

UC Berkeley

UC Berkeley Previously Published Works

Title

Agentic but not warm: Age-gender interactions and the consequences of stereotype incongruity perceptions for middle-aged professional women

Permalink

<https://escholarship.org/uc/item/99s3p1k6>

Authors

Chatman, Jennifer A
Sharps, Daron
Mishra, Sonya
[et al.](#)

Publication Date

2022-11-01

DOI

10.1016/j.obhdp.2022.104190

Copyright Information

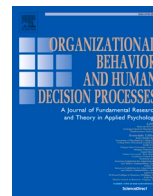
This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed



Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp

Agentic but not warm: Age-gender interactions and the consequences of stereotype incongruity perceptions for middle-aged professional women[☆]

Jennifer A. Chatman^{a,*}, Daron Sharps^b, Sonya Mishra^a, Laura J. Kray^a, Michael S. North^c

^a University of California, Berkeley, United States

^b Pinterest, United States

^c New York University, United States

ARTICLE INFO

Keywords:

Gender-age interactions
Discrimination
Social categories
Careers
Longitudinal analysis

ABSTRACT

We propose that perceptions of professional women change differently than perceptions of men as they age. Drawing inspiration from intersectionality theory, we examine the interaction of age and gender, finding that professional women are seen as more agentic, but also maximally incongruent with the gender-intensified prescription of being communal, in middle age. Our experiment showed that middle-aged women were perceived as agentic, like men, but also as declining more in warmth between young adulthood and middle age. Our field study also showed that middle-aged professional women are viewed as similarly agentic but less warm than men. Our longitudinal within-person study showed that these perceptions have consequences: Unlike men, middle-aged women (professors) received lower performance evaluations compared to their younger selves. Further, a linguistic analysis showed that middle-aged women professors were acknowledged to be more agentic, but also criticized for violating communal stereotype prescriptions, which mediated the link between age and women's, but not men's, performance evaluations.

"Ageism is alive and well. It is okay for men to get older, because men become more desirable by being powerful. With women, it's all about...trying to stay young" – Jane Fonda (2015).

Women, compared to men, face different expectations as they age, particularly with respect to agency and warmth, two central dimensions on which people are perceived and evaluated (Fiske, 2018). Though research has often focused on contrasting young and old people, we suggest that, for professional women, middle age is a particularly consequential, yet often overlooked, period. The pattern of changing perceptions of men and women over the lifespan may uniquely influence evaluations of women's effectiveness at work and potentially harm their career advancement. Therefore, we develop a theory explaining how perceptions of men's and women's agency and warmth evolve over time, and why women face the most significant consequences of those changing perceptions, lower performance evaluations, in middle age.

Understanding how workers are perceived at different ages is important given the demographic changes in the workforce (e.g., U.S. Bureau of Labor Statistics, 2022) and research showing that older workers have fewer opportunities to be hired and promoted, and are

more likely to be fired (e.g., North & Fiske, 2015). We suggest that there may be a significant age-gender interaction, as Jane Fonda references in the quotation above, in which aging may be associated with more negative consequences for women than for men. Identifying differences in perceptions between professional men and women as they age is important because "trying to stay young" as Fonda suggests, is only metaphorically possible.

Perceptions of warmth and agency dominate interpersonal judgments (e.g., Bakan, 1966) and gender differences in perceptions of warmth and agency have been well established (e.g., Eckes, 2002). Because agentic women are perceived as deviating from central gender role prescriptions (that women should be less agentic and more communal), they are viewed as discrepant, atypical, and counter-stereotypical (e.g., Chatman et al., 2008). Once women are perceived as counter-stereotypical, perceivers also reason that they are likely to be less communal too (e.g., Okimoto & Brescoll, 2010). Compared with identically behaving men, agentic women are judged as similarly competent, but less likable and less hireable. In contrast, agentic men are not perceived to have violated gender role prescriptions and thus do not

[☆] We are immensely grateful to Jay Stowsky for supporting this research. We also thank Robert Pickett for providing statistical advice, Mikki Hebl and Juan Madera for sharing their dictionaries, and Charles O'Reilly for comments on an earlier draft of the paper.

* Corresponding author.

E-mail address: chatman@berkeley.edu (J.A. Chatman).

<https://doi.org/10.1016/j.obhdp.2022.104190>

Received 27 October 2020; Received in revised form 15 August 2022; Accepted 1 September 2022

0749-5978/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

experience the concurrent penalty of being seen as necessarily less communal.

We take this logic a step further by considering whether gender interacts with age in shaping social perceptions. We draw inspiration from intersectionality theory, which offers a framework for examining interconnections between social categories and their power dynamics (e.g., Atewologun, 2018), typically at the intersection of different fixed, stigmatized identities.¹ Our theory does not explicitly incorporate power dynamics (Bowleg & Bauer, 2016), but introduces an interactional lens to understanding how women are perceived as they age and the consequences of those perceptions. We implicate middle age as a critical phase, recognizing that age is a liminal yet impermanent series of stages through which all people pass (North, 2019).

Prototypical career progress involves gaining expertise and status with age (e.g., Altonji & Pierret, 2001), attributes that are associated with higher agency (Goller & Billett, 2014). This progress should occur for both men and women professionals in the transition from young adulthood to middle age; and we theorize that, just like men, women will be perceived as more agentic as they gain work experience and status as they age. But agentic women will also be perceived as less prototypical and as violating rule-based gender stereotypes (Rudman, 1998). And, if women are judged to be more agentic in middle age as they gain expertise and status compared to when they are younger, perceivers may see them as less prototypical. This, we suggest, includes seeing them as less warm. Being seen as less warm, in turn, may have unique negative consequences for middle-aged professional women.

In parallel, older people are seen as less warm when their behavior is interpreted as failing to cede pragmatic and symbolic resources to younger generations such as not retiring from coveted jobs (North & Fiske, 2013). This penalty seems to target older men more than older women, because older men historically have controlled more resources and are seen as a bigger impediment to younger people who seek to acquire those resources (Martin et al., 2019). By contrast, this line of work has largely overlooked how similar penalties may target *middle-aged* women compared to men. As such, it is also less clear whether the perceptions of warmth and agency that can trigger negative consequences for women vary according to their age.

We develop and test a theory of how perceptions of professional women and men change as they age and the consequences of those perceptions. We suggest that middle-aged women, like men, are perceived as displaying higher levels of agency compared to when they are younger, which grants middle-aged workers the higher status that they typically are due relative to younger workers because of their more advanced career stage. We also suggest that women will be judged as declining more in warmth between young adulthood and middle-age than men, consistent with the importance of warmth as a gender intensified prescription for women (Prentice & Carranza, 2002). As such, we suggest that professional women are perceived as the most counter-stereotypic and consequently, judged as performing worse in middle-age compared to when they were younger.

To test our theory, we conducted an experiment (Study 1), examining whether middle-aged professional women (but not men) are

¹ Intersectionality theory recognizes that people are members of multiple social categories, that these combined categories influence their lived experience, that category membership is, in turn, linked to systemic power structures and inequality (Bowleg & Bauer, 2016), and that different categories may be evoked by different contexts, situations, and traits, and are therefore fluid and dynamic (Else-Quest & Hyde, 2016). In recognizing that people exist at the intersection of many different social identity groups, Moss-Racusin (2021, p. 1437) states, “the historical under-investigation of intersectional identities has contributed to the ‘intersectional invisibility’ of many individuals.” Thus, intersectional research explores the individual and combined effects of membership in multiple social categories, whether “additive” (e.g., “double jeopardy,” Lincoln & Allen, 2004) or multiplicative (e.g., Hancock, 2007).

perceived to be significantly higher in agency and lower in warmth compared to young-adult women (and men) with identical profiles; a cross-sectional investigation (Study 2) to conceptually replicate the experiment and consider the generalizability of the phenomenon by examining professionals working in a wide range of industries; and a longitudinal within-person archival field study (Study 3) showing how perceptions of *the same woman or man* changed as they aged, and specifically that middle-aged women professors, but not men, were penalized in their performance evaluations by being viewed as less warm. We also show that perceptions of middle-aged women’s lower warmth (but not men’s) partially mediates the link between age and performance evaluations.

Our theory offers at least three new insights into research on gender inequality. First, we examine how perceptions of women’s behavior is uniquely shaped by their age and influences perceptions of their performance at work (e.g., Martin et al., 2019). Although research has established the universality of warmth and competence (Cuddy et al., 2008), our theory is among the first to identify the unique influence of age on these social perceptions for professional women. Second, we identify *when* working women are most likely to be perceived as maximally incongruent with sex-role stereotypes. They are most agentic and least warm when they are middle-aged. Third, we consider the consequences of these perceptions for women’s careers, offering a novel explanation for why gender inequality persists (Smith-Barrow, 2015). Our theory highlights middle age (which naturally is often closely correlated with mid-career; Grandey et al., 2020) as a perilous time for professional women, in which the potential for both career attainment and negative perceptions derived from stereotype incongruity are peaking.

We begin by discussing the relative importance of agency and warmth as primary dimensions of social perception and normative expectations for both men and women. Next, we articulate our theory that professional women are perceived differently as they age, with increasing perceptions of agency coupled with decreasing perceptions of warmth as they reach middle-age. Finally, we discuss how these evolving perceptions make middle-aged women especially susceptible to negative consequences such as lower performance evaluations.

1. Gender stereotypes: describing prescribing warmth in women

People consider another person’s intentions to try to determine whether that person will be warm, trustworthy, cooperative, and honest. The opposite or the absence of these attributes signals competitiveness, or an intention to put one’s own needs above those of others. People also want to know if others have the capacity to enact their intentions toward us—or whether others are agentic, intelligent, skilled, and confident (Fiske et al., 2002).

Research has consistently shown that women are expected to be warm, nurturing, and communal, while men are expected to be agentic, competent, and assertive (e.g., Eagly & Karau, 2002), perhaps traceable to historical divisions of labor tied to women’s unique biological capacity to bear children and men’s capacity, based on generally greater physical strength, to protect and provide for their family (e.g., Hyde, 2005). The functional relevance of these caregiving versus breadwinning roles is largely obsolete in modern society, but the associated stereotypes linger because they increase predictability in behavior and maintain the existing power structure, which has been dominated by men (e.g., Carothers & Reis, 2013).

Stereotypes are not just descriptive, they are also prescriptive, regulating how members of social groups are allowed to, or prohibited from, behaving (Prentice & Carranza, 2002). As such, women are penalized for violating pervasive gender stereotypes because they are seen as threatening the existing social order (e.g., Ridgeway & Correll, 2004). Agency is associated with high status, and agentic women are especially likely to be perceived as threatening to power structures dominated by men (e.g., Acker, 2006), men’s primacy as earners (Eckes,

2002), and, in organizations, the prevalent “think manager, think man” association (Schein & Davidson, 1993). For example, when study participants were primed with a threat to their national political system (Rudman et al., 2012) or an upward social comparison (Parks-Stamm et al., 2008), they essentialized gender differences (e.g., Brescoll et al., 2013), and endorsed prescriptive gender stereotypes more strongly (Brescoll & LaFrance, 2004). Conversely, violating these stereotypes evokes social and economic reprisals against agentic women (Rudman, 1998).

Experiments have shown that women who violate gender stereotypes are generally viewed as less likable (Heilman & Okimoto, 2007), receive negative evaluations and lower hiring recommendations (e.g., Phelan et al., 2008), are considered worse leaders (Livingston et al., 2012), and are viewed as less worthy of political support (Okimoto & Brescoll, 2010). Further, both men and women raters are equally likely to sanction agentic women (e.g., Rudman & Glick, 1999), with penalties ranging from disapproval (Rudman et al., 2012) to a lower likelihood of being hired (e.g., Bowles et al., 2007). Taken together, this research shows that women working in high-status roles are censured for engaging in the agentic behavior mandated by their role because these behaviors violate feminine niceness prescriptions (Rudman & Phelan, 2008). This work has assumed, however, that stereotype incongruities and the associated penalties are static across women’s lifetimes, which masks potential variability in perceived gender stereotypicality at different points in the lifespan.

2. The age-gender interaction: middle-aged women maximize stereotype incongruity

We consider how two social categories interact to influence how women are perceived as they age. In our interactional approach, we suggest that middle-aged professional women make up a category that has been overlooked. To date, research on women’s career progress has focused on cross-sectional perceptions of and, only occasionally, outcomes for women and men at a particular age. Moreover, prior work has primarily focused on young-adult women – demonstrating, for example, how younger women in simulated impression management dilemmas or entry-level job interviews face negative perceptions for being agentic (e.g., Phelan et al., 2008). But women in young-adulthood may be perceived differently than middle-aged women—less agentic and warmer than middle-aged women—making them more congruent with prevailing stereotypes. In contrast, middle-aged women are thought to be more agentic but less warm (Fiske, 2017).

Over the course of their careers, and as they age, workers tend to increase their agency by accruing the experience and knowledge necessary to achieve mastery in their roles (Goller & Billett, 2014) and potentially, to advance to leadership positions. Increasing agency typically corresponds to increases in status and respect, which are consistent with gender stereotypes among men (Ridgeway, 2001). For women, however, increasing agency creates a gap between their prescribed status and their achieved status (e.g., Joshi et al., 2015).

Eagly and Karau’s (2002) role congruity theory suggests that because agentic women are perceived to have deviated from central gender role prescriptions (that women should be less agentic and more communal), they are viewed as discrepant, atypical, and counter-stereotypical (e.g., Chatman et al., 2008). Once women are perceived as counter-stereotypical, perceivers also reason that they are likely to be less communal too (e.g., Okimoto & Brescoll, 2010). As Phelan and Rudman (2010, p. 807) highlight, “rule-based gender stereotypes stipulate acceptable behaviors for women and men and when these rules are violated, perceivers react negatively.” Existing research illustrates this phenomenon, primarily using vignette studies (e.g., Livingston et al., 2012) and confederates enacting scripts (Rudman, 1998), showing that agentic women are consistently seen as less warm and less likable than are less agentic women—and agentic men.

Women who seek and inhabit masculine-typed roles, such as high-

status professions and powerful leadership roles within organizations, must display agency to qualify for and perform effectively in these positions (e.g., Rudman & Phelan, 2008), but such perceptions of increased agency tend to decrease women’s, but not men’s, likability, that is, warmth and agency serve compensatory functions and, perceptually, are inversely related in women but not men (e.g., Holoien & Fiske, 2013). Specifically, when women are perceived as agentic, particularly in professional roles, they are also seen as men’s competitors.² We suggest that it is in middle-age that women’s rising status and demonstrated experience pose the greatest threat to the existing gender hierarchy, which has historically granted men more power and status than women.

