ROYAL SOCIETY OPEN SCIENCE

A registered report on how implicit pro-rich bias is shaped by the perceiver's gender and socioeconomic status

Journal:	*Royal Society Open Science*
Manuscript ID	RSOS-191232.R1
Article Type:	Registered Report - Stage 1
Date Submitted by the Author:	n/a
Complete List of Authors:	Mattan, Bradley; University of Pennsylvania, Annenberg School for Communication Cloutier, Jasmin; University of Delaware, Department of Psychological and Brain Sciences
Subject:	psychology < BIOLOGY
Keywords:	Implicit Bias, Social Status, Gender, Income, Education, Individual Differences
Subject Category:	Psychology and cognitive neuroscience

A registered report on how implicit pro-rich bias is shaped by the perceiver's gender and socioeconomic status

Bradley D. Mattan[1] & Jasmin Cloutier[2]

Affiliations:

[1] Annenberg School for Communication, University of Pennsylvania, Philadelphia, PA, USA

[2] Department of Psychological and Brain Sciences, University of Delaware, Newark, DE, USA

Corresponding author:

Bradley D. Mattan

Annenberg School for Communication

University of Pennsylvania

3620 Walnut St.

Philadelphia, PA 19104

Telephone: +1 (312)-450-9194

Email: brad.mattan@gmail.com

Abstract

Although high status is often considered a desirable quality, this may not always be the case. Different factors may moderate the value of high status along a dimension such as wealth (e.g., gender, perceiver income/education). For example, studies suggest men may value wealth and control over resources more than women. This may be especially true for high-income men who already have control over substantial resources. Other work suggests that low-income men and women may have different experiences in educational contexts compared to their richer peers who predominate at higher levels of education. These experiences may potentially lead to different attitudes about the wealthy among low-income men and women. In this stage-1 registered report, we propose two key predictions based on our review of the literature and analyses of pilot data from the Attitudes, Identities, and Individual Differences (AIID) Study (n>1,000): (H1) Increasing income will be associated with increased pro-wealthy bias for men more than for women, and (H2) income will also moderate the effect of education on implicit pro-wealthy bias, depending on gender.

Keywords

Implicit Bias, Social Status, Gender, Income, Education, Individual Differences

Background

Sometimes subtle (Bjornsdottir & Rule, 2017; Hall, Coats, & LeBeau, 2005; Oh, Shafir, & Todorov, 2019), signs of social status are readily observed in others, ultimately influencing how we evaluate people (Kraus, Park, & Tan, 2017). With some exceptions (Cloutier, Ambady, Meagher, & Gabrieli, 2012; Cloutier & Gyurovski, 2014; Cuddy, Fiske, & Glick, 2007; Garcia, Weaver, & Chen, 2018; Horwitz & Dovidio, 2017), we generally perceive high socioeconomic status (SES) in others as a positive characteristic (Fiske, 2010; Mattan, Kubota, & Cloutier, 2017; Mattan, Wei, Cloutier, & Kubota, 2018; Varnum, 2013). Even at the level of implicit associations, evidence suggests that higher SES (among other status dimensions: see Boukarras, Era, Aglioti, & Candidi, 2019) is associated with increasing positivity (Mattan, Kubota, Li, Venezia, & Cloutier, 2019). Moreover, greater wealth—a component of SES—frequently elicits positive implicit evaluations (Cunningham, Nezlek, & Banaji, 2004; Horwitz & Dovidio, 2017; Rudman, Feinberg, & Fairchild, 2002). These implicit status-based biases can have a real-world impact, shaping the evaluations and judgments we make in circumstances characterized by ambiguity and/or time pressure (Horwitz & Dovidio, 2017; Moore-Berg, Karpinski, & Plant, 2017). Critically, social status cues are seldom perceived in isolation from other visually salient attributes such as gender. Numerous studies have demonstrated that assumptions about status are ingrained into gendered roles (Eagly, 1987; Eagly & Wood, 1982, 1991; Fiske, 1993; Fiske, Dupree, Nicolas, & Swencionis, 2016) and interactions (Berger, Cohen, & Zelditch, 1972; Lockheed & Hall, 1976; Ridgeway, 1991; Ridgeway & Balkwell, 1997), such that women are assumed to be subordinate to men. Dovetailing with this work on perceived gender roles, the literature on masculinity suggests that the control over social and financial capital that high status entails is important to men's gender identity (Connell, 1987; Connell & Messerschmidt, 2005;

Mahalik et al., 2003). Weaving together these diverse strands of research, this registered report introduces and tests the hypothesis that men show greater implicit pro-wealthy bias than do women in a large online sample. However, we also predict that gender-based positive associations with wealth will further depend on the perceiver's own position in the social hierarchy (Mattan et al., 2017). Specifically, we will examine how this masculine preference for the wealthy is associated with the perceiver's own SES (e.g., income, education), predicting that both income and education will shape men's (vs. women's) preferences for the wealthy.

