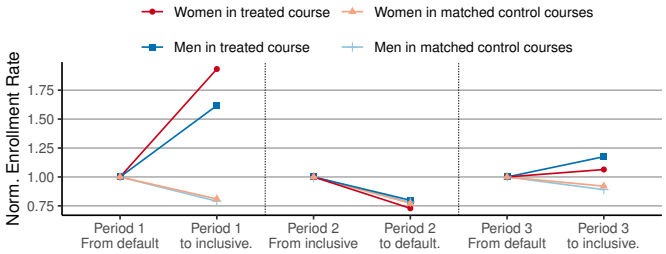


Figure 6: Normalized enrollment rate for men (blues) and women (reds) in each period for the target course (dark shade) and matched control courses (light shade) showing the change in enrollment numbers relative to the beginning of each study period.

each period relative to the start of the period. We can now see that underlying the effect in period 1 in the treated course is a 93% increase in enrollments from women relative to a 62% increase from men; while at the same time enrollments slightly dropped in the matched courses (-19% for women; -21% for men). In period 2, when the psychologically inclusive cues were removed, enrollment from women dropped 27% while enrollment from men dropped by only 20% (-23% for women and -22% for men in matched courses). In the final period, enrollments increased in the target course by 6.4% for women and 18% for men, compared to 8% and 11% enrollment drops, respectively, in the matched courses.

To summarize, the psychologically inclusive visual and verbal cues were effective at increasing enrollments from women in two out of three periods. This effect can be attributed to a greater increase (decrease) of enrollments from women relative to men. However, changes in the rate of men enrolling in the target course were in the same direction as for women, albeit smaller in magnitude in the first two periods. Thus, there is no evidence to suggest that male enrollment rates were adversely affected by the inclusivity cues.

7 DISCUSSION

Psychologically inclusive design is a theoretically grounded approach to creating digital environments that afford equal opportunities to diverse learners by strategically manipulating content and design cues [32]. We introduced a taxonomy of psychologically inclusive cues to constrain the complexity that arises from the expansive, multifaceted design space that digital environments afford. We then tested the effectiveness of the approach in two large field experiments. By demonstrating the behavioral consequences of subtle cues in online environments, this research extends prior work on the influence of environmental cues on psychological barriers to women's participation in STEM education. It represents one of the first published studies estimating the behavioral consequences of inclusive cues in real-world environments and thus contributes to both theory and practice. The broad theoretical foundation and evidence base of psychologically inclusive design means that the approach can also be applied in other contexts to improve inclusivity for people affected by psychological barriers, such as ethno-racial minorities or those with low socio-economic status. How effective the approach is in other contexts with other populations warrants scientific attention.

Overall, our findings show that (subtle) visual and verbal cues influence enrollment behaviors in online courses, specifically women's enrollment in online STEM courses. Certain cues, such as cold, "techy" images on online course enrollment pages, appear to have activated identity-based concerns which reduce anticipated belonging and discourage women from enrolling in STEM courses [10, 44, 49]. However, cues based on psychologically inclusive design (e.g., warmer pictures depicting female students as role models) can alleviate identity-based concerns and stimulate information-seeking and increase women's enrollment in STEM courses [42]. This research advances our understanding of the opportunities and limitations of psychologically inclusive design in practice.

7.1 Implications

Given the rapid expansion of online learning both nationally and internationally, and concerns about the participation of women in STEM disciplines, our work has practical implications for how to market and design platforms for online courses and even degree programs. A practical implication of our findings is that updating marketing collateral and the content of online course enrollment portals according to psychologically inclusive design principles can raise the participation of women in STEM courses. However, we found in Study 1 that changing visual cues alone may not send a strong enough signal to change construals of STEM learning environments as a "chilly", or psychologically threatening place [21, 69]. Updating an online advertisement with only an inclusive visual content cue or verbal content cue did not significantly raise enrollment behaviors relative to the default advertisement (though the verbal cue showed more promise than the visual cue in terms of the effect size). It was only when we updated both visual and verbal content cues that women's enrollment behaviors significantly increased. Thus, designers should consider adapting both visual and verbal cues to maximize the impact of psychologically inclusive design interventions.