Using an interactional lens allows us to identify the points in a woman’s lifespan when she is perceived to violate gender stereotypes most acutely, and to understand the consequences that arise from viewing women as less communal as they increase in agency. We propose that women are viewed as more and less conforming to the gender intensified communality prescription at different points in their lives, given distinct age-based prescriptions that interact with gender (e.g., Francioli & North, 2021). We therefore theorize that professional women are perceived as maximally counter-stereotypic in middle-age, compared to in young-adulthood, because of their heightened agency (Fiske, 2017). While men can be both agentic and warm, agentic women are necessarily viewed as less warm due to the gender role incongruity that their agency poses (e.g., Heilman & Okimoto, 2007).

Building on the wealth of literature documenting how women are perceived negatively and penalized for being agentic (for a meta-analysis, see Williams & Tiedens, 2016), we expect perceptions of women’s warmth to drop more steeply than men’s as they move from young-adulthood to middle-age. It is critical to establish first, whether middle-aged women are perceived as peaking in agency compared to when they were younger, and compared to men, and second, whether their heightened agency in middle-age comes along with perceptions of lower warmth. Thus, we predict that both professional men and women will be perceived as increasing in agency from young adulthood to middle age, but that perceptions of women’s warmth will decline more as they move from young-adult to middle age than perceptions of men’s warmth will, making middle-aged women higher in agency and lowest in warmth compared to young-adult women.

Hypothesis 1 (H1). *Professional men and women will be perceived as increasingly agentic from young adulthood to middle age.*

Hypothesis 2 (H2). *Professional men and women will be viewed differently in warmth in middle age. As professional women move from young-adulthood to middle age, perceptions of their warmth will decrease more than will perceptions of men’s warmth.*

3. Consequences of differential perceptions of middle-aged professional women

Given intensified gender prescriptions for women to be higher in warmth (Prentice & Carranza, 2002), if women are perceived as decreasing in warmth as they move from young adulthood to middle age, they may experience consequences, including social and economic penalties for defying a key stereotypic expectation (Eagly & Karau, 2002). Research has shown that violating feminine niceness prescriptions makes it less likely that women will be evaluated as hireable (e.

² Rosette et al. (2016: 431) further distinguish agency based on competence, or skills and talents that reveal a person’s functional and instrumental capabilities to accomplish individual and organizational goals, from agency based on dominance, defined as interpersonal displays of ambition, assertiveness, and power. Importantly, Rosette et al. (2016) acknowledge that both competence and dominance are incongruent with gender stereotypes of women, and thus both types of agency have the potential to trigger perceptions that an agentic woman is less communal.

g., Bowles et al., 2007) or having leadership potential (e.g., Livingston et al., 2012). Thus, being perceived as incongruent with the gender intensified prescription of warmth impacts women's career advancement.

If perceptions of women are more influenced by their age as compared to men, the interaction of these attributes is particularly important for women's hierarchical advancement and equality at work. Initial findings paint a somewhat dire picture, at least when it comes to older women in the workplace. At the level of entry, audit studies—in which researchers mail out matched resumes that vary by age (young/old) and gender (men/women)—find that older women fare the worst in obtaining callbacks (e.g., Neumark et al., 2015). Supporting this, sociological perspectives argue that older women face “double jeopardy,” in which aging incurs greater penalties among women than it does among men (e.g., Handy & Davy, 2007).

Nevertheless, how age-gender interactions generate consequences for middle-aged working women remains a largely open question. Recent research indicates that women's careers are significantly shorter than are men's in fields such as academia (Huang et al., 2020), but it remains unclear if this pattern stems from middle-age or older-age afflictions. Gender aside, middle-age tends to correlate with maximal status, resources, influence, and agency—both in perception and reality (e.g., North & Fiske, 2012). Because middle-age is the most prototypical age in the category of “working professional” (Krueger et al., 1995), and professional women are seen as more agentic and less warm (Fiske et al., 2002), and because warmth is a gender intensified prescription (Prentice & Carranza, 2002), it stands to reason that middle-aged women might be at particular risk for warmth deficit penalties.

Support for the idea that middle-aged women experience the most significant consequences for stereotype violations arises from comparing them to younger and older women. For example, compared to younger women, middle-aged women are perceived as less feminine (e.g., Singh & Young, 1995), less attractive (Pliner et al., 1990), and less likable (Fiske, 2017). Compared to older women, middle-aged women are seen as less gender-role-congruent because they do not yet qualify for the more communal stereotypes associated with old-age (Brewer et al., 1981) or the “invisibility” that provides older women with the leeway to act with agency (Martin et al., 2019).

Taken together, we suggest that middle-aged professional women, because of their perceived increase in agency coupled with a decrease in warmth, will be evaluated as less effective in their professional roles compared to when they are younger and older.³ First, we predict that middle-aged women will be evaluated as performing worse than when they are younger (and older) even when doing the same work across both periods. Second we predict that, as professional women move from young adulthood to middle age, they will be evaluated as performing worse compared to men during the same period. And third, we predict that there will be a curvilinear relationship between women's age and performance evaluations with women being evaluated as performing better when they are younger and older rather than middle-aged, and that this will be accounted for by middle-aged women's perceived warmth deficits. Finally, we predict that the higher agency of middle-

³ We distinguish our theory from the so-called “motherhood penalty” (e.g., Cuddy et al., 2004), in which middle-aged women are penalized for their perceived fulfillment of feminine niceness prescriptions. While motherhood corresponds to stereotypes of women having high warmth and low competence, in line with traditional patriarchal beliefs that women are nurturing, we predict a drop in middle-aged women's performance evaluations that occurs because women at this high-status stage are perceived to be in violation of feminine niceness prescriptions.

aged women will lead to higher perceptions of warmth deficits, which will mediate the relationship between women's but not men's age and performance evaluations.⁴

Hypothesis 3A. *Middle-aged women's job performance will be evaluated more negatively compared to the performance of younger (and older) women.*

Hypothesis 3B. *The effect of age on evaluated performance will be more negative as women move from young adulthood to middle age compared to the same period for men.*

Hypothesis 4. *The effect of age on perceived warmth deficits will mediate the curvilinear relationship between women's age and performance evaluations such that perceived warmth deficits will account for women's lower evaluations in middle age.*

Hypothesis 5. *For women (but not men), aging from young adulthood to middle age will be associated with higher perceptions of agency and, in turn, higher perceptions of warmth deficits. This serial path will mediate the relationship between women's age and performance evaluations.*

We test Hypotheses 1 and 2 in Studies 1 and 2, and all six Hypotheses in Study 3. We pre-registered our studies and provide data and other materials at the Open Science Foundation: <https://osf.io/ntvgs/> and AsPredicted: <https://aspredicted.org/x4re4.pdf>.

4. Study 1: perceptions of agency and warmth in men and women across ages

We designed Study 1 to test our predictions that, while men and women will both be perceived as increasing in agency from young adulthood to middle age (H1), only women will be perceived as less warm between the two ages (H2). Participants read detailed information about a fictitious man or woman target working in the technology industry corresponding to a pre-tested photo of that person.⁵ The most stringent test of our hypotheses examines whether *the same person* is viewed as changing in their levels of agency and warmth as they age. Therefore, we measured age-based changes in targets' perceived agency and warmth, comparing each target's “current self” to their “past” or “future” self. This simulated within-person test allows us to examine whether women and men are subject to different perceptions as they age.

4.1. Participants and design

We recruited a sample of 999 U.S. residents from the survey site Prolific ($M_{age} = 34.0$, $SD = 12.0$ (range = 18–79), 53% women) for our 2 (man/woman target manager) \times 2 (middle-aged target imagined as young-adult/young-adult target imagined as middle-aged) experimental design. We counterbalanced the study so half of the participants viewed a middle-aged, 46-year-old man or woman target, and were asked to imagine that same person's younger, 29-year-old self, while the remaining half of participants viewed a younger, 29-year-old man or woman target, and were asked to imagine the target's middle-aged, 46-year-old self. We chose these ages based on prior research delineating young adulthood and middle age (Grandey et al., 2020).

4.2. Procedure and materials

We adapted a paradigm originally developed by Heilman and

⁴ Stronger evidence for these consequences has been observed among agentic White women than among agentic Black women (Livingston et al., 2012). Thus, as a reasonable starting point, we focus on the gender-age interaction while holding race constant or controlling for race.

⁵ Please see Supplemental Materials for stimuli pilot (S1.A), and pilot results of each photo (S1.B and Table S1.B1).

Early-adult Photos



Middle-aged Photos



Fig. 1. Photo stimuli of early-adult and middle-aged and working professionals used in Study 1.

Okimoto (2007). Participants were asked to imagine that they were employees at a technology company called C.A.S. Corporation, and provided with details about the company's products, culture, and history. Participants were told that, "Your manager is part of a leadership training program and is reflecting on their career trajectory—the growth, changes, and challenges they have experienced or will experience over time. An outside consultant is asking for anonymous, upward feedback to share with your manager about how you perceive their characteristics and abilities over time." Participants then viewed an individual headshot photo of their supervisor, a Product Manager named Steve Wilson (man) or Sue Miller (woman), in professional attire (see Fig. 1). The photo was accompanied by a detailed profile containing information about the manager's education, past work experience, and interests. Participants then read a letter from C.A.S. Corporation's CEO to the product management division, giving additional information about Steve Wilson's or Sue Miller's experience and qualifications.

Next, participants were asked to "please compare [Steve Wilson/Sue Miller] as [he/she] is now, at age [29/46], compared to when [he/she] [will be age 46/was age 29]. In considering [Steve/Sue] at these different ages, which characteristics [will be/were] more true of [Steve/Sue] [in middle age (age 46)/at a younger age (age 29)], compared to [Steve/Sue] now (age [29/46])?" (-3 = younger extremely more, -2 = younger mostly more, -1 = younger barely more, 1 = middle-aged barely more, 2 = middle-aged mostly more, 3 = middle-aged extremely more). We counterbalanced the order of the choice set to avoid order effects. Before analyzing and reporting these responses, we reverse-coded them so that higher numbers in our analyses always indicate greater stereotype endorsement for the middle-aged group relative to the younger group.

Participants rated the targets on the 40 characteristics of the Bem Sex Roles Inventory (SRI) (1974), which is composed of masculine (agency) and feminine (warmth) scales, including terms like "forceful" or "gentle", respectively. We tested the agency ($\alpha_{\text{woman targets}} = 0.93$, $\alpha_{\text{man targets}} = 0.92$) and warmth ($\alpha_{\text{woman targets}} = 0.91$, $\alpha_{\text{man targets}} = 0.90$) scales (20 items each).

4.3. Results

See Table 1 for Study 1 descriptive statistics and correlations. Supporting our prediction in H1, participants rated both the woman target ($M = 0.60$, $SD = 1.26$), $t = 10.56$, $p < .001$, $d = 0.47$) and the man target ($M = 0.42$, $SD = 1.22$), $t = 7.67$, $p < .001$, $d = 0.34$) as increasing in agency as they aged (see Fig. 2).⁶ To test H2, we conducted a one-way ANOVA predicting target warmth, with condition (man vs woman) as the between-subjects factor. We found a main effect of condition, $F(1, 997) = 6.72$, $p = .010$, $\eta_p^2 = 0.007$, indicating that changes in perceptions of woman targets' warmth as they age are greater than changes in perceptions of man targets' warmth as they age. Participants perceived the woman as decreasing in warmth as she moved from young adulthood to middle age ($M = -0.22$, $SD = 1.07$), $t = -4.64$, $p < .001$, $d = 0.21$, but did not perceive the man as changing in warmth as he aged ($M = -0.05$, $SD = 1.03$), $t = -1.07$, $p = .285$, $d = 0.05$ (see Fig. 2), supporting H2.

4.4. Discussion

Study 1 provides an initial demonstration that, while both men and women managers are perceived as becoming more agentic from when they are younger to when they are middle-aged, only women are viewed as becoming less warm than men from young adulthood to middle age. Consistent with our predictions, we uncovered a significant interaction between gender and age in which a woman manager, whose perceived agency increased just like an identical man, was perceived as declining in warmth as she aged. The male manager was not perceived to decline in warmth. This suggests a perceptual change that, due to the experimental design, occurs absent any actual changes in the woman's behavior, or differences in her background or personality. This study had the advantage of holding constant all aspects of the target except for

⁶ We also broke down the agency items by dominance and competence, consistent with Rosette et al. (2016). As expected, the identical pattern emerged for both dimensions of agency (see Supplemental Materials S1.C).

Table 1
Descriptive statistics and correlations among Study 1 variables.

| | M | SD | 1 | 2 | 3 | 4 |
|--|------------------------|------|---------|---------|-------|---------|
| (1) Condition (0 = male manager, 1 = female manager) | 50% male 50% female | – | – | | | |
| (2) Participant gender (0 = male, 1 = female) | 47% male 53% female | – | 0.03 | – | | |
| (3) Participant age | 34.0 | 12.0 | 0.01 | 0.00 | – | |
| (4) Agency | 0.51 | 1.24 | 0.07* | 0.10** | –0.02 | – |
| (5) Warmth | –0.14 | 1.06 | –0.08** | –0.11** | 0.03 | –0.35** |

* $p < 0.05$, ** $p < 0.01$.

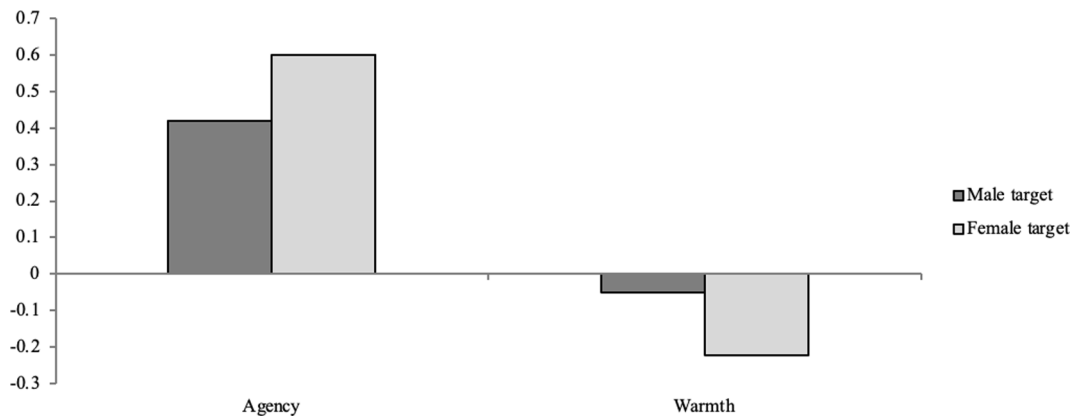


Fig. 2. Agency and warmth ratings comparing an early-adult vs middle-aged man and woman, in Study 1. Positive numbers indicate greater stereotype endorsement for the middle-aged target.

their gender and age, and using random assignment, allowing us to identify a causal relationship between the age-gender interaction and perceptions of warmth. That said, the hypothetical nature of the vignette raises questions about external validity, which we address in Study 2.