Gender Roles and Status Incongruity

The literature on gender roles provides a helpful starting point for the present investigation of how one's gender and SES shape implicit preferences for the wealthy. The gender roles account starts from the premise that social expectations about what is appropriate for each sex predicts gender differences in behavior (Eagly & Wood, 1991). Specifically, women are expected to be communal and focused on supporting the family, whereas men are expected to be agentic and focused on the public sphere (Eagly, 2009; but see Twenge, 2009). Due to their stereotypic agency and greater historical participation in the paid workforce, men are often presumed to occupy superior positions relative to women (Eagly & Steffen, 1984; Eagly & Wood, 1982; Rudman & Glick, 2008; Rudman & Kilianski, 2000; Twenge, 2001). This gender differential in the perceived status is reinforced by a collection of ambivalent attitudes and beliefs about women (Glick & Fiske, 1996, 2001). On the one hand, hostile sexism reflects negative attitudes and stereotypes about women with potential negative consequences irrespective of the context (Goh & Hall, 2015; Masser & Abrams, 2004; Masser, Viki, & Power, 2006). On the other hand, benevolent sexism reflects relatively warmer attitudes toward women based in beliefs that women are less capable than and require the protection of men (Glick &

Fiske, 1996, 2001). Despite its association with greater warmth toward women (Glick & Fiske, 1996, 2001; Goh & Hall, 2015), benevolent sexism can also have negative consequences for the advancement of women due to implied stereotypes (e.g., low competence/agency) that give these warm attitudes a decidedly paternalistic flavor (Fiske, 1993; Vescio, Gervais, Snyder, & Hoover, 2005). Indeed, women who violate gender roles by pursuing high-status positions frequently face backlash in the form of negative social evaluations (Eagly & Karau, 2002; Rudman, 1998; Rudman & Fairchild, 2004; Rudman & Glick, 1999; Rudman, Moss-Racusin, Phelan, & Nauts, 2012). Like women, men have also been shown to face backlash for violating gender roles (viz., by being modest: Moss-Racusin, Phelan, & Rudman, 2010; Rudman, 1998). Taken together, these findings provide support for the status incongruity hypothesis, which proposes that violating status-based gender stereotypes has negative consequences for perceived employability and agreeableness (Eagly & Karau, 2002; Moss-Racusin et al., 2010; Rudman et al., 2012).

Backlash resulting from gender-status incongruity may also have consequences for how men and women of varying socioeconomic positions value cues related to SES, such as wealth. Existing evidence suggests that people generally show implicit bias favoring high-status targets; however this bias is most pronounced in those belonging to a privileged group (Mattan et al., 2019; Rudman et al., 2002). Given that gender roles can confer on men a privileged status over women, it follows that implicit pro-rich bias may be greater for men than it is for women. However, it remains an open question whether wealthy or highly educated women may also show an implicit preference for status that is similar to that of their wealthy/educated male counterparts. In one experiment examining a similar question in the domain of race (Mattan et al., 2019), we found an interaction between perceiver race and SES in predicting positive implicit associations with high-SES people such that status-congruent individuals (low-SES

Black and high-SES White participants) showed reliable pro-high-SES bias. However, individuals who were status-incongruent vis-à-vis their race (high-SES Black and low-SES White participants) showed relatively unreliable implicit bias favoring high-SES targets. This may be explained by status-incongruent individuals having ambivalent feelings about status. On the one hand, achieving higher rank is generally desirable (but see Anderson, Willer, Kilduff, & Brown, 2012). On the other hand, status-incongruent Black and White participants may also experience status-based threats (e.g., denigration, exclusion) from high-status White people who do not see them as equals (Kunstman, Plant, & Deska, 2016; Mendoza-Denton, Downey, Purdie, Davis, & Pietrzak, 2002), making the experience of being high in rank seem less desirable. Previous work suggests that competing (i.e., ambivalent) associations can result in attenuated implicit bias (Petty, Briñol, Tormala, & Jarvis, 2006). Accordingly, one possibility is that status-incongruent women and men may show ambivalent attitudes about the rich and attenuated pro-rich implicit bias relative to wealthy or highly educated men and poor or high-school-educated women.