Another implication of our findings is that psychologically inclusive cues may "backfire" with the majority group. Contrary to prior work (and our hypothesis 1b) that consistently finds men's attitudes to be unchanged by inclusivity cues targeting women [9, 10, 40], we found men's behavior to be negatively affected. Study 1 revealed a significant decrease in enrollment behaviors among men when they were targeted with a verbal cue designed to be psychologically inclusive for women. In the first two periods of Study 2, the rate of enrollment among men changed in the same direction as for women but at a slower pace. Likewise, Brooks et al. [6] found lower participation of men when exposed to female cues. To our knowledge, it is the only other published field experiment testing the impact of inclusive cues on behavior. This suggests that cues can have different behavioral effects for men and women in real-world settings. Moreover, it may be the case that the relationship between the strength of cues and behavioral effects is not linear. Stronger cues or ones where the intent is too obvious may cause psychological reactance (for a review, see [5]), which has been identified as an explanation for why some persuasion techniques (e.g., pop-up ads [19]) backfire. We therefore recommend that designers consider potential unintended consequences of cues, especially if only strong verbal cues are used, and when possible, to test messages with non-targeted populations for negative reactance.

A third implication of our results is that the effects of cues in real-world settings may be variable over time and influenced by other contextual factors. In Study 2, we essentially replicated the same intervention three times over the course of 18 months due to the ABAB design. A significant effect of inclusive cues on women's enrollment was detected two out of three times. The third time the effect was not only null but significantly reversed: the percentage of women's enrollments dropped. It may be the case that unrelated changes to the online environment such as adjustments to the search engine optimization or enrollment portal design elements may have affected the impact and efficacy of the psychologically inclusive cues. The positive effect of the cues was strongest when

the course was included in a direct email campaign and the course tile was displayed prominently on the course enrollment portal. Psychologically inclusive cues may be more effective the more salient and/or accessible they are relative to other cues in the online environment. Alternatively, the larger audience that accessed the course enrollment page when the course was prominently displayed may have included more women who are receptive to the cues. Although the overall effect across the entire study was significant in the predicted direction, the reliability of the effects of psychologically inclusive cues warrants further investigation.

7.2 Limitations

We conducted field studies to estimate the behavioral effects of psychologically inclusive cues with high ecological validity. However, to do so we had to give up some internal validity compared to laboratory studies. In particular, we had less control over the instrumentation of the research. In Study 1, we used Facebook's business advertising platform which provides a set of metrics to evaluate ad experiments. Our analysis focuses on a key metric in digital advertising, the click-through rate (CTR), which is the number of people who click on the ad when they are exposed. People may see an advertisement multiple times giving them repeated opportunities to click. Yet the platform provides only aggregate counts, which do not allow us to account for this source of dependence in the data. Using the unique number of clicks (i.e. a platform-provided *estimate* of how many people ever clicked on the ad) in the numerator instead of the total number of clicks yielded equivalent results. The platform also provides a 'reach' statistic, the *estimated* number of unique people exposed to the ad, but without information on the sampling variability. Using this estimate in the denominator renders the effects statistically insignificant.

Another limitation related to instrumentation is that we could not assess the psychological processes posited by theory in this research. Ideally, we could have tested the entire process in the wild, namely that psychologically inclusive cues affect enrollment behaviors by increasing anticipated belonging and reducing gender-identity threat (e.g. [33]). Metaxa-Kakavouli et al. [42] adapted a number of instruments for online learning environments, such as ones for ambient belonging and gender-related anxiety, that would help to confirm in future work whether psychologically inclusive design cues affect the key mechanisms identified by theories of social belonging and identity threat.