5. Study 2: comparison of cross-sectional perceptions of men and women professionals

In Study 2, we conducted a between-subjects investigation of how different targets were evaluated by different evaluators, cross-sectionally. We used a field setting of professionals working in a wide range of industries and roles.

5.1. Research context, sample, and variables

We obtained data on 476 students enrolled in a selective public West Coast university MBA-degree program’s executive leadership class, taught by the same instructor, between 2012 and 2019. At the time of admission, 26% of participants worked in the technology industry, 17% in consulting, 13% in financial services, 10% in healthcare, 6% in the non-profit sector, 4% in retail and hospitality, 4% in energy, 4% in government, with the remaining 12% in manufacturing, education, sports and entertainment, advertising, and law. Industry was unknown for 5%. The professional MBA students (77% of the sample) had an average of 9.5 years of post-university work experience and full-time MBA students (23% of the sample) had an average of five years of post-university full-time work experience.

Dependent Variables: Agency and Warmth. As part of a leadership class assignment, participants invited colleagues to evaluate them on several measures. Participants were instructed to select cross-evaluators who were current or former colleagues (e.g., co-workers, bosses, or subordinates) and who were familiar with how the focal participant typically behaved in professional settings. Of the 3,194 unique evaluations, 50% of ratings were completed by peers, 16% of ratings were completed

by managers, 13% of ratings were completed by subordinates, 13% of ratings were completed by classmates, 2% were completed by friends of the focal participants, and 6% were unknown. Each participant was evaluated by an average of seven raters (minimum number of raters = 1, maximum number of raters = 31).

Participants were evaluated on several characteristics. To assess agency, we created a 4-item composite variable ($\alpha = 0.70$) based on prior research (e.g., Abele et al., 2008): assertive, dominant, forceful, and self-confident. We used the single-item measure of agreeableness from the Ten Item Personality Inventory (TIPI - Gosling et al., 2003), “Sympathetic, Warm,” to measure warmth. Items were measured on a 1 (very inaccurate description of this person) to 7 (very accurate description of this person) Likert-type scale.

For each variable of interest, we performed hierarchical linear regression analysis via the “lmer” function in R. Since participants were rated by multiple evaluators, we created a dummy variable in which each unique target was given an ID number. We included a random intercept for target ID in all our regression analyses.

Independent Variables - Gender and Age. We obtained participants’ self-reported gender from the school’s database, with 64% identifying as men (0) and 36% identifying as women (1). We calculated age by subtracting each participant’s birth year from the year in which they took the course. Women participants ranged in age from 23 to 55 years ($M = 34.1$, $SD = 5.8$) and men ranged from 25 to 58 ($M = 35.4$, $SD = 6.6$). We treated participant age as a continuous variable and employed a linear model to test H1 and H2.⁷

Control Variables. We controlled for a number of variables in the

⁷ Although the maximum age in this sample is 58 years, there were only seven women (and 23 men) older than 46 (our middle-age foci in Study 1). We did not truncate the sample and examined age as a continuous variable. We also ran all of our analyses using logged age and our pattern of results remained the same.

Table 2
Descriptive statistics and correlations among Study 2 variables.

| | M | SD | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|--|------|--------|---------|---------|--------|---------|--------|---------|
| (1) MBA program type (0 = full-time MBA, 1 = MBA for working professionals) | 23% full-time MBA 77% MBA for working professionals | – | | | | | | | |
| (2) Target U.S. citizenship (0 = U.S. citizen, 1 = non-U.S. citizen) | 61% U.S. citizen 39% non-U.S. citizen | – | –0.01 | | | | | | |
| (3) Target race (0 = White or Asian, 1 = Black, Hispanic, or Native American) | 89% ethnic majority student 11% URM student | – | 0.03 | 0.16** | | | | | |
| (4) Year of course | – | – | 0.61** | 0.11** | –0.02 | | | | |
| (5) Target gender (0 = male, 1 = female) | 64% male 36% female | 0.48 | 0.03 | –0.11** | –0.01 | –0.02 | | | |
| (6) Target age | 34.75 | 6.08 | 0.40** | –0.07** | –0.09** | 0.17** | –0.13** | | |
| (7) Agency | 4.55 | 1.09 | –0.00 | –0.03 | 0.05** | –0.03 | 0.02 | 0.06** | |
| (8) Warmth | 5.45 | 1.39 | 0.02 | 0.06** | 0.01 | 0.03 | 0.05** | 0.01 | –0.21** |

* $p < 0.05$, ** $p < 0.01$.

Table 3
Hierarchical linear model of target gender and target age predicting perceived agency in Study 2.

| | Model 1: Control variables | Model 2: Control variables with independent variables |
|---|-------------------------------|--|
| Intercept (Agency) | 34.436 (36.172) | 29.033 (36.158) |
| MBA program type (0 = full-time MBA, 1 = MBA for working professionals) | 0.071 (0.099) | –0.017 (0.108) |
| Target U.S. citizenship (0 = U.S. citizen, 1 = non-U.S. citizen) | –0.088 (0.067) | –0.078 (0.067) |
| Target race (0 = White or Asian, 1 = Black, Hispanic, or Native American) | 0.141 (0.106) | 0.155 (0.106) |
| Year of course | –0.015 (0.018) | –0.012 (0.018) |
| Target gender (0 = male, 1 = female) | – | 0.041 (0.068) |
| Target age | – | 0.012* (0.006) |
| Observations | 3,193 | 3,193 |
| Number of Targets | 476 | 476 |

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.01$.

regression analyses to rule out alternative explanations: We coded subjects' race (0 or 1) as White or Asian (89%) or Black, Hispanic, or Native American (11%); nationality as U.S. citizens (61%) or non-U.S. citizens (39%); the year that the participant was enrolled in the course (21% in 2012, 5% in 2014, 11% in 2015, 19% in 2016, 21% in 2017, 8% in 2018, and 15% in 2019); and enrollment in the MBA Program for Working Professionals (1) and Full-Time MBA Program (0).

5.2. Results

Table 2 displays descriptive statistics and correlations among study variables. Model 1 in Table 3 shows that none of the control variables influenced perceptions of participants' agency. Model 2 in Table 3 reveals a main effect for participant age ($\beta = 0.012, p = .045$); older participants were perceived as higher in agency compared to younger participants, offering support for H1. As expected, no main effect for participant gender emerged ($\beta = 0.041, p = .540$).

Model 1 in Table 4 shows that non-U.S. citizens were perceived as warmer than U.S. citizens. Model 2 in Table 4 shows a main effect of participant gender ($\beta = 1.083, p = .014$); women were perceived to be warmer than men overall. A significant main effect of age indicated that older participants were perceived as warmer ($\beta = 0.016, p = .030$). And, as predicted, a significant interaction between participant gender and participant age in predicting perceived warmth ($\beta = -0.027, p = .034$ – see Fig. 3) emerged, supporting H2. Contrary to H2, however, the simple slope effect for women was not significant ($\beta = -0.010, p = .342$), but it

Table 4
Hierarchical linear model of target gender and target age predicting perceived warmth (as measured by the item "Sympathetic, Warm") in Study 2.

| | Model 1: Control variables | Model 2: Control variables with independent variables with two-way interaction |
|---|-------------------------------|---|
| Intercept (Warmth) | –16.620 (41.010) | –19.371 (40.789) |
| MBA program type (0 = full-time MBA, 1 = MBA for working professionals) | 0.066 (0.112) | 0.005 (0.122) |
| Target U.S. citizenship (0 = U.S. citizen, 1 = non-U.S. citizen) | 0.162* (0.076) | 0.179* (0.076) |
| Target race (0 = White or Asian, 1 = Black, Hispanic, or Native American) | –0.004 (0.120) | 0.005 (0.120) |
| Year of course | 0.011 (0.020) | 0.012 (0.020) |
| Target gender (0 = male, 1 = female) | – | 1.083* (0.440) |
| Target age | – | 0.016* (0.008) |
| Target gender x Target age | – | –0.027* (0.013) |
| Observations | 3194 | 3194 |
| Number of Targets | 476 | 476 |

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.01$.

was for men ($\beta = 0.016, p = .030$), and in a positive direction, indicating that middle-aged men were perceived as warmer than young-adult men but that there was no difference in warmth perceptions for younger compared to middle-aged women. Fig. 3 shows the relationship and, interestingly, though perceptions of women's warmth across the age span visibly declines, the relationship is not significant. Thus, H2 is only partially supported; men and women are viewed as differing in warmth as they age, but it is men, and not women who are perceived as changing more, and in a positive direction.⁸

5.3. Discussion

Study 2 provides further support for H1, that middle-aged

⁸ These findings were robust to excluding control variables: For perceived agency, the main effect for participant age became marginally significant ($\beta = 0.010, p = .062$). When controlling for the year that the participant was enrolled in the course, the main effect for participant age significantly predicted perceived agency ($\beta = 0.011, p = .037$). For perceived warmth, upon dropping our controls, the interaction between participant gender and participant age remained significant ($\beta = -0.028, p = .027$).

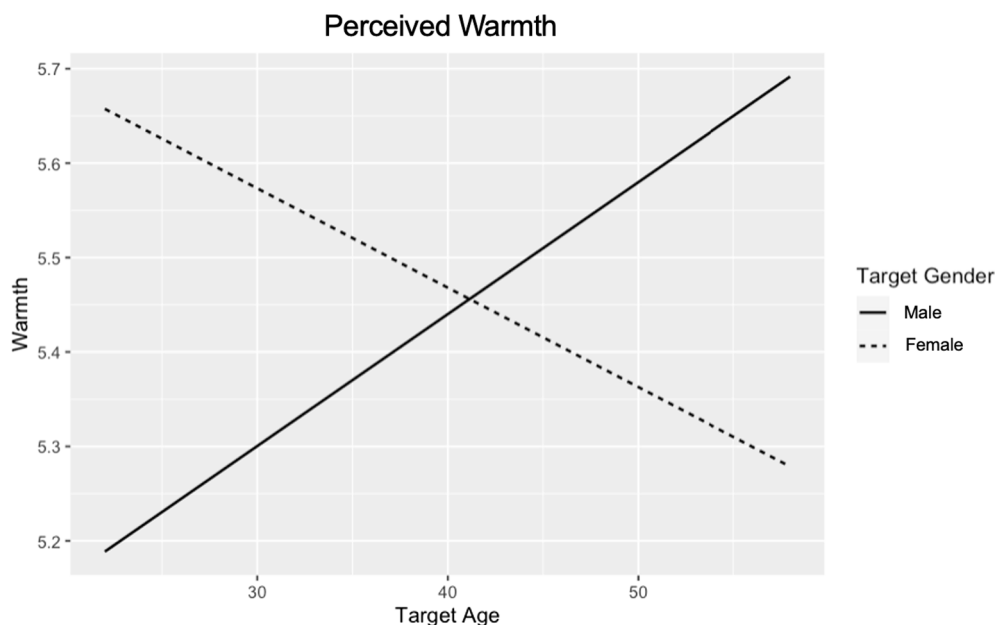


Fig. 3. Target age and target gender predicting perceived warmth in Study 2.

professionals are seen as higher in agency compared to younger professionals while H2, that gender and age interact to influence perceived warmth, was only partially supported. On the one hand, men and women were judged differently at different ages in their level of warmth, but the test for H2 in Study 2 did not follow our prediction that women would be judged as declining in warmth; instead, men were viewed as *increasing* in warmth as they aged while there was no change in perceptions of women's warmth as they aged (in the range that existed in this sample for women).

Given the downward trend for women from younger-adult to middle-age shown in Figure 3, H2 might have been confirmed in a larger sample and particularly-one with more female targets, who made up only 36% of participants in this sample. It is also worth noting that targets selected their own evaluators and as such, may have received more favorable warmth evaluations. Thus, our results may be more conservative (e.g., less stark for women) than if cross-evaluators were not chosen by targets. These results still offer some support for our theory in that they show that age-based prescriptions operate distinctly for women and men, and an expectation that people will generally increase in warmth as they age, with women being seen as failing to conform to this expectation (Chopik & Kitayama, 2018).

6. Study 3: within-person performance evaluations of men and women professors

6.1. Research context, sample, and procedure

In Study 3 we examined how evaluations of the same target person compare between young adulthood and middle age (like Study 1) and how evaluations of younger women (men) as a group compare to evaluations of middle-aged women (men) as a group (like Study 2), using a longitudinal, within-person analysis of the teaching evaluations of tenure-track faculty. These professors were working at a selective U.S. business school and rated by their Masters of Business Administration (MBA) students over a 15-year period. This setting represents a profession dominated by men: among the top 20 MBA programs, the proportion of women professors ranges from 16 to 29 percent (Financial Times Global MBA Ranking, 2019). During the relevant data analysis period, the percentage of women tenure-track professors was comparable at this business school, ranging from 19 to 25 percent.

We chose teaching evaluations to test our hypotheses pertaining to

evaluated performance because professors teach comparable courses (often the same course over many years), are evaluated by multiple raters using the same metrics over multiple years, are evaluated using a standardized assessment, and are assessed with teaching evaluations that are consequential for their career outcomes⁹ (Simpson & Sigauw, 2000). Students are asked to anonymously evaluate each of their professors at the end of each course and professors cannot view course evaluations until they have submitted their students' grades. Importantly, teaching requires a skill set that typically improves with experience, but is also one in which the specific skill set demanded does not change markedly over time, in contrast to professions that require different skill sets over the course of a career (Sanner & Bunderson, 2015).¹⁰ Our general expectation is that, if perceptions of gender-stereotype incongruity and associated penalties in terms of lower performance evaluations did not exist, both women and men professors would be evaluated as having about the same or perhaps higher teaching performance as they increased in age and experience; at least those evaluations should not decline markedly from young adulthood, when professors are just starting their careers, to middle age, when they have gained significantly more extensive teaching experience. Since students' perceptions of professor likability predict higher ratings of teaching effectiveness (Delucchi, 2000), the perceived communal deficit of middle-aged women compared to men evident in Study 1 and partially

⁹ Despite ongoing debates about the validity of teaching evaluations—questioning whether they in fact measure student learning (Greenwald, 1997; Stark & Freishtat, 2014)—they remain consequential for faculty, both in terms of career advancement and personal well-being (Stark & Freishtat, 2014). Publications continue to carry the most weight in academic faculty advancement decisions (e.g., Tang & Chamberlain, 1997), but teaching effectiveness constitutes at least 26% of the total tenure decision (May 2005), a meaningful proportion even in research-focused universities. Top-tier business schools are even more focused on faculty teaching performance because their revenue models depend on satisfied MBA students paying premium tuition to attend (e.g., Bonsoms, 2016). Moreover, for the focal business school in Study 3, administrators reward faculty who achieve higher teaching scores with higher pay and research resources. Even if their research records are stellar, faculty whose teaching evaluations fall below the required level are not paid this premium. Thus, teaching performance, though not the only metric of performance for professors, is a significant factor for faculty in most top tier business schools, and the one that was the subject of Study 3.

in Study 2 is likely to correspond to the lowest teaching ratings for middle-aged women.