Social Hierarchy and Masculine Identity

Complementing the gender roles literature, which focuses on how socially prescribed roles reward the behaviors of men and women, the psychology of gender identity (e.g., masculinity, femininity) focuses more closely on identification with one's chosen and/or ascribed gender. Of considerable relevance to the present project, the theory of precarious manhood suggests that masculinity is distinct from femininity in its intimate relationship with status-related concerns (Vandello & Bosson, 2013). Unlike femininity, which is generally considered stable and resulting from physical maturity, masculinity is seen as a form of status awarded through achievement of cultural standards of manhood (Gilmore, 1990; Vandello, Bosson,

Cohen, Burnaford, & Weaver, 2008). Not only are these standards generally demanding, the status of men is precarious in the sense that it can be lost when one fails to measure up to these standards (Moss-Racusin et al., 2010; Vandello et al., 2008). Restoring one's public image as a man when this image is threatened (i.e., gender threat: Caswell, Bosson, Vandello, & Sellers, 2014; Vandello et al., 2008) is thought to require public action to restore perceptions of control or competence (Weaver, Vandello, & Bosson, 2013; Weaver, Vandello, Bosson, & Burnaford, 2010), among other hallmarks of hegemonic masculinity (see Connell, 1987; Connell & Messerschmidt, 2005; Mahalik et al., 2003; Williams, Paluck, & Spencer-Rodgers, 2010).

Consistent with the picture that men's status is fragile, it stands to reason that men may be more vigilant toward and desirous of opportunities to advance their own status in comparison to women. Indeed, evidence suggests that men compared to women more readily display and pursue higher status when status is construed in terms of social or economic influence (Baumeister & Sommer, 1997; Dovidio, Ellyson, Keating, Heltman, & Brown, 1988; Eagly, 2009; Van Vugt, De Cremer, & Janssen, 2007; van Vugt & Tybur, 2015). Even among early adolescents, boys' preference for popularity has a greater impact on their affiliations with peers than it has for girls (Shin, 2017). In adulthood, men also tend to report valuing wealth (Desmarais & Curtis, 1997; Furnham, 1984; Martin & Kirkcaldy, 1998; Prince, 1993; Tang & Gilbert, 1995) and power (i.e., control over resources: Buss, 1981; Hays, 2013) more than women.

In sum, findings from the literature on masculinity suggest that one reason why men may evaluate high SES more positively is because SES is communicated easily and publicly through various cues (e.g., posture, voice, clothing: Kraus, Torrez, Park, & Ghayebi, 2019; Oh et al., 2019; Shariff & Tracy, 2009; Tracy, Shariff, Zhao, & Henrich, 2013) that convey control over

social and financial capital. Publicly conveying such an image would be important for claiming, restoring, and/or maintaining one's identity as a man (Vandello & Bosson, 2013). Because one's status as a man may be easily challenged, public reminders of high SES (e.g., wealth, professional degrees) could therefore be of greater value to men (vs. women) occupying any position on the social hierarchy. However, this is perhaps especially the case for men whose only claim to status is their income, which is more fluid than educational attainment.

Gendered Preference for the Rich Based on Hierarchical Position

To more clearly identify how gender shapes implicit pro-rich bias, this study examined the moderation of this bias by the perceiver's own income and education. Examining the contribution of individual differences like income and education is important because our relative position along these distinct dimensions of status may differentially impact how we evaluate wealth in ourselves and others (Cheng & Tracy, 2013; Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013; Cloutier et al., 2012; Cloutier & Gyurovski, 2014; Mattan et al., 2017, 2019; Rudman et al., 2002).

Income. Despite some evidence that men may desire and value wealth more with increasing income levels (Desmarais & Curtis, 1997) and that wealth is stereotypically tied to masculinity (Williams et al., 2010), few studies have examined how positive evaluations of the rich are moderated by the perceiver's own gender and income. One possibility is that status-congruent individuals (viz., high-income men, low-income women) will show stronger implicit pro-rich bias than status-incongruent individuals (viz., high-income women, low-income men). This would result in a Gender × Income interaction, with increasing pro-rich bias in men and decreasing pro-rich bias in women as a function of increasing income. Greater pro-rich bias among rich men would be consistent with previous work showing greater preferences for high-

status groups (e.g., the rich) among individuals who consider themselves members of those groups (B. Mullen, Brown, & Smith, 1992; Rudman et al., 2002). The attenuated pro-rich bias among status-incongruent individuals may be due to ambivalent feelings about the rich (see Petty et al., 2006)—a hypothesis we intend to explore. Based in part upon this literature and pilot data reported below, our first prediction (H1) is that men will show greater pro-wealthy bias than do women, particularly as a function of their own increasing income (i.e., Gender × Income interaction).