Moreover, the study population (targeted Facebook users and people familiar with the Stanford online course platform) and study materials (the specific verbal and visual content cues) limit the generalizability to other populations and contexts. In fact, we found that the exact same cues in the same context induced divergent effects in Study 2. Arguably, this suggests that our ability to reliably replicate behavioral effects in this area of research is bounded by our limited scientific understanding of the role of other contextual factors.

Finally, no quasi-experimental study (i.e., Study 2) is without threats to internal validity that may bias the causal inference. We combined three econometric methods to reduce potential bias and approximate an as-if random comparison. Nevertheless, three ways that the current design could be improved are: (1) manipulate more than one course to increase generalizability, (2) match courses before randomization such as in a matched-pair design to increase statistical power and cleanly separate study design from analysis, and (3) observe outcomes for shorter periods to reduce variance due to temporal variation in enrollment.

7.3 Future Work

Findings of this research are consistent with extant social-psychological theory and empirical results on social belonging and social identity threat, as well as recent work that has examined how cues in online environments can raise psychological barriers and identity-based concerns for certain groups of people. Consistent with prior work on social-identity contingencies, findings of Study 1 shows a positive effect of a verbal cue that states a commitment to diversity. Purdie-Vaughns et al. [49] independently manipulated two identity-safety cues in a brochure about a potential work setting: the company's diversity statement and the numeric representation of minorities in a picture. They found that so long as there was a statement embracing the value of diversity, it did not raise significant concerns among minorities if the photo showed minorities underrepresented. Thus, different cues can raise or lower threat, and even counteract the effects of other cues to render an environment psychologically inclusive [48].

Further research is necessary to inform the optimal choice of cues to implement and understand how the efficacy of different cues depends on audience characteristics and other contextual factors (i.e. heterogeneous treatment effects). Furthermore, while this research focuses on supporting the STEM pipeline with psychologically inclusive design, more work is needed to investigate the role of design cues in the subsequent steps of women's STEM trajectories. A natural extension of our work in this area is to track click-stream and performance data throughout a course after adding psychologically inclusive cues before enrollment. The study would build on prior work showing that lifting psychological barriers and changing construals of non-belonging in academic environments can give rise to virtuous recursive cycles [11, 68].

In both of our studies, we found evidence that men's behavior was affected by psychologically inclusive cues meant to increase women's participation. As discussed above, this stands in contrast to many previous studies that found men to remain unaffected by such cues. We specifically found in Study 1 evidence of a backfiring effect whereby male enrollments were significantly decreased by verbally inclusive cues. This raises both practical concerns and questions about whether contextual factors may reveal boundary conditions in extant theory and research. Future research is needed to reveal whether, for example, certain contexts and types and intensity of verbal cues may result in a backfiring effect for men.

Future work should also investigate the effects of the other two categories of cues presented in the taxonomy of psychologically inclusive design: visual design and interaction design cues. These were not tested in the current research, though they are important in the design of digital learning environments. In particular, a factorial design that experimentally manipulates all four types of cues would greatly advance our understanding of the relationships between cues in the taxonomy.

8 CONCLUSION

This research contributes to a small but growing scientific literature documenting that, just as in face-to-face settings, cues in online spaces can activate psychological barriers for certain segments of the population. Our studies demonstrate that psychologically inclusive design principles can be applied in real-world settings to ameliorate these barriers and influence behaviors such as increasing female participation in online STEM courses. In addition, we present a taxonomy that can inform future work exploring the behavioral consequences of different types psychologically inclusive cues: verbal content, visual content, visual design, and interaction design cues. These studies also reveal that cues may not operate in real-world settings as they do in the laboratory, and we have highlighted potential exogenous factors to consider in future work. Given the consequences of non-inclusive design, such as the perpetuation of enrollment and participation gaps for women in STEM disciplines, we implore researchers to systematically test and practitioners to earnestly explore psychologically inclusive design in various online learning contexts.