Because teaching evaluations have a unique impact on the status and performance reviews of tenure track professors, and because these professors are comparable in being employed full-time by a single university, we excluded lecturers, graduate student instructors, and other non-tenure-track instructors from our analysis. To maintain comparability in teaching level, we restricted our analysis to teaching evaluations for courses in the school’s MBA-degree programs, which included full-time students (61% of courses), and part-time students (working professionals, 39% of courses). Each professor taught an average of 11.2 courses during the span of the dataset ($SD = 9.3$) over an average range of 6.1 years ($SD = 4.0$). The database we constructed, using records kept by the school’s academic dean’s office, contains digitized information about 126 professors who taught 1402 courses between Fall 2003 and Spring 2017, all relatively evenly distributed across six subject areas (e.g., finance, marketing). Our dataset comprises 59,600 student ratings ($M = 42.0$ ratings per class, $SD = 15.5$).

We estimated models using random effects with clustered standard errors regression (see Supplemental Materials S3.A). We also report results from fixed effects models in the Supplemental Materials (S3.B and Table S3.B1). Because of the proportion of professors that at some point in the study would be categorized as older adults, and older adults are viewed as warmer than other ages (Fiske, 2018), we tested for a curvilinear effect in which evaluated teaching effectiveness was the dependent variable, and gender, squared age, and their interaction were the independent variables.

We also analyzed MBA students’ open-ended, free-response comments using Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2015) analyzing words that pertained to an agentic or communal deficit category about their professors to test H1, H2, H4 and H5. We created separate dictionaries, consistent with previous studies, for positively valenced agency- and negatively valenced warmth-related words (e.g., Lindholm & Yzerbyt, 2018). The variables pertaining to the text analysis, positive agentic language and niceness deficit, were the dependent variables with respect to the gender-age interaction (H1 and H2) and mediator variables linking the gender-age interaction to performance evaluations (H4 and H5).

6.2. Dependent variable: performance evaluation

We operationalized each professor’s evaluation as their average assigned student rating in a given course. We used the school’s database of anonymous student evaluations, collected during the final session of each course, to determine mean teaching evaluation ratings for each of the 1402 courses taught by professors in the sample. We focused on the question that is used in the university’s professor merit and promotion reviews: “Considering both the limitations and possibilities of the subject matter and course, how would you rate the overall teaching effectiveness of this professor?” Students evaluated their professor on a Likert-type scale from 1 (not effective) to 7 (extremely effective) ($M = 5.82$, $SD = 0.82$).

6.3. Independent variables

Gender. Seventy-nine percent of the 126 professors were men (0) and 21% were women (1). Professors in the sample ranged in age from 27 to 65 ($M = 44.8$, $SD = 9.2$). Women professors taught 21% of the 290 courses in our sample. For the linguistic moderated mediation analysis, 81 percent of comments pertained to men professors and 19 percent to women professors. There were no differences between the comment word count for women ($M = 12.39$, $SD = 9.68$) and men professors ($M = 12.18$, $SD = 9.84$), $t(9, 622.2) = 1.50$, *n.s.*, $d = 0.02$.

Age. We calculated professor age at time of teaching each course by subtracting the professor’s birth year from the year in which the course was taught. We then mean-centered the age variable (average age across

Table 5
Comment characteristics and professor demographics in Study 3.

| Characteristics in comment data <i>n</i> = 32,377 comments | Percentage or mean |
|---|--------------------------|
| Professor gender | |
| Male | 75% (<i>n</i> = 56) |
| Female | 25% (<i>n</i> = 19) |
| Mean age | |
| Male | 44.5 |
| Female | 42.0 |
| Average number of comments | |
| Male | 77.2 (<i>SD</i> = 22.2) |
| Female | 74.9 (<i>SD</i> = 23.5) |
| Mean teaching score (7-point scale) | |
| Male | 5.80 |
| Female | 5.52 |
| Tenure status | |
| Tenured | 66% |
| Untenured | 34% |
| Left institution | |
| Remained at institution | 67% |
| Left institution | 33% |
| Evaluation period | |
| Fall 2003 – Spring 2009 | 10 semesters |

all faculty, or 44.83 years) for all reported regressions.

To examine H3A and H3B, we squared the mean-centered age of the professor when the course was being taught,^{10 1112}

6.4. Moderated mediation variables

Communal deficit language and positive agentic language. We operationalized the target evaluation as the language students used to evaluate each professor. The dean’s office transcribed 32,377 student comments made about 75 professors who were a part of the longitudinal analysis, over the course of 10 semesters teaching 500 courses, from Fall 2003 to Spring 2009, (after 2009 the handwritten forms were given directly to faculty without transcribing them). Descriptive statistics for the moderated mediation analysis data can be found in Table 5; descriptive statistics of the variables in the longitudinal analysis and the moderated mediation analysis are compared in the Supplemental Materials (S3.D).

The LIWC program computes the percentage of words from a given unit of analysis that fall into a predefined linguistic category. Our unit of analysis was each individual student’s entire written comment (average word count = 12.22 words) about a focal professor. We adapted the LIWC custom dictionary which comprises 85 words relating to agency and communality (Madera et al., 2009), in two ways. First, we expanded

¹⁰ We mean-centered to shift the parabola, centering the curvature at the mean age. If age was not centered before squaring, the curvature would be restricted to continually linear increases or decreases in teaching score as a person’s age advances, precluding a test of our hypotheses that age has a nonlinear effect on scores and specifically that women’s teaching evaluations have a U-shaped distribution in which teaching evaluations are higher at younger and older ages than in middle-age, while men’s do not.

¹¹ While our theory is primarily concerned with changes in perceptions of women and men from young adulthood to middle age, previous research has suggested that older women are seen as higher in warmth and lower in agency, unlike those in middle-age or young-adulthood (Fiske, 2018). Because we did not want to remove any of the data years of the professors in the sample, even when they exceeded middle-age (30% of the sample), we expected agency to exhibit an inverted U-shaped pattern and warmth to exhibit a U-shaped pattern as professors aged, which we modeled accordingly.

¹² An alternative approach for categorizing careers among academic, tenure-track positions would be to focus on rank, rather than age (assistant professor, associate professor, full professor). A detailed consideration of this possibility, along with an analysis of these variables, can be found in Supplemental Materials (S3.C).

the dictionary to detect terms commonly found in professor comments using MAXQDA's frequency evaluation tool (Kuckartz & Rädicker, 2019). This produced a comprehensive list of 608 adjectives that appeared at least once in the professor comments. Two independent coders, blind to the study hypotheses, assigned a linguistic category of agentic or niceness deficit to each word (ICC = 0.91). This produced a final dictionary of 373 words (including 85 words from Madera et al. (2009) and 288 words derived from the actual teaching comments about professors). We operationalized agency as the presence of positive agentic words (114 words including assertive, bright, and competent) and feminine niceness deficit as the count of negative communal words (84 words including callous, harsh, and insensitive), for men and women professors. We transformed these data into a binary format by converting all nonzero values (in which the LIWC output was greater than 0, meaning the identified words were present) to 1 while keeping all zero values, indicating that the review did not contain any agentic or communal language. Positive agentic and communal deficit consisted of 13,039 and 1,031 comments respectively.

6.5. Control variables

Core or elective course. We created a dummy variable to indicate whether a course was required in the core MBA curriculum (0), or the course was an optional elective course (1). Fifty-four percent of courses were required classes.

Qualitative or quantitative course. We created a dummy variable to indicate whether a course focused on quantitative content (0), such as finance and accounting, or more qualitative material (1), such as marketing and management. Seventy percent of courses were quantitative.

Program. We created a dummy variable for each type of MBA program to account for average differences in students' age, current work status, and work experience (0 = *working professional MBA*, 1 = *full-time MBA*). Sixty-one percent of courses were in the full-time MBA.

Tenure status. We created a dummy variable for whether the professor was tenured at the time of the course, controlling for career-stage-related factors pertinent to the specific teaching context but not directly to our predictions. Faculty who had not yet been granted tenure at the time of the course were coded as 0 (31%) and those who had tenure were coded as 1 (69%).

Departure from the institution. We created a dummy variable for whether the professor left the institution in any year of the 2003–17 period, as departures could relate to teaching quality. Faculty who did not leave the institution by the end of spring 2017 were coded as 0 (81%), whereas those who left by the end of spring 2017 were coded as 1 (19%).

Nationality. We created a dummy variable for the target's nationality (0 = *not a U.S. citizen*, 1 = *U.S. citizen*). Sixty-seven percent of courses were taught by U.S. citizens, while 33% of courses were taught by non-U.S. citizens. None of the faculty during the study period were considered an underrepresented minority as defined in the U.S. cultural context.

Citation count. We used Google Scholar's profile tool, which catalogs citations per year, to ascertain a professor's citation count, a proxy for research productivity, in the year that he or she taught the focal course ($M = 679.9$, $SD = 1459.3$).

Child care leaves. Women professors' teaching performance could decline in middle-age because child-bearing or child-rearing responsibilities reduced their ability to devote effort to teaching. The proportion of faculty who took the child leave (16% took one leave, 10% took two leaves) is comparable to the proportion of women in the sample, suggesting that, with some women having no children, most of the women took the leave.

6.6. Results

Please see Table 6 for descriptive statistics and correlations among

Study 3 variables.

Random effects analysis. Table 7 displays the base equation (model 1) and the within-person longitudinal test of H3A and H3B (model 2). Among the control variables, teaching in the MBA for Professionals ($\beta = -0.183$, $p < .05$) and citation counts are positively related to teaching evaluations ($\beta = 0.000$, $p < .01$). Relevant to our hypothesis tests, women receive more negative teaching evaluations ($\beta = -0.662$, $p < .001$), and lower evaluations when they are middle-aged compared to when they were younger and when they are older, as indicated by the significant coefficient for squared age ($\beta = -0.002$, $p = .036$). Moreover, the interaction between squared age and gender is significant ($\beta = 0.005$, $p < .001$). The form of this interaction in Fig. 4 reveals a significant decline in women's teaching evaluations from young adulthood to middle age (and a rebound from middle age to older adulthood) supporting H3B, that women, *even when compared to their own earlier and later performance on a learned skill*, are rated the lowest when they are middle-aged. This contrasts with the pattern for men professors, as demonstrated in Table 7 and Fig. 4, whose teaching evaluations *increase* from young adulthood to middle age, supporting H3A. Cross-sectional and fixed effects analyses revealed the same pattern of results, which we report in S3.E and S3.B and Table S3.B1, respectively. It is worth noting that our predicted results emerge even including child-leaves in the models. Specifically, the number of child leaves taken did not significantly predict teaching evaluations ($\beta = -0.052$, $p = .570$).

Moderated Mediation analysis: To examine H4, we first tested whether professor gender and professor age-squared interacted to influence the likelihood of professors receiving positive agency and communal deficit comments. Given the open-ended nature of the prompts and the wide range of topics that students could comment on, a number of comments were not captured by our linguistic categories. Additionally, since only a subset of the data contained written comments, we only had an average of three years of linguistic data for our sample of 76 professors. Taken together with the fact that our data contained a large proportion of zeros, we did not have adequate statistical power to conduct longitudinal analysis to examine changes over a given professor's lifetime. Therefore, we used logistic regression analysis, which can account for large proportions of zeros arising when none of the communal deficits or positive agentic items are contained in a student's evaluation (e.g., Hoetker, 2007), in a cross-sectional analysis, to estimate the probability of communal deficits and agentic language appearing in a given student comment as a function of the professor's squared age \times gender interaction.

Agency comments. To model the distribution of age in this data set (given the presence of younger, middle-aged, and older professors), we examined whether a curvilinear relationship exists between target age and perceived agency, testing H1. Table 8 shows the results of our cross-sectional logit analysis. Model 1 shows the base equation estimating an individual professor's probability of receiving positive agentic comments as an intercept; teaching a core course, a quantitative course, teaching in the MBA for Professionals Program, and being tenured are associated with receiving fewer positive agency comments. Model 2 includes the independent variables of age, gender, and age-squared. In testing H1, squared age was significantly associated with a higher likelihood of positive agency comments ($\beta = -0.001$, $p < .001$) such that middle-aged professors had a higher likelihood of receiving agency comments compared to young-adult (and older) professors (see Fig. 5A). Notably, and as expected, a main effect of gender did not emerge ($\beta = -0.043$, $p = .192$).

Communal deficit comments. In testing H2, Table 9 shows the results of our cross-sectional logit analysis. Model 1 shows the base equation estimating an individual professor's probability of communal deficit language as an intercept. Communal deficit comments were more common among instructors teaching elective courses or full-time MBAs, as well as those who were not tenured at course time, and those who left the university during our sample period. Model 2 includes the independent variables of age, gender, and squared age. Model 3 includes the

Table 6
Descriptive statistics and correlations among Study 3 variables.

| | <i>M</i> | <i>SD</i> | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|--|--|--|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| (1) MBA program type (0 = <i>full-time MBA</i> , 1 = <i>MBA for working professionals</i>) | FTMBA: 61% MBA for working professionals: 39% | – | – | | | | | | | | | | | | |
| (2) Target U.S. citizenship (0 = <i>non-U.S. citizen</i> , 1 = <i>U.S. citizen</i>) | Non-U.S. citizen: 33% U.S. citizen: 67% | – | –0.10** | – | | | | | | | | | | | |
| (3) Course requirement (0 = <i>required course</i> , 1 = <i>elective course</i>) | Elective: 46% Core: 54% | – | 0.02 | 0.03 | – | | | | | | | | | | |
| (4) Course type (0 = <i>qualitative course</i> , 1 = <i>quantitative course</i>) | Qual: 30% Quant: 70% | – | 0.01 | 0.07** | 0.18** | – | | | | | | | | | |
| (5) Tenure status (0 = <i>not tenured at course time</i> , 1 = <i>tenured at course time</i>) | Assistant: 31% Tenured: 69% | – | –0.09** | 0.26** | –0.04 | 0.07 | – | | | | | | | | |
| (6) Left institution (0 = <i>did not leave institution</i> , 1 = <i>departed institution</i>) | Stayed: 81% Left institution: 19% | – | –0.02 | –0.10** | –0.16** | –0.20** | –0.32** | – | | | | | | | |
| (7) Citation count | 679.80 | 1459.26 | –0.02 | 0.12** | 0.02 | 0.09** | 0.24** | –0.09** | – | | | | | | |
| (8) Number of child leaves | 0.30 | 0.64 | 0.07** | –0.16** | 0.15** | –0.01 | –0.40** | 0.04 | –0.14** | – | | | | | |
| (9) Target gender (0 = <i>male</i> , 1 = <i>female</i>) | Male: 79% Female: 21% | – | 0.09** | 0.09** | –0.16** | –0.11** | –0.09** | 0.19** | –0.09** | 0.15** | – | | | | |
| (10) Target age | 44.83 Male: 45.00 Female: 44.19 | 9.17 Male: 9.26 Female: 8.80 | 0.03 | 0.33** | –0.03 | 0.14** | 0.62** | –0.23** | 0.35** | –0.35** | –0.04 | – | | | |
| (11) Target squared age | 2094.06 Male: 2110.68 Female: 2030.16 | 863.07 Male: 870.48 Female: 832.31 | 0.04 | 0.34** | –0.04 | 0.14** | 0.59** | –0.22** | 0.34** | –0.34** | –0.04 | 1.00** | – | | |
| (12) Agentic positive (<i>N</i> = 32,377 comments) | 0.40 | 0.49 | 0.01* | 0.02** | –0.02** | –0.02** | –0.00 | 0.00 | 0.00 | 0.01 | 0.02** | –0.02** | –0.02** | – | |
| (13) Communal deficit (<i>N</i> = 32,377 comments) | 0.03 | 0.18 | –0.01** | 0.01 | 0.03** | 0.00 | –0.00 | 0.01 | 0.00 | –0.01 | –0.00 | 0.00 | 0.00 | –0.06** | – |
| (14) Average teaching score | 5.82 Male: 5.87 Female: 5.62 | 0.82 Male: 0.81 Female: 0.84 | 0.07* | –0.01 | –0.03 | –0.03 | 0.00 | –0.04 | 0.05 | –0.06* | –0.12** | –0.14** | –0.15** | 0.07** | –0.05** |

* $p < 0.05$, ** $p < 0.01$.

Table 7
Random effects model (using clustered standard error estimates) of professor and course characteristics predicting teaching evaluations in Study 3.

| | Model 1: Controls | Model 2: Controls with IVs | Model 3: Controls with IVs and interaction |
|---|----------------------------|---------------------------------------|---|
| Intercept (Average teaching score) | 5.947*** (0.179) | 5.759*** (0.219) | 5.786*** (0.218) |
| MBA program type (0 = MBA program for working professionals, 1 = full-time MBA program) | -0.087 (0.093) | -0.175* (0.080) | -0.183* (0.076) |
| Target U.S. citizenship (0 = non-U.S. citizen, 1 = U.S. citizen) | 0.023 (0.140) | 0.130 (0.138) | 0.087 (0.133) |
| Course requirement (0 = required course, 1 = elective course) | -0.046 (0.120) | -0.101 (0.116) | -0.138 (0.114) |
| Course type (0 = qualitative course, 1 = quantitative course) | -0.173 (0.139) | -0.130 (0.157) | -0.111 (0.154) |
| Tenure status (0 = not tenured at course time, 1 = tenured at course time) | -0.060 (0.126) | 0.110 (0.143) | 0.166 (0.145) |
| Left institution (0 = did not leave institution, 1 = departed institution) | 0.031 (0.197) | 0.028 (0.193) | -0.047 (0.183) |
| Citation count | 0.000 (0.000) | 0.000**¹ (0.000) | 0.000** (0.000) |
| Number of child leaves | -0.046 (0.077) | -0.065 (0.087) | -0.052 (0.092) |
| Target age | - | -0.022 (0.012) | -0.018 (0.012) |
| Target gender (0 = male, 1 = female) | - | -0.282 (0.152) | -0.662*** (0.181) |
| Target squared age | - | -0.001 (0.001) | -0.002 (0.001) |
| Target age × Target gender | - | - | -0.029 (0.018) |
| Target squared age × Target gender | - | - | 0.005*** (0.001) |
| Observations – Courses | 1402 | 1402 | 1402 |
| Groups – Professors | 126 | 126 | 126 |
| Adjusted R ² | 0.0125 | 0.0767 | 0.1083 |

Clustered standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

¹ The precise estimate is 0.00005612, and is the effect *per citation*. So, for example, if someone had 2000 citations, $2000 \times 0.00005612 = 0.112$, which is a significant effect on a 7-point scale.

interaction term, squared age and gender. A main effect of gender ($\beta = 0.283, p = .027$) indicated that women were more likely to receive communal deficit comments, while no main effect of squared age emerged ($\beta = -0.001, p = .182$).

Most germane to testing H2, a significant interaction emerged between squared age and gender ($\beta = -0.003, p = .025$) showing that middle-aged women were more likely to receive communal deficit comments compared to younger women to a greater extent than when comparing middle-aged men to younger men. Middle-aged women were also perceived as less warm (more communal deficit comments) compared to young-adult (and older) women. Fig. 5B shows a curvilinear pattern, in which the probability of communal deficit language increases for women between young-adulthood and middle-age (and then decreases for women in older adulthood; $\beta = -0.004, p = .005$). In contrast, the probability of communal deficit language remains relatively constant for men as the slope for squared age for men does not change significantly across various ages ($\beta = -0.001, p = .182$). Thus, H2 is supported in Study 3.¹³

¹³ For teaching evaluations, the interaction between gender and squared age remained significant ($\beta = 0.004, p = .002$) upon dropping all control variables. For perceived agency, the main effect of squared age remained significant ($\beta = -0.001, p < .001$). For perceived warmth, the interaction between participant gender and squared age remained significant ($\beta = -0.003, p = .025$).

Moderated mediation analyses of communal deficits. Binary mediators are not permitted in moderated mediation analysis (Hayes, 2013), so we used the proportion of comments falling into the communal deficit category for a given professor. We calculated these proportions by summing the total number of comments a professor received for a given course number that fell into the communal deficit category and dividing each sum by the total number of comments the professor received for the given course. Thus, we examined 497 unique instances (the proportion of comments a professor received for the communal deficit category) where the unit of analysis was broken down by the professor, professor age, and the course.

We ran a moderated mediation model using Hayes (2013) PROCESS macro (Model 8 with 5,000 resamples) to test whether gender moderates the relationship between squared age and increased perceptions of communal deficit. We entered squared age as the independent variable, target gender as the moderator, proportion of comments containing communal deficit language as the mediator, and mean teaching evaluation score as the dependent variable. The direct effect of squared age on the proportion of comments containing communal deficit language was significant for women (effect = -0.0001, SE = 0.0000, $p = .014$) but not for men (effect = 0.0000, SE = 0.0000, $p = .424$), indicating that middle-aged women, but not middle-aged men, were significantly more likely to receive comments containing communal deficit language compared to young-adult and older-adult professors. The proportion of comments containing communal deficit language significantly mediated the effect of squared age on performance evaluations for women (indirect effect = 0.0009, SE = 0.0003, 95% CI [0.0003, 0.0017]) indicated by the confidence interval, which does not include zero, but not for men (95% CI [-0.0002, 0.0004]).

The direct and indirect relationship between women’s squared age, the proportion of comments containing communal deficit language, and women’s performance evaluations are shown in Fig. 6A. The regression coefficient represents the direct relationship between squared age and mean performance evaluation score decrease (from 0.201 to 0.053) when the mediator (the proportion of comments containing communal deficit language) was included in the model. Moreover, when we included communal deficit language as a mediator, the conditional direct effect of squared age on performance evaluations became insignificant for women (indirect effect = 0.0006, SE = 0.0010, 95% CI [-0.0015, 0.0026]), but not for men (indirect effect = -0.0019, SE = 0.0004, 95% CI [-0.0027, -0.0011]), as shown in Fig. 6B. The regression coefficient representing the direct relationship between squared age and mean performance evaluation score remains largely unchanged and significant (from -0.246 to -0.260) when the proportion of comments containing communal deficit language was included in our model. These findings demonstrate that the proportion of comments containing communal deficit language significantly mediates women’s performance evaluations, supporting H4 and offering insight into the mechanism—perceptions that middle-aged women (but not women in young- and late-adulthood) lack communality, which, in turn, drives lower performance evaluations.¹⁴

Serial mediation analysis. To test H5, we ran a serial mediation model using Hayes (2013) PROCESS macro (Model 6 with 5,000 resamples) separately for men and women. We entered squared age as the independent variable, proportion of comments containing positive agentic language as the first mediator, proportion of comments containing communal deficit language as the second mediator, and mean teaching evaluation score as the dependent variable.

For women, the indirect effect of the full serial mediation model was not significant (indirect effect = 0.0000, SE = 0.0001, 95% CI [-0.0002,

¹⁴ We conducted exploratory analysis to see if agency comments mediated between age and performance evaluations, for both men and women targets, and they did not. Please refer to S3.F, S3.G, S3.H, and S3.I in our Supplemental Materials.

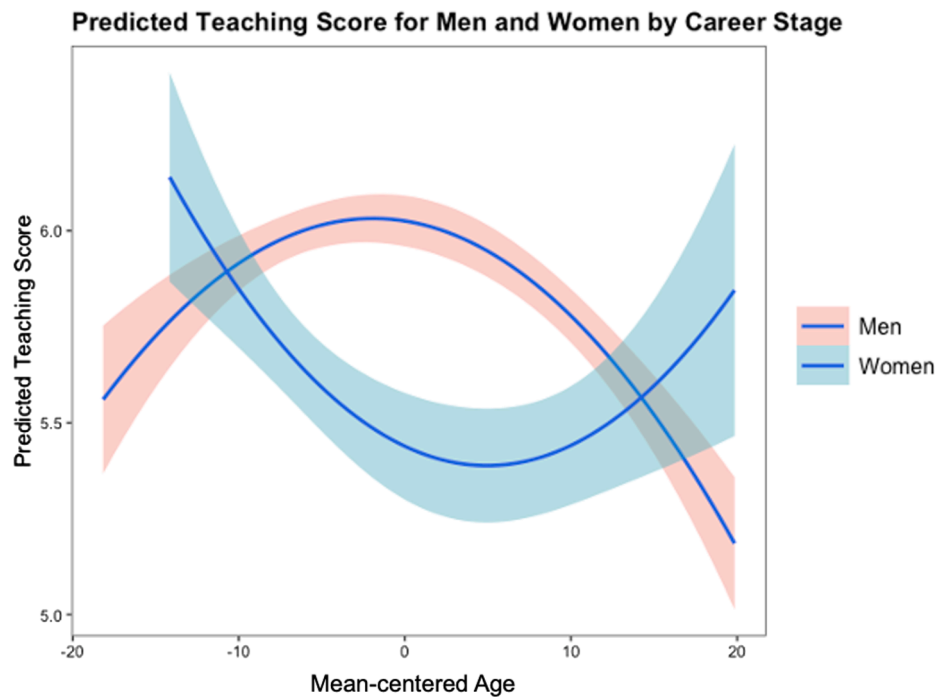


Fig. 4. Clustered standard error graph of predicted teaching score, by professor age and gender, in Study 3.

Table 8

Cross-sectional differences in positive agentic perceptions of men and women in Study 3 using logistic regression.

| | Model 1: Controls | Model 2: Controls with independent variables |
|---|----------------------|--|
| Intercept (Positive agentic language) | -0.218*** (0.044) | -0.196*** (0.050) |
| MBA program type (0 = MBA program for working professionals, 1 = full-time MBA program) | -0.076** (0.026) | -0.115*** (0.027) |
| Target U.S. citizenship (0 = non-U.S. citizen, 1 = U.S. citizen) | 0.120*** (0.026) | 0.194*** (0.029) |
| Course requirement (0 = required course, 1 = elective course) | -0.112*** (0.026) | -0.119*** (0.027) |
| Course type (0 = qualitative course, 1 = quantitative course) | -0.113*** (0.026) | -0.127*** (0.026) |
| Tenure status (0 = not tenured at course time, 1 = tenured at course time) | -0.078* (0.032) | -0.030 (0.039) |
| Left institution (0 = did not leave institution, 1 = departed institution) | -0.052 (0.030) | -0.054 (0.031) |
| Citation count | 0.000 (0.000) | 0.000 (0.000) |
| Number of child leaves | 0.050* (0.024) | 0.034 (0.025) |
| Target age | — | -0.007** (0.002) |
| Target gender (0 = male, 1 = female) | — | -0.043 (0.033) |
| Target squared age | — | -0.001*** (0.000) |
| Target age × Target gender | — | — |
| Target squared age × Target gender | — | — |
| Observations – Evaluations | 32,377 | 32,377 |

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

0.0001]). Only the indirect relationship between squared age, proportion of comments containing communal deficit language, and teaching evaluations reached significance (indirect effect = 0.0010, SE = 0.0005, 95% CI [0.0002, 0.0021]). For men, the indirect effect of the full serial mediation model was also not significant (indirect effect = 0.0000, SE = 0.0000, 95% CI [-0.0001, 0.0000]) while only the indirect relationship

between squared age, proportion of comments containing positive agentic language, and teaching evaluations reached significance (indirect effect = -0.0004, SE = 0.0001, 95% CI [-0.0007, -0.0001]). These results suggest that for women, being middle-aged is associated with higher perceived communal deficits, which corresponds with lower performance evaluations. By contrast, for men, being middle-aged is associated with higher perceived agency, which is associated with higher teaching evaluations.

6.7. Discussion

Study 3 explored whether women would be evaluated as performing worse in middle age compared to men (H3A) and to their younger (and older) selves (H3B). Our random effects analyses supported H3A and our fixed effects analyses further compared women and men professors to their own earlier and later performance. For women, evaluation scores of their teaching—the same activity that they engaged in and were evaluated similarly to men earlier in their careers—declined from young adulthood to middle age, supporting H3B, whereas men did not experience this decline, supporting H3A. Contrary to the idea that teaching ability is a skill that is likely to improve with experience, women were uniquely evaluated as significantly worse teachers even after they had gained additional experience and were perceived as more agentic.