ACKNOWLEDGMENTS

We are grateful to Candace Thille who granted us permission to experiment with one of her online courses. We also thank J. Gray, J. Ly, I. Manasherob, B. Merkel, and V. Singh for help with implementing the advertising experiment.

REFERENCES

[1] Roy F Baumeister, C Nathan DeWall, Natalie J Ciarocco, and Jean M Twenge. 2005. Social exclusion impairs self-regulation. *Journal of personality and social psychology* 88, 4 (2005), 589–604.

[2] Roy F Baumeister and M R Leary. 1995. The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin* 117, 3 (may 1995), 497–529.

[3] Eric P Bettinger, Lindsay Fox, Susanna Loeb, and Eric S Taylor. 2017. Virtual classrooms: How online college courses affect student success. *American Economic Review* 107, 9 (2017), 2855–75.

[4] Jim Blascovich, S. J. Spencer, D. Quinn, and Claude M. Steele. 2001. African Americans and High Blood Pressure: The Role of Stereotype Threat. *Psychological Science* 12, 3 (2001), 225–229.

[5] Sharon S Brehm and Jack W Brehm. 2013. *Psychological reactance: A theory of freedom and control.* Academic Press.

[6] Christopher Brooks, Josh Gardner, and Kaifeng Chen. 2018. How Gender Cues in Educational Video Impact Participation and Retention. In *Proceedings of the 13th Intl. Conference of the Learning Sciences*. 1835–1842.

[7] Ryan P. Brown and Elizabeth C. Pinel. 2003. Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology* 39, 6 (2003), 626–633.

[8] Margaret Burnett, Simone Stumpf, Jamie Macbeth, Stephann Makri, Laura Beckwith, Irwin Kwan, Anicia Peters, and William Jernigan. 2016. GenderMag: A method for evaluating software's gender inclusiveness. *Interacting with Computers* 28, 6 (2016), 760–787.

[9] Sapna Cheryan, Andrew N. Meltzoff, and Saenam Kim. 2011. Classrooms matter: The design of virtual classrooms influences gender disparities in computer science classes. *Computers & Education* 57, 2 (2011), 1825–1835.

[10] Sapna Cheryan, Victoria C Plaut, Paul G Davies, and Claude M. Steele. 2009. Ambient belonging: how stereotypical cues impact gender participation in computer science. *Journal of personality and social psychology* 97, 6 (2009), 1045–1060.

[11] Geoffrey L Cohen, Julio Garcia, Valerie Purdie-Vaughns, Nancy Apfel, and Patricia Brzustoski. 2009. Recursive processes in self-affirmation: intervening to close the minority achievement gap. *Science (New York, N.Y.)* 324, 5925 (2009), 400–403.

[12] J Contrera. 2016. The end of "shrink it and pink it": A history of advertisers missing the mark with women. *The Washington Post* (8 June 2016). https://www.washingtonpost.com/lifestyle/style/the-end-of-shrink-it-or-pink-it-a-history-of-advertisers-missing-the-mark-with-women/2016/06/08/3bcb1832-28e9-11e6-ae4a-3cdd5fe74204_story.html

[13] J.-C. Croizet and Theresa Claire. 1998. Extending the Concept of Stereotype Threat to Social Class: The Intellectual Underperformance of Students from Low Socioeconomic Backgrounds. *Personality and Social Psychology Bulletin* 24, 6 (1998), 588–594. arXiv:0803973233

[14] Henry a. Danso and Victoria M. Esses. 2001. Black Experimenters and the Intellectual Test Performance of White Participants: The Tables Are Turned. *Journal of Experimental Social Psychology* 37, 2 (2001), 158–165.

[15] Nilanjana Dasgupta. 2011. Ingroup Experts and Peers as Social Vaccines Who Inoculate the Self-Concept: The Stereotype Inoculation Model. *Psychological Inquiry* 22, 4 (oct 2011), 231–246. https://doi.org/10.1080/1047840X.2011.607313

[16] Paul G Davies, Steven J Spencer, D. M. Quinn, and R. Gerhardstein. 2002. Consuming Images: How Television Commercials that Elicit Stereotype Threat Can Restrain Women Academically and Professionally. *Personality and Social Psychology Bulletin* 28, 12 (2002), 1615–1628.