The cross-sectional linguistic analysis showed that middle-aged women were perceived as more agentic compared to young-adult (and older-adult) women, lending support to H1. Consistent with our theory, middle-aged women were described as more deficient in communality compared to women who were younger (or older), reflected in the significant increase in negative communal language appearing in middle-aged woman professors' comments compared to the comments younger (and older) women professors received, supporting H2. In contrast, perceptions of men professors' warmth were unrelated to their age. It is notable that perceptions of agency are highest for women when they are middle-aged. Along with our moderated mediation analyses, this pattern suggests that ratings of agency in middle age, which men also experience, cannot directly account for the negative consequences that middle-aged women uniquely face. Instead, we found that perceived communal deficits mediated lower teaching evaluations for

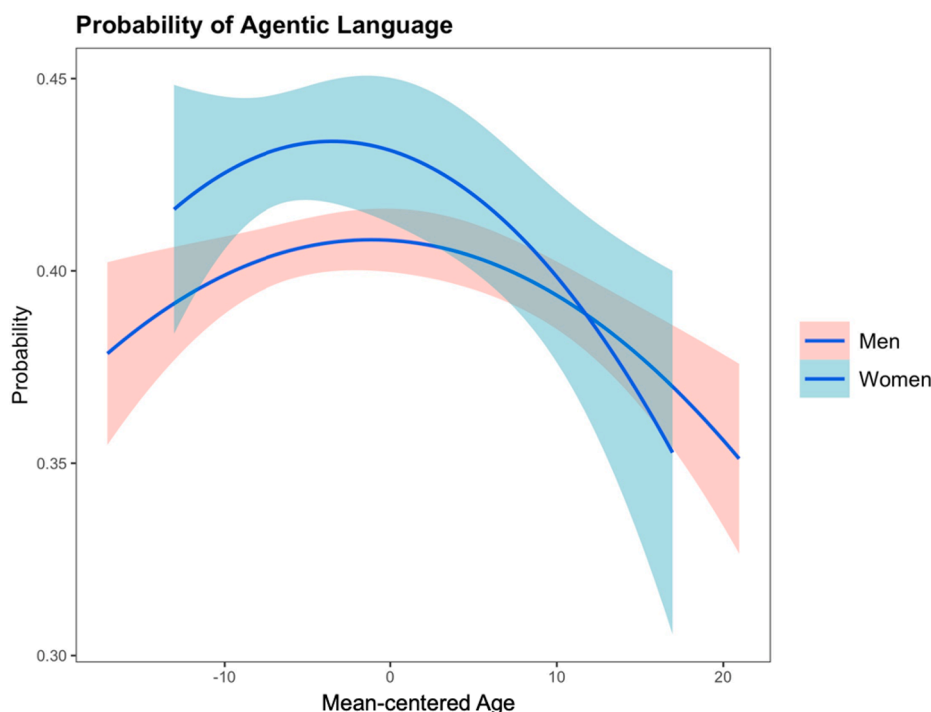


Fig. 5A. Graph of probability of positive agentic language for male and female professors by age in Study 3 (confidence intervals indicated in pink and blue).

professional women in middle age, supporting H4, but did not mediate between age and performance evaluations among men.

Surprisingly, we did not detect evidence supporting H5, which predicted a serial mediation. Instead, we found evidence suggesting that men were more likely to be perceived as highly agentic in middle age, driving up their teaching evaluations, while women were more likely to be perceived as high in communal deficits (lower in warmth) in middle age, driving down their teaching evaluations. Thus, while the serial mediation proposed in H5 was unsupported, this analysis reveals how middle age invokes competing stereotypes as a function of target gender, such that for women, middle age is associated with lower perceived warmth, while for men, it is associated with higher perceived agency. It also shows that these perceptions have consequences in terms of differences in men's and women's performance evaluations.

A possible explanation for why we did not detect evidence for serial mediation could be that warmth is considered the primary dimension of social perception, followed by agency (Fiske et al., 2002; Fiske et al., 2007) and that warmth is judged before agency. As such, stereotypes about middle-aged women's low warmth may lead to lower performance evaluations, even in the context of high agency, which cannot compensate for perceived communal deficits. Thus, while it is possible to capture the inverse relationship between warmth and agency in experimental settings that manipulate agency and measure warmth, in real-world scenarios such as judgments of professors, observers may form warmth judgments before they begin to assimilate information or draw conclusions about agency. The real-world context of in Study 3 may have made it difficult to detect the hypothesized inverse relationship between warmth and agency for women.

Another possible reason for the lack of serial mediation may relate to our data, in which we examine the proportion of comments that contain only either positive agentic language or communal deficit language in the context of a full course evaluation. By its very nature, the course evaluation invited a broader range of comments than simple warmth or agency judgments, which made the data both more externally valid and likely more "noisy."

Finally, our serial mediation prediction might not hold because the relationship between agency and communality is complex and varies

across settings and attributes. For example, Fiske and colleagues (2007) has suggested that agency and warmth are orthogonal dimensions, with no necessary relationship between a person's level of agency and how much they are liked by others. But, as Williams and Tiedens (2016, p. 167) note, research has shown that groups that are perceived as being in competition with, and threatening to the dominant group are viewed as low in warmth (e.g., Cuddy et al., 2008). Indeed, if agentic women leaders are viewed as highly discrepant, contrast effects would suggest that they would also get low ratings on communality (Ponce de Leon & Rosette, 2022). Adding to this logic, Correll et al. (2020) and others (e.g., Okimoto & Brescoll, 2010) show that it is agentic behaviors such as acting aggressively or seeking positions of power that cause women to be viewed as less likable and warm.

Study 3 limitations include, first, that while our ability to integrate the quantitative teaching ratings with the linguistic data enabled us to test our mediation hypothesis, the data were not linked at the individual rater level and instead, at the level of the class. We assumed that the underlying perceptual process leading individual raters nested within course sections to make the comments we analyzed would be associated with the final ratings made by the total set of raters in those course sections. Still, future research could dispense with anonymous ratings to more tightly link perceptions and performance ratings, and assess our mediation prediction at the individual rater level. Second, correlational mediation models do not allow causal claims or ruling out alternative explanations (Bullock et al., 2010; Fiedler et al., 2011). Finally, the context of academia includes multiple criteria for evaluating performance; we only included perceptual measures of teaching effectiveness. Thus, our ability to determine whether these evaluations ultimately determined key career outcomes is limited (e.g., Rothausen-Vange et al., 2005).

7. General discussion

Research has consistently demonstrated that professional women face unique challenges compared to their male peers. In the current research, we investigated how women professionals are perceived differently as a function of their age, as well as the negative

Table 9
Cross-sectional differences in communal deficit perceptions of men and women in Study 3 using logistic regression.

| | Model 1: Controls | Model 2: Controls with independent variables | Model 3: Controls with independent variables and interaction |
|---|-----------------------------|---|--|
| Intercept (Communal deficit language) | -3.910*** (0.133) | -3.792*** (0.150) | -3.784*** (0.150) |
| MBA program type (0 = MBA program for working professionals, 1 = full-time MBA program) | 0.195** (0.075) | 0.225** (0.077) | 0.246** (0.078) |
| Target U.S. citizenship (0 = non-U.S. citizen, 1 = U.S. citizen) | 0.110 (0.073) | 0.096 (0.082) | 0.100 (0.083) |
| Course requirement (0 = required course, 1 = elective course) | 0.438*** (0.077) | 0.451*** (0.079) | 0.442*** (0.079) |
| Course type (0 = qualitative course, 1 = quantitative course) | -0.036 (0.073) | 0.014 (0.076) | 0.002 (0.079) |
| Tenure status (0 = not tenured at course time, 1 = tenured at course time) | -0.036 (0.093) | -0.203 (0.112) | -0.240* (0.114) |
| Left institution (0 = did not leave institution, 1 = departed institution) | 0.172* (0.086) | 0.213* (0.088) | 0.182* (0.090) |
| Citation count | 0.000 (0.000) | 0.000 (0.000) | 0.000 (0.000) |
| Number of child leaves | -0.160* (0.071) | -0.187* (0.074) | -0.183* (0.074) |
| Target age | - | 0.016** (0.006) | 0.016* (0.006) |
| Target gender (0 = male, 1 = female) | - | 0.078 (0.093) | 0.283* (0.128) |
| Target squared age | - | -0.001* (0.000) | -0.001 (0.000) |
| Target age × Target gender | - | - | 0.005 (0.012) |
| Target squared age × Target gender | - | - | -0.003* (0.001) |
| Observations – Evaluations | 32,377 | 32,377 | 32,377 |

Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

consequences of those changing perceptions that middle-aged women experience in evaluations of their work. Our findings are generally consistent with our theory that gender and age interact such that middle-aged women are at heightened risk of negative career consequences. We offer multi-method evidence for our theory indicating a more complex interplay between gender and age than previous research has considered, controlling for differences in men’s and women’s behavior (Study 1), examining many focal men and women of varying ages working in many industries (Study 2), and showing that women receive their lowest performance evaluations in middle-age (compared to when they were younger), when they are seen as agentic yet not fulfilling feminine niceness prescriptions (Study 3). Both middle-aged women and men are acknowledged for their agency, but middle-aged women are rated as significantly less effective compared to middle-aged men in actual performance, even when that comparison is made to their own performance when they were younger (and older), on a skill that generally increases with experience.

It is possible that women’s increasing professional stature as they move from young adulthood to middle age may be threatening to a system that has traditionally excluded them from positions of power and status (e.g., Kanter, 1977). As our results generally show, another way to reduce this threat, beyond penalizing women by giving them poorer

performance evaluations (e.g., Rudman & Phelan, 2008), is to question the legitimacy of their success and their trustworthiness by casting them as lower in warmth (Fiske et al., 2002).

7.1. Theoretical contributions

The results from our studies offer several theoretical contributions. First, although not speaking to intersectionality directly (since our current data do not explicitly manipulate power dynamics that inhibit individuals living at intersecting identities; Bowleg & Bauer, 2016), our findings at least introduce an interactional age-gender lens to understanding how women are perceived as they age and the consequences of those perceptions. In so doing, we implicate middle age as a critical phase. Generally, intersectional approaches situate gender in relation to other key social identities that women hold, and acknowledge qualitatively different expressions of womanhood and gender discrimination (Warner et al., 2016). Further, our age-gender interaction approach allowed us to examine whether the relationship between agency and warmth differs as a function of the interplay of these two social categories. Specifically, we discovered that agency and warmth appear to be more related for middle-aged women than for middle-aged men, a finding that may help account for ambiguity in their relationship in prior research (e.g., Wojciszke & Abele, 2008). Given this apparent importance of middle age, future research may more directly explore age-gender intersectionality by viewing age as a power mechanism, and/or ascertaining the liminality of age as an impermanent status, which could alter certain assumptions of intersectionality theory.

Identifying middle age as when women are most susceptible to changing perceptions that generate negative on-the-job consequences surfaces a number of insights. First, it helps to explain why women’s career advancement has stalled even as their overall workplace representation increases; since middle age is when essential opportunities exist to advance to, and be groomed for, the highest status roles (e.g., Simonton, 1988). Second, while older women are known to benefit from “intersectional escape,” middle-aged women may be uniquely vulnerable to “double jeopardy” (i.e., negative consequences for violating age and gender agency prescriptions, e.g., Rosette et al., 2018). Overall, whereas studies on age perceptions tend to cast middle-age as the default, or comparison, our studies highlight that middle age should be a focus for scholars interested in gender inequality at work.

Second, our results generally suggest that perceived violations of feminine niceness prescriptions is a novel explanation for negative performance evaluations among middle-aged women. In Study 1, we focused on a single target’s longitudinal changes in personality, which revealed a pattern in which a younger woman was seen as warmer than the same woman in middle-age. In Study 2, even though cross-sectional perceptions of middle-aged professional women’s warmth were not significantly lower than in young adulthood, perceptions of their warmth failed to increase in a manner that would be consistent with well-established changes in traits like agreeableness as a function of age (Chopik & Kitayama, 2018), while perceptions of men’s warmth did increase as they aged. In Study 3, our longitudinal, within-person analysis of students’ comments about professors suggests that it is middle age itself, rather than the women who happen to be middle-aged, that is the driver of low-warmth perceptions. In contrast to a motherhood penalty prediction in which middle-aged women would be penalized for exhibiting high warmth and low agency (Cuddy et al., 2004), the current findings generally suggest that middle-aged women are perceived as higher in agency but lower in warmth, and are, accordingly, penalized for this maximum stereotype incongruity with lower performance evaluations, even when accounting for their child-bearing status. In contrast to young-adult women, who have not had time to amass the same credentials, women in middle age have marshaled the qualifications to compete for masculine-typed high-status roles, potentially challenging men’s primacy as earners and prevailing power structures (Eckes, 2002).

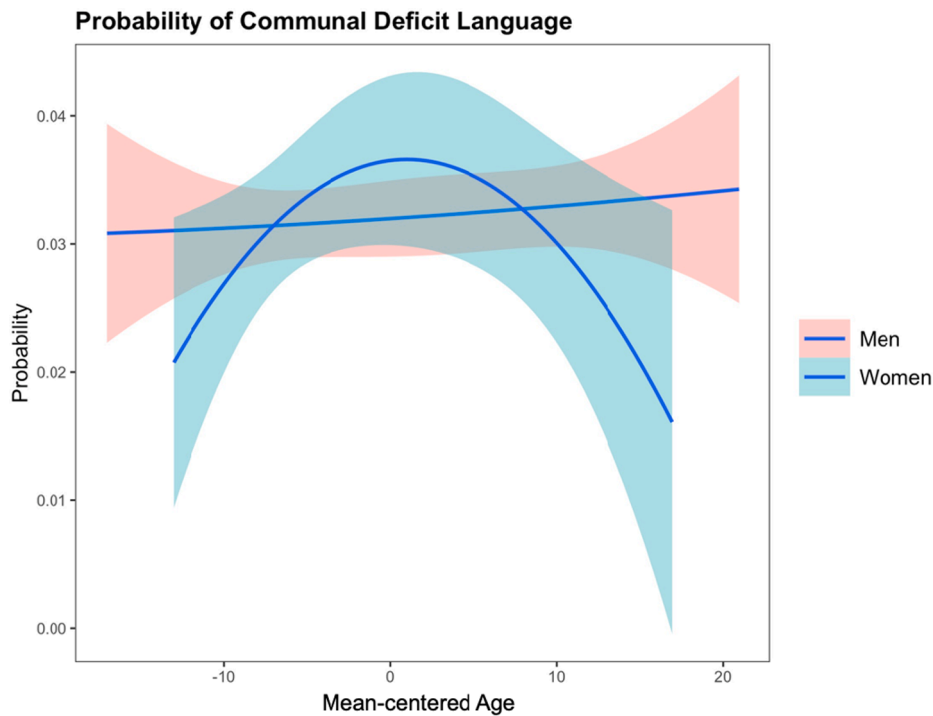


Fig. 5B. Graph of probability of communal deficit language for male and female professors by age in Study 3 (confidence intervals indicated in pink and blue).