[17] Paul G Davies, Steven J Spencer, and Claude M. Steele. 2005. Clearing the air: identity safety moderates the effects of stereotype threat on women's leadership aspirations. *Journal of personality and social psychology* 88, 2 (2005), 276–287.

[18] D Deming, C Goldin, LF Katz, and N Yuchtman. 2014. Can online learning bend the cost curve of higher education. In *American Economic Review*.

[19] Steven M Edwards, Hairong Li, and Joo-Hyun Lee. 2002. Forced exposure and psychological reactance: Antecedents and consequences of the perceived intrusiveness of pop-up ads. *Journal of Advertising* 31, 3 (2002), 83–95.

[20] Maya Escueta, Vincent Quan, Andre Joshua Nickow, and Philip Oreopoulos. 2017. Education Technology: An Evidence-Based Review. (2017), 102. https://doi.org/10.3386/w23744

[21] Faye Flam. 1991. Still a "Chilly Climate" for Women? *Science* 252, 5013 (1991), 1604–1606. https://doi.org/10.1126/science.252.5013.1604

[22] Scott A Ginder, Janice E Kelly-Reid, and Farrah B Mann. 2017. Enrollment and Employees in Postsecondary Institutions, Fall 2016; and Financial Statistics and Academic Libraries, Fiscal Year 2016: First Look (Provisional Data). NCES 2018-002. *National Center for Education Statistics* (2017).

[23] J. D. Hansen and J. Reich. 2015. Democratizing education? Examining access and usage patterns in massive open online courses. *Science (New York, N.Y.)* 350, 6265 (2015), 1245–1248.

[24] Judith M Harackiewicz, Elizabeth A Canning, Yoi Tibbetts, Cynthia J Giffen, Seth S Blair, Douglas I. Rouse, and Janet S Hyde. 2014. Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology* 106, 2 (2014), 375–389.

[25] John R Hauser, Glen L Urban, Guilherme Liberali, and Michael Braun. 2009. Website Morphing. *Marketing Science* 28, 2 (mar 2009), 202–223. https://doi.org/10.1287/mksc.1080.0459 arXiv:1606.07988

[26] Michael Inzlicht and Talia 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 (2000), 365–371.

[27] Michael Inzlicht and Talia Ben-Zeev. 2003. Do high-achieving female students underperform in private? The implications of threatening environments on intellectual processing. *Journal of Educational Psychology* 95, 4 (2003), 796–805.

[28] Shanna Smith Jaggars. 2014. Democratization of Education for Whom? Online Learning and Educational Equity. *Diversity & Democracy* 17 (2014). Issue 1. https://www.aacu.org/diversitydemocracy/2014/winter/jaggars

[29] Samantha Jaroszewski, Danielle Lottridge, Oliver L Haimson, and Katie Quehl. 2018. Genderfluid or Attack Helicopter: Responsible HCI Research Practice with Non-binary Gender Variation in Online Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 307.

[30] Michael Johns, Toni Schmader, and Andy Martens. 2005. Knowing Is Half the Battle Women's Math Performance. *Psychological Science* 16, 3 (2005), 175–179.

[31] Matthew Kay, Cynthia Matuszek, and Sean A Munson. 2015. Unequal representation and gender stereotypes in image search results for occupations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 3819–3828.

[32] Rene F. Kizilcec. 2017. *Identity, threat, and belonging in online learning environments.* Ph.D. Dissertation. Stanford University.

[33] René F Kizilcec, Glenn M Davis, and Geoffrey L Cohen. 2017. Towards equal opportunities in MOOCs: affirmation reduces gender & social-class achievement gaps in China. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale*. ACM, 121–130.

[34] René F. Kizilcec and Sherif Halawa. 2015. Attrition and Achievement Gaps in Online Learning. In *Proceedings of the Second ACM Conference on Learning@Scale*. ACM, 57–66.

[35] René F Kizilcec, Andrew J Saltarelli, Justin Reich, and Geoffrey L Cohen. 2017. Closing global achievement gaps in MOOCs. *Science (New York, N.Y.)* 355, 6322 (2017), 251–252.

[36] Laura J Kray, Adam D Galinsky, and Leigh Thompson. 2002. Reversing the Gender Gap in Negotiations: An Exploration of Stereotype Regeneration. *Organizational Behavior and Human Decision Processes* 87, 2 (2002), 386–409.

[37] Christine Logel, Gregory M Walton, Steven J Spencer, Emma C Iserman, William von Hippel, and Amy E Bell. 2009. Interacting with sexist men triggers social identity threat among female engineers. *Journal of personality and social psychology* 97, 4 (2009), 578.

[38] Steve Lohr. 2018. Facial recognition is accurate, if you're a white guy. *New York Times* (9 Feb. 2018). https://www.nytimes.com/2018/02/09/technology/facial-recognition-race-artificial-intelligence.html

[39] David M. Marx and Phillip Atiba Goff. 2005. Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology* 44, 4 (dec 2005), 645–657.

[40] Allison Master, Sapna Cheryan, Andrew N Meltzoff, Allison Master, Sapna Cheryan, and Andrew N Meltzoff. 2015. Computing Whether She Belongs: Stereotypes Undermine Girls' Interest and Sense of Belonging in Computer Science. *Journal of Educational Psychology* 107, 3 (2015).

[41] Rodolfo Mendoza-Denton, Lindsay Shaw-Taylor, Serena Chen, and Eunice Chang. 2009. Ironic effects of explicit gender prejudice on women's test performance. *Journal of Experimental Social Psychology* 45, 1 (2009), 275–278.

[42] Danaë Metaxa-Kakavouli, Kelly Wang, James A Landay, and Jeff Hancock. 2018. Gender-Inclusive Design: Sense of Belonging and Bias in Web Interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 614.

[43] Akira Miyake, 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 (New York, N.Y.)* 330, 6008 (2010), 1234–1237.

[44] Mary C Murphy, Claude M. Steele, and James J Gross. 2007. Signaling Threat. *Psychological science* 18, 10 (2007), 879–885.

[45] Mary C. Murphy and Valerie Jones Taylor. 2011. The Role of Situational Cues in Signaling and Maintaining Stereotype Threat. In *Stereotype Threat: Theory, Process, and Applications*, M. Inzlicht and T. Schmader (Eds.). Oxford University Press, New York, NY, USA, Chapter 2, 17–33.

[46] Hannah-Hanh D Nguyen and Ann Marie Ryan. 2008. Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *The Journal of applied psychology* 93, 6 (2008), 1314–1334.

[47] Donald A. Norman. 2004. *Emotional Design: Why we love (or hate) everyday things*. Basic Books.

[48] Daphna Oyserman, Daniel Brickman, Deborah Bybee, and Aaron Celious. 2006. Fitting in matters: Markers of in-group belonging and academic outcomes. *Psychological Science* 17, 10 (2006), 854–861.

[49] Valerie Purdie-Vaughns, Claude M. Steele, Paul G Davies, Ruth Ditlmann, and Jennifer Randall Crosby. 2008. Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology* 94, 4 (2008), 615–630.

[50] Katharina Reinecke. 2013. Many People , Many Eyes : Aggregating Influences of Visual Perception on User Interface Design. (2013), 3299–3302.

[51] Katharina Reinecke and Abraham Bernstein. 2011. Improving Performance, Perceived Usability, and Aesthetics with Culturally Adaptive User Interfaces. *Transactions on Computer-Human Interactions (TOCHI)* 18, 2 (2011), 1–29.