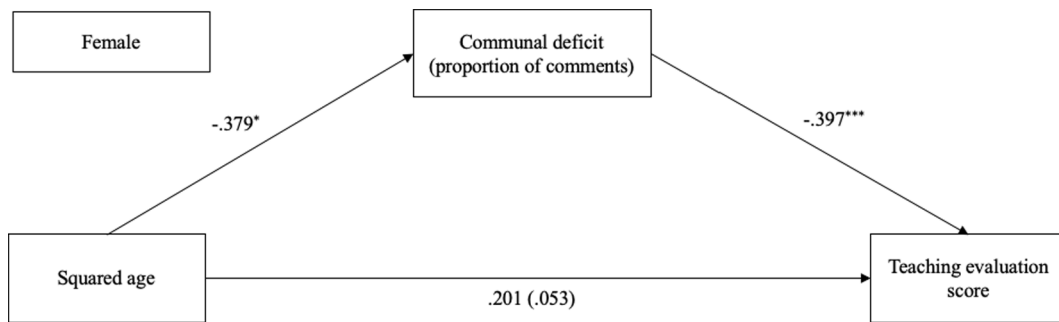


Fig. 6A. Mediation analysis of female professors' proportion of communal deficit comments (Study 3). Coefficients represent unstandardized coefficients using z-transformed variables in linear regression analyses. * $p < .05$. ** $p < .01$. *** $p < .001$.

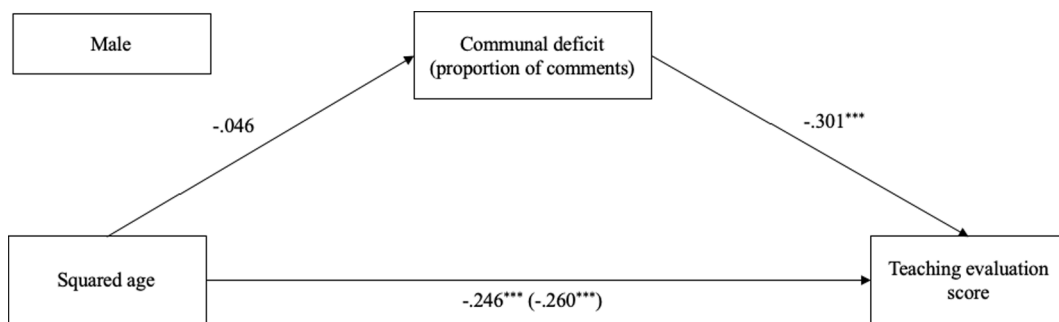


Fig. 6B. Mediation analysis of male professors' proportion of communal deficit comments (Study 3). Coefficients represent unstandardized coefficients using z-transformed variables in linear regression analyses. * $p < .05$. ** $p < .01$. *** $p < .001$.

Third, our findings from Study 3 offer some insight into the consequences of changing perceptions of women as counter-stereotypic over time, going beyond static, cross-sectional findings and experimental vignettes. In the teaching context, one in which knowledge and experience should be a benefit, performance would likely increase or, at the very least, remain relatively steady from young adulthood to middle age.

But, this intuitive pattern emerges for men only, while women are viewed as performing worse in middle age, even accounting for parental status and research productivity, which points to deviation from gender prescriptions as the culprit.

7.2. Study limitations, practical implications, and future directions

We acknowledge a host of limitations in our studies. First, although we examined data from multiple settings—a hypothetical manager in high-technology, business professionals enrolled in an MBA program, and professors in academia—our performance evaluation data are limited to the longitudinal study of professors, which does not offer causal evidence of the relationship between the age-gender interaction, declining warmth ratings in middle-age women, and their lower performance ratings. Future research should undertake longitudinal field experiments in which men and women managers are matched in terms of job roles, age, and experience and tracked over time using objective and subjective evaluations of their work. Experimental work could also manipulate the agency of women and men at different ages (e.g., using actors) to see if different levels of perceived agency influence perceived warmth. To examine the generalizability of our findings beyond business professionals and professors (who are typically perceived as high in agency), future research might examine whether the relationship between perceived feminine niceness deficits and performance evaluations is weaker for women perceived to have lower agency, such as those in non-professional or blue-collar roles.

Second, we tested six hypotheses across multiple studies (ten tests in total). All were supported with two exceptions: H2, that perceptions women's warmth will decrease more than will perceptions of men's warmth from young adulthood to middle age, was supported in Studies 1 and 3, but perceptions of women's warmth did not decrease in Study 2 as predicted, whereas perceptions of men's warmth increased from young adulthood to middle age. That said, the net effect of perceptions of warmth increasing for middle-aged men but not middle-aged women still partially supports our underlying theory that perceptions of men's and women's warmth differ in middle age. Nonetheless, this inconsistency may have arisen from the different methodological approaches and measures we used across studies. Future research should standardize measures of agency, warmth, and performance across research contexts and examine our hypotheses in gender balanced workplaces.

The second exception was that we proposed a serial mediation (H5) that, for women (but not men), aging from young adulthood to middle age will be associated with higher perceptions of agency and, in turn, higher perceptions of warmth deficits. Instead instead we found an unexpected pattern: that men were perceived as highly agentic in middle-age, which corresponded with increasing teaching evaluations, whereas women were perceived as high in communal deficits in middle-age, which corresponded with decreasing teaching evaluations. A number of possibilities could explain why the hypothesized serial mediation was not observed. Empirically, and perhaps most simply, the relationship between warmth and agency varies widely across contexts and research approaches (e.g., ambivalent: Durante et al., 2013; orthogonal: Kervyn et al., 2013; Wiggins, 1979; curvilinear: Imhoff & Koch, 2017; dependent on valence: Suitner & Maass, 2008), even in our investigation, with the r s in our studies being: Study 1: -0.35 , Study 2: -0.21 , Study 3: -0.06 . Further, the significant but small inverse raw correlation between warmth and agency in Study 3 may represent an artifact of the archival data, in which warmth and agency perceptions were derived from a context of broader course evaluations which may have generated more unexplained variance than would direct inquiries about warmth and agency.

Theoretically, a primary explanation for why we did not detect evidence to support H5 could be that warmth is considered the focal dimension of social perception, followed sequentially by agency (Fiske et al., 2007). In real-world scenarios such as judgments of professors, students may fall back on age-gender stereotypes to form warmth judgments before they begin to assimilate information or draw conclusions about agency. Relatedly, field data, with its inherently rich context and repeated interactions, simply may not hold such sharp trade-offs between agency and warmth as those found in experimental research settings. Nevertheless, although the current research does not yet

resolve the puzzle of why women are perceived as decreasing in warmth as they reach middle age, our results clearly show that: 1) women are perceived as decreasing in warmth from young adulthood to middle age; and 2) increases in perceived agency may not be the culprit, as the changes in perceived warmth seem to occur independent of perceived agency. Future research should develop field experiments in which the benefits of random assignment by age and gender could be combined with the external validity of a field setting to examine ratings of men's and women's warmth and agency as they relate to evaluated performance at the two age stages.

A third limitation is that our studies focus primarily on White targets. We were not able to develop a theory of or test a three-way gender, age, race interaction. Given previous research showing that Black women escape dominance penalties incurred by White women (Livingston, et al., 2012) and theories suggesting women of different races face different expectations (e.g., Rosette et al., 2016), it would be useful to conduct additional research that includes other identity differences, such as race, sexual orientation, and disability status, as we expect that these interactions are, like age and gender, also sources of persistent inequity at work.

Finally, our research is unable to rule out the possibility that the perceptual differences among middle-aged professionals reflect reality rather than biased perceptions. The possibility that stereotypes are accurate was first acknowledged by Allport (1954), yet he focused on how qualities that may, on average, be true of a group are exaggerated and overapplied in a pernicious fashion that leads perceivers to be resistant to counter-stereotypical information. Our field data cannot speak to the accuracy of these perceptions (Studies 2 and 3). Yet, given that we find this pattern in hypothetical vignettes in which the targets are described as identical (Study 1), we suggest that these group-based generalizations emerge even in the absence of any behavioral "truth" at the individual-level. Thus our data generally show a novel stereotype that women, more so than men, become less warm as they move from young-adulthood to middle-age.

Although it is possible that this warmth gap in middle-age reflects reality, we know of no developmental theory that could explain why this would be so for women but not men. The idea that women but not men become "objectively" less warm as they age is inconsistent with work on agreeableness, which increases rather than decreases as individuals age (Chopik & Kitayama, 2018). Another possible explanation is that there is a "kernel of truth" deriving from self-fulfilling prophecies (Jussim, 1986). If perceivers expect an individual to be low in warmth, then perceivers may act less warm to targets, which draws out less warmth from targets. In other words, if there is an observable warmth deficit in middle-aged women, it is possible that it reflects the more negative treatment they experience from those who expect them to be unkind. Future research should seek behavioral data in a controlled setting to test whether middle-aged women are less warm compared to young-adult women and men.

Despite these limitations, our findings offer practical insights into how gender inequality at work might be remedied. Specifically, by identifying middle age as a period in which women are uniquely susceptible to perceptions that they are failing to satisfy gender stereotypical prescriptions about warmth, managers can develop strategies for surfacing such stereotype prescriptions and offering explicit support for women at this career stage. For example, organizations can alert evaluators of middle-aged women of the systematic difference in perceptions uncovered here and offer training and tools to enable evaluations to focus more on work-relevant expectations than gender prescribed stereotypes, thereby examining performance more objectively (e.g., Correll, 2017).

8. Conclusion

Professional women must show masculine agency to get ahead—but, in so doing, are perceived as lacking feminine warmth (Rudman &

Phelan, 2008). Our findings generally show that this balance appears particularly challenging for women in middle-age, not necessarily because of their own behavior but because of how they are perceived as they gain agency. Ultimately, this paper introduces a timely, interactional, and multi-method approach to investigating those changing perceptions of women as they age and the consequences of those perceptions in which women in middle-age perceived as maximally incongruous with the gender-intensified prescription of communality, in the context of their earned agency. Recognizing this age-based challenge unique to women should advance scholarly insight into social category interactions, and inform practical solutions as women of all ages strive to overcome the constraints to attaining the highest-echelon roles in professional settings.

CRedit authorship contribution statement

Jennifer A. Chatman: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Investigation, Validation, Visualization, Writing - original draft, Writing - review & editing. **Daron Sharps:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. **Sonya Mishra:** Data curation, Formal analysis, Methodology, Validation, Visualization. **Laura J. Kray:** Conceptualization, Methodology, Resources, Supervision, Validation, Writing - review & editing. **Michael S. North:** Conceptualization, Methodology, Resources, Supervision, Validation, Writing - review & editing.

Funding sources

Haas School of Business, University of California, Berkeley Stern School of Business, New York University The Eckles Fund for Diversity and Social Impact.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.obhdp.2022.104190>.

References

- Abele, A. E., Uchronski, M., Suitner, C., & Wojciszke, B. (2008). Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. *European Journal of Social Psychology, 38*(7), 1202–1217. <https://doi.org/10.1002/ejsp.575>
- Acker, J. (2006). Inequality regimes: Gender, class, and race in organizations. *Gender and Society, 20*, 441–464. <https://doi.org/10.1177/0891243206289499>
- Allport, F. H. (1954). The structuring of events: Outline of a general theory with applications to psychology. *Psychological Review, 61*(5), 281. <https://doi.org/10.1037/h0062678>
- Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics, 116*(1), 313–350. <https://doi.org/10.1162/003355301556329>
- Atewologun, D. (2018). Intersectionality theory and practice. In *Oxford Research Encyclopedia of Business and Management*. <https://doi.org/10.1093/acrefore/9780190224851.013.48>
- Bakan, D. (1966). *The Duality of Human Existence: Isolation and Communication in Western Man*. Boston, MA: Beacon Press.
- Bonsoms, D. (2016). America's Wealthiest Business Schools. *Poets & Quants*. Retrieved from: <http://poetsandquants.com/2016/04/04/businessschoolendowments>.
- Bowleg, L., & Bauer, G. (2016). Invited reflection: Quantifying intersectionality. *Psychology of Women Quarterly, 40*(3), 337–341. <https://doi.org/10.1177/0361684316654282>
- Bowles, H. R., Babcock, L., & Lai, L. (2007). Social incentives for gender differences in the propensity to initiate negotiations: Sometimes it does hurt to ask. *Organizational Behavior and Human*.
- Brescoll, V. L., & LaFrance, M. (2004). The correlates and consequences of newspaper reports of research on sex differences. *Psychological Science, 15*, 515–520. <https://doi.org/10.1111/j.0956-7976.2004.00712.x>
- Brescoll, V. L., Ullman, E. L., & Newman, G. E. (2013). The effects of system-justifying motives on endorsement of essentialist explanations for gender differences. *Journal of Personality and Social Psychology, 105*, 891–908. <https://doi.org/10.1037/a0034701>
- Brewer, M. B., Dull, V., & Lui, L. (1981). Perceptions of the elderly: Stereotypes as prototypes. *Journal of Personality and Social Psychology, 41*, 656–670. <https://doi.org/10.1037/0022-3514.41.4.656>
- Bullock, J. G., Green, D. P., & Ha, S. E. (2010). Yes, but what's the mechanism? (don't expect an easy answer). *Journal of Personality and Social Psychology, 98*(4), 550. <https://doi.org/10.1037/a0018933>
- Carothers, B. J., & Reis, H. T. (2013). Men and women are from Earth: Examining the latent structure of gender. *Journal of Personality and Social Psychology, 104*(2), 385–407. <https://doi.org/10.1037/a0030437>
- Chatman, J. A., Boinsnier, A. D., Spataro, S. E., Anderson, C., & Berdahl, J. L. (2008). Being distinctive versus being conspicuous: The effects of numeric status and sex-stereotyped tasks on individual performance in groups. *Organizational Behavior and Human Decision Processes, 107*(2), 141–160. <https://doi.org/10.1016/j.obhdp.2008.02.006>
- Chopik, W. J., & Kitayama, S. (2018). Personality change across the life span: Insights from a cross-cultural, longitudinal study. *Journal of Personality, 86*(3), 508–521. <https://doi.org/10.1111/jopy.12332>
- Correll, S. J. (2017). Reducing gender biases in modern workplaces: A small wins approach to organizational change. *Gender and Society, 31*, 725–750. <https://doi.org/10.1177/0891243217738518>
- Correll, S. J., Weisshaar, K. R., Wynn, A. T., & Wehner, J. D. (2020). Inside the black box of organizational life: The gendered language of performance assessment. *American Sociological Review, 85*(6), 1022–1050. <https://doi.org/10.1177/0003122420962080>
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2004). When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues, 60*, 701–718. <https://doi.org/10.1111/j.0022-4537.2004.00381.x>
- Cuddy, A. J., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology, 40*, 61–149. [https://doi.org/10.1016/s0065-2601\(07\)00002-0](https://doi.org/10.1016/s0065-2601(07)00002-0)
- Delucchi, M. (2000). Don't worry, be happy: Instructor likability, student perceptions of learning, and teacher ratings in upper-level sociology courses. *Teaching Sociology, 28*, 220–231. <https://doi.org/10.2307/1318991>
- Durante, F., Fiske, S. T., Kervyn, N., Cuddy, A. J., Akande, A., Adetoun, B. E., ... Storari, C. C. (2013). Nations' income inequality predicts ambivalence in stereotype content: How societies mind the gap. *British Journal of Social Psychology, 52*(4), 726–746. <https://doi.org/10.1111/bjso.12005>
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*, 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>
- Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles, 47*(3), 99–114. <https://doi.org/10.1023/A:1021020920715>
- Else-Quest, N. M., & Hyde, J. S. (2016). Intersectionality in quantitative psychological research: I. Theoretical and epistemological issues. *Psychology of Women Quarterly, 40*(2), 155–170. <https://doi.org/10.1177/0361684316629797>
- Fiedler, K., Schott, M., & Meiser, T. (2011). What mediation analysis can (not) do. *Journal of Experimental Social Psychology, 47*(6), 1231–1236. <https://doi.org/10.1016/j.jesp.2011.05.007>
- Financial Times. (2019). Global MBA Ranking. Accessed at <http://rankings.ft.com/businessschoolrankings/global-mba-ranking-2019>.
- Fiske, S. T. (2017). Prejudices in cultural contexts: Shared stereotypes (gender, age) versus variable stereotypes (race, ethnicity, religion). *Perspectives on Psychological Science, 12*(5), 791–799. <https://doi.org/10.1177/1745691617708204>
- Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science, 27*, 67–73. <https://doi.org/10.1177/0963721417738825>
- Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*, 77–83. <https://doi.org/10.1016/j.tics.2006.11.005>
- Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*, 878–902. <https://doi.org/10.1037/pspa0000163>
- Francioli, S. P., & North, M. S. (2021). The older worker: Gender and age discrimination in the workplace. In K. W. Schaie, & S. L. Willis (Eds.), *Handbook of the Psychology of Aging* (9th Ed.). London: Elsevier. <https://doi.org/10.1016/b978-0-12-816094-7.00013-1>.
- Goller, M., & Billett, S. (2014). Agentic behaviour at work: Crafting learning experiences. In *Discourses on Professional Learning* (pp. 25–44). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-7012-6_3.
- Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality, 37*(6), 504–528. [https://doi.org/10.1016/s0092-6566\(03\)00046-1](https://doi.org/10.1016/s0092-6566(03)00046-1)

- Grandey, A. A., Gabriel, A. S., & King, E. B. (2020). Tackling taboo topics: A review of the three Ms in working women's lives. *Journal of Management*, 46(1), 7–35. <https://doi.org/10.1177/0149206319857144>
- Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist*, 52(11), 1182. <https://doi.org/10.1037/0003-066x.52.11.1182>
- Hancock, A. M. (2007). Intersectionality as a normative and empirical paradigm. *Politics & Gender*, 3(2), 248–254. <https://doi.org/10.1017/s1743923x07000062>
- Handy, J., & Davy, D. D. (2007). Gendered ageism: Older women's experiences of employment agency practices. *Asia Pacific Journal of Human Resources*, 1, 85–99. <https://doi.org/10.1177/1038411107073606>
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.
- Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communality deficit. *Journal of Applied Psychology*, 92(1), 81. <https://doi.org/10.1037/0021-9010.92.1.81>
- Holoien, D. S., & Fiske, S. T. (2013). Downplaying positive impressions: Compensation between warmth and competence in impression management. *Journal of Experimental Social Psychology*, 49, 33–41. <https://doi.org/10.1016/j.jesp.2012.09.001>
- Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal*, 28, 331–343. <https://doi.org/10.1002/smj.582>
- Huang, J., Gates, A., Sinatra, R., & Barbasi, A. (2020). Historical comparison of gender inequality in scientific careers across countries and disciplines. *Proceedings of the National Academy of Sciences*, 117: 4609–4616. <https://doi.org/10.1073/pnas.1914221117>
- Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, 60, 581–592. <https://doi.org/10.1037/0003-066x.60.6.581>
- Imhoff, R., & Koch, A. (2017). How orthogonal are the Big Two of social perception? On the curvilinear relation between agency and communion. *Perspectives on Psychological Science*, 12(1), 122–137. <https://doi.org/10.1177/1745691616657334>
- Joshi, A., Son, J., & Roh, H. (2015). When can women close the gap? A meta-analytic test of sex differences in performance and rewards. *Academy of Management Journal*, 58, 1516–1545. <https://doi.org/10.5465/amj.2013.0721>
- Jussim, L. (1986). Self-fulfilling prophecies: A theoretical and integrative review. *Psychological Review*, 93(4), 429. <https://doi.org/10.1037/0033-295x.93.4.429>
- Kanter, R. M. (1977). Some effects of proportions on group life. In *The Gender Gap in Psychotherapy* (pp. 53–78). Boston, MA: Springer. https://doi.org/10.1007/978-1-4684-4754-5_5
- 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>
- Krueger, J., Heckhausen, J., & Hundertmark, J. (1995). Perceiving middle-aged adults: Effects of stereotype-congruent and incongruent information. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 50, 82–93. <https://doi.org/10.1093/geronb/50B.2.P82>
- Kuckartz, U., & Rädiker, S. (2019). *Analyzing Qualitative Data with MAXQDA: Text, Audio, and Video*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-15671-8>
- Lincoln, A. E., & Allen, M. P. (2004). Double jeopardy in Hollywood: Age and gender in the careers of film actors, 1926–1999. In *Sociological Forum* (Vol. 19, No. 4, pp. 611–631). Kluwer Academic Publishers-Plenum Publishers. <https://doi.org/10.1007/s11206-004-0698-1>
- Lindholm, T., & Yzerbyt, V. Y. (2018). When being nice or being smart could bring you down: Compensatory dynamics in strategic self-presentation. *International Review of Social Psychology*, 31, 16. <https://doi.org/10.5334/irsp.136>
- Livingston, R. W., Rosette, A. S., & Washington, E. F. (2012). Can an agentic Black woman get ahead? The impact of race and interpersonal dominance on perceptions of female leaders. *Psychological Science*, 23, 354–358. <https://doi.org/10.1177/0956797611428079>
- Madera, J. M., Hebl, M. R., & Martin, R. C. (2009). Gender and letters of recommendation for academia: Agentic and communal differences. *Journal of Applied Psychology*, 94, 1591–1599. <https://doi.org/10.1037/a0016539>
- Martin, A. E., North, M. S., & Phillips, K. W. (2019). Intersectional escape: Older women elude agentic prescriptions more than older men. *Personality and Social Psychology Bulletin*, 45, 342–359. <https://doi.org/10.1177/0146167218784895>
- May, D. C. (2005). *The nature of School of Education faculty work and materials for promotion and tenure at a major research university*. University of Pittsburgh. Doctoral dissertation.
- Moss-Racusin, C. A. (2021). Psychology of gender: Addressing misconceptions and setting goals for the field. *American Psychologist*, 76(9), 1429–1441. <https://doi.org/10.1037/amp0000930>
- Mullkerrins, J. 2015. Jane Fonda: Men want young women. For us, it's about trying to stay young. *The Telegraph*, May 22, 2015.
- Neumark, D., Burn, I., & Button, P. (2015). Is it harder for older workers to find new jobs? New and improved evidence from a field experiment. *National Bureau of Economic Research*. <https://doi.org/10.1086/701029>
- North, M. S. (2019). A GATE to understanding “older” workers: Generation, age, tenure, experience. *Academy of Management Annals*, 13(2), 414–443. <https://doi.org/10.5465/annals.2017.0125>
- North, M. S., & Fiske, S. T. (2012). An inconvenienced youth? Ageism and its potential intergenerational roots. *Psychological Bulletin*, 138, 982–997. <https://doi.org/10.1037/a0027843>
- North, M. S., & Fiske, S. T. (2013). Act your (old) age: Prescriptive, ageist biases over succession, consumption, and identity. *Personality and Social Psychology Bulletin*, 39(6), 720–734. <https://doi.org/10.1177/0146167213480043>
- North, M. S., & Fiske, S. T. (2015). Intergenerational resource tensions in the workplace and beyond: Individual, interpersonal, institutional, international. *Research in Organizational Behavior*, 35, 159–179. <https://doi.org/10.1016/j.riob.2015.10.003>
- Okimoto, T. G., & Brescoll, V. L. (2010). The price of power: Power seeking and backlash against female politicians. *Personality and Social Psychology Bulletin*, 36, 923–936. <https://doi.org/10.1177/0146167210371949>
- Parks-Stamm, E. J., Heilman, M. E., & Hearn, K. A. (2008). Motivated to penalize: Women's strategic rejection of successful women. *Personality and Social Psychology Bulletin*, 34, 237–247. <https://doi.org/10.1177/0146167207310027>
- Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2015). *Linguistic Inquiry and Word Count: LIWC2015*. *Pennebaker Conglomerates*. www.LIWC.net
- Phelan, J. E., Moss-Racusin, C. A., & Rudman, L. A. (2008). Competent yet out in the cold: Shifting criteria for hiring reflect backlash toward agentic women. *Psychology of Women Quarterly*, 32, 406–413. <https://doi.org/10.1111/j.1471-6402.2008.00454.x>
- Phelan, J. E., & Rudman, L. A. (2010). Prejudice toward female leaders: Backlash effects and women's impression management dilemma. *Social and Personality Psychology Compass*, 4(10), 807–820. <https://doi.org/10.1111/j.1751-9004.2010.00306.x>
- Pliner, P., Chaiken, S., & Flett, G. L. (1990). Gender differences in concern with body weight and physical appearance over the life span. *Personality and Social Psychology Bulletin*, 16, 263–273. <https://doi.org/10.1177/0146167290162007>
- Ponce de Leon, R., & Rosette, A. S. (2022). “Invisible” Discrimination: Divergent Outcomes for the Non-Prototypicality of Black Women. *Academy of Management Journal*, (ja). <https://doi.org/10.5465/amj.2020.1623>
- Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, 26, 269–281. <https://doi.org/10.1111/1471-6402.t01-1-00066>
- Ridgeway, C. L. (2001). Gender, Status, and Leadership. *Journal of Social Issues*, 57, 637–655. <https://doi.org/10.1111/0022-4537.00233>
- Ridgeway, C. L., & Correll, S. J. (2004). Unpacking the gender system: A theoretical perspective on gender beliefs and social relations. *Gender and Society*, 18, 510–531. <https://doi.org/10.1177/0891243204265269>
- Rosette, A. S., Koval, C. Z., Ma, A., & Livingston, R. (2016). Race matters for women leaders: Intersectional effects on agentic deficiencies and penalties. *The Leadership Quarterly*, 27(3), 429–445. <https://doi.org/10.1016/j.leaqua.2016.01.008>
- Rosette, A., Ponce de Leon, R., Koval, C. Z., & Harrison, D. (2018). Intersectionality: Connecting experiences of gender with race at work. *Research in Organizational Behavior*, 38, 1–22. <https://doi.org/10.1016/j.riob.2018.12.002>
- Rothausen-Vange, T. J., Marler, J. H., & Wright, P. M. (2005). Research productivity, gender, family, and tenure in organization science careers. *Sex Roles*, 53(9), 727–738. <https://doi.org/10.1007/s11199-005-7737-0>
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, 74, 629–645. <https://doi.org/10.1037/0022-3514.74.3.629>
- Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, 77, 1004–1010. <https://doi.org/10.1037/0022-3514.77.5.1004>
- Rudman, L. A., Moss-Racusin, C. A., Phelan, J. E., & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against female leaders. *Journal of Experimental Social Psychology*, 48, 165–179. <https://doi.org/10.1016/j.jesp.2011.10.008>
- Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, 28, 61–79. <https://doi.org/10.1016/j.riob.2008.04.003>
- Sanner, B., & Bunderson, J. S. (2015). When feeling safe isn't enough: Contextualizing models of safety and learning in teams. *Organizational Psychology Review*, 5, 224–243. <https://doi.org/10.1177/2041386614565145>
- Schein, V. E., & Davidson, M. J. (1993). “Think manager—think male”—Managerial sex typing among U.K. business students. *Management Development Review*, 6, 24–28. <https://doi.org/10.1108/eum00000000000738>
- Simonton, D.K. (1988). Creativity, leadership, and chance. In R.J. Sternberg (Ed.), *The nature of creativity: Contemporary Psychological Perspectives*: 386–426.
- Simpson, P. M., & Sigauw, J. A. (2000). Student evaluations of teaching: An exploratory study of the faculty response. *Journal of Marketing Education*, 22, 199–213. <https://doi.org/10.1177/0273475300223004>
- Singh, D., & Young, R. K. (1995). Body weight, waist-to-hip ratio, breasts, and hips: Role in judgments of female attractiveness and desirability for relationships. *Ethology and Sociobiology*, 16, 483–507. [https://doi.org/10.1016/0162-3095\(95\)00074-7](https://doi.org/10.1016/0162-3095(95)00074-7)
- Smith-Barrow, D. (2015). MBA organizations strive for gender diversity in the classroom. Accessed at <https://www.usnews.com/education/best-graduate-schools/top-business-schools/articles/2015/11/17/how-mba-organizations-push-for-gender-diversity>.

- Suitner, C., & Maass, A. (2008). The role of valence in the perception of agency and communion. *European Journal of Social Psychology, 38*(7), 1073–1082. <https://doi.org/10.1002/ejsp.525>
- Stark, P. B., & Freishtat, R. (2014). An evaluation of course evaluations. *ScienceOpen*. <https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1>
- Tang, T. L. P., & Chamberlain, M. (1997). Attitudes toward research and teaching: Differences between administrators and faculty members. *The Journal of Higher Education, 68*(2), 212–227. <https://doi.org/10.1080/00221546.1997.11780882>
- U.S. Bureau of Labor Statistics. (2022). Labor force statistics from the current population survey. <https://www.bls.gov/cps/demographics.htm>.
- Warner, L. R., Settles, I. H., & Shields, S. A. (2016). Invited reflection: Intersectionality as an epistemological challenge to psychology. *Psychology of Women Quarterly, 40*, 171–176. <https://doi.org/10.1177/0361684316641384>
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37*(3), 395. <https://doi.org/10.1037/0022-3514.37.3.395>
- Williams, M. J., & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin, 142*(2), 165. <https://doi.org/10.1037/bul0000039>
- Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology, 38*(7), 1139–1147. <https://doi.org/10.1002/ejsp.549>