[52] Jens Riegelsberger, M. Angela Sasse, and John D. McCarthy. 2005. The mechanics of trust: A framework for research and design. *International Journal of Human Computer Studies* 62, 3 (2005), 381–422.

[53] Toni Schmader and Michael Johns. 2003. Converging evidence that stereotype threat reduces working memory capacity. *Journal of personality and social psychology* 85, 3 (2003), 440–452.

[54] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. 2005. Reflective design. In *Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility*. ACM, 49–58.

[55] Jenessa R Shapiro and Amy M Williams. 2012. The role of stereotype threats in undermining girls' and women's performance and interest in STEM fields. *Sex Roles* 66, 3-4 (2012), 175–183.

[56] Steven J. Spencer, Steven Fein, Connie T. Wolfe, Christina Fong, and Meghan A. Dunn. 1998. Automatic Activation of Stereotypes: The Role of Self-Image Threat. *Personality and Social Psychology Bulletin* 24 (1998), 1139–1152. Issue 11.

[57] Steven J. Spencer, Claude M. Steele, and Diane M. Quinn. 1999. Stereotype Threat and Women's Math Performance. *Journal of Experimental Social Psychology* 35, 1 (1999), 4–28.

[58] Claude M. Steele. 1997. A threat in the air. How stereotypes shape intellectual identity and performance. *The American Psychologist* 52, 6 (1997), 613–629.

[59] Claude M. Steele. 2010. *Whistling vivaldi*. WW Norton & Company, New York, N.Y.

[60] Claude M. Steele, Steven J Spencer, and Joshua Aronson. 2002. Contending with group image: The psychology of stereotype and identity threat. *Advances in Experimental Social Psychology* 34 (2002), 379–440.

[61] Nicole M. Stephens, Stephanie A. Fryberg, Hazel Rose Markus, Camille S. Johnson, and Rebecca Covarrubias. 2012. Unseen disadvantage: How American universities' focus on independence undermines the academic performance of first-generation college students. *Journal of Personality and Social Psychology* 102, 6 (2012), 1178–1197.

[62] H Tajfel and JC Turner. 1986. The Social Identity Theory of Intergroup Behavior. In *Psychology of Intergroup Relations*, S. Worchel and WG Austin (Eds.). Nelson-Hall Publishers, Chicago, I.L., 7–24.

[63] Valerie Jones Taylor and Gregory M Walton. 2011. Stereotype threat undermines academic learning. *Personality and social psychology bulletin* 37, 8 (2011), 1055–1067.

[64] M. Mitchell Waldrop. 2013. Online learning: Campus 2.0. *Nature* 495, 7440 (2013), 160–163.

[65] Gregory M Walton and Shannon T Brady. 2017. The many questions of belonging. In *Handbook of Competence and Motivation (2nd Edition): Theory and Application*, Andrew J Elliot, Carol S Dweck, and David S Yeager (Eds.). Guilford Press, New York, NY, USA, Chapter 15, 272–293.

[66] Gregory M Walton and Geoffrey L. Cohen. 2003. Stereotype lift. *Journal of Experimental Social Psychology* 39, 5 (2003), 456–467.

[67] Gregory M Walton and Geoffrey L Cohen. 2007. A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology* 92, 1 (2007), 82–96.

[68] Gregory M Walton and Geoffrey L Cohen. 2011. A brief social-belonging intervention improves academic and health outcomes of minority students. *Science* 331, 6023 (2011), 1447–1451.

[69] Gregory M Walton, Christine Logel, Jennifer M Peach, Steven J Spencer, and Mark P Zanna. 2015. Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology* 107, 2 (2015), 468–485.

[70] Gayna Williams. 2014. Are you sure your software is gender-neutral? *Interactions* 21, 1 (2014), 36–39.

[71] Daryl A Wout, Margaret J Shih, James S Jackson, and Robert M Sellers. 2009. Targets as perceivers: How people determine when they will be negatively stereotyped. *Journal of Personality and Social Psychology* 96, 2 (2009), 349–362.