Education. In contrast to income, it is perhaps less clear how education levels may affect pro-wealthy bias, implicit or otherwise. Although education is associated with income, it can also be considered a reliable index of SES on its own (Oakes & Rossi, 2003) that is readily ascertained in social interactions (e.g., through speech: Kraus, Torrez, et al., 2019). Education affords greater occupational opportunities (Andersen & van de Werfhorst, 2010; Posselt & Grodsky, 2017; Sewell, Haller, & Ohlendorf, 1970) and upward economic mobility (Brand & Xie, 2010; Breen & Jonsson, 2005). This is especially the case for affiliates of elite academic institutions (Rivera, 2011), which further enhance one's prestige (Gordon, 1980; Peters & Ceci, 1982). Conceptualized in this way as a relatively public and direct representation of one's socioeconomic rank, one might expect education to function similarly to income, increasing pro-rich bias among those who are most privileged (viz., the highly educated). Beyond the direct status-conveying aspects of education (e.g., Kraus, Torrez, et al., 2019), advanced education as a component of SES also has a cultural component (Kraus, Callaghan, & Ondish, 2019; Kraus, Piff, & Keltner, 2011), which may further enhance positivity toward the rich among the well-educated. For example, students who pursue undergraduate or graduate degrees tend to experience greater exposure to a high-SES cultural environments (viz., academia: Jury, Aelenei,

Chen, Darnon, & Elliot, 2019; Lott, 2012), potentially developing more positive attitudes toward wealth and the wealthy over time as a result of their contact with a generally high-SES population (cf. Henry & Hardin, 2006; Kubota, Peiso, Marcum, & Cloutier, 2017).

Income and education. We have so far considered the contributions of perceiver income and education to implicit wealth-based bias separately. However, one of the objectives of the present study is to determine whether these two factors may differentially impact the way women and men evaluate the rich and poor. Based on an integration of the literature reviewed above, we offer two accounts of how income, education, and gender may together shape implicit associations about wealth. These accounts focus separately on women and men and are therefore not mutually exclusive.

Cultural fit account. The first account is motivated by the cultural component of education. As mentioned above, pursuing higher education involves immersion in a typically middle- to upper-class environment, which may lead to greater affinity toward the rich by proxy. However, the implications of contact with the rich likely differ depending on the income of the student or the student's family. In low-income but highly educated individuals, pro-wealthy associations brought about by familiarity with upper class peers may be offset to some degree by difficulties in navigating cultural differences across class lines (Jack, 2019; Jury et al., 2019; Stephens, Townsend, & Dittmann, 2019; Swencionis, Dupree, & Fiske, 2017). Indeed, research has identified the academic context as a potential threat to the class identity of middle-class or first-generation students at elite universities, resulting in greater stress (Stephens, Townsend, Markus, & Phillips, 2012) and need for self-regulation (Johnson, Richeson, & Finkel, 2011). In sum, the cultural component of education's contribution to status suggests that lower-income individuals may experience a combination of increased positive contact with the rich (e.g., as

friends and colleagues) and negative challenges arising from differences in cultural fit within academe from which their richer peers are sheltered. Such negative experiences may serve to attenuate implicit pro-wealthy bias in low-income but highly educated individuals (Petty et al., 2006). Notably, this prediction runs counter to the prediction based on the prestige component of education as a direct representation of social rank, at least for low-income individuals. The prestige-based component would predict generally enhanced rather than attenuated pro-rich bias as a function of increasing education.

Extending the cultural fit account into the domain of gender, it seems likely that low-income women may have ambivalent associations with academia not just because of their income but also because of their gender. Although women receive undergraduate and postgraduate degrees in similar if not greater numbers than men (A. L. Mullen & Baker, 2015; Posselt & Grodsky, 2017), gender disparities persist in educational environments. For example, in high-status disciplines such as science, technology, engineering, and mathematics (STEM), women are marginalized professionally (Blickenstaff, 2005; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012; Sheltzer & Smith, 2014; Xie, Fang, & Shauman, 2015; Zhang, Schmader, & Forbes, 2009). Even outside of STEM disciplines, women may find themselves underrepresented at the most prestigious academic institutions (A. L. Mullen & Baker, 2015) and face additional hurdles in the pursuit of advanced degrees (Posselt & Grodsky, 2017). In light of this evidence, one possibility is that having both low income and higher education may lead to ambivalence about status and therefore greater attenuation of pro-rich bias in women compared to men.

Precarious manhood account. The literature on precarious manhood suggests a complementary effect of education for low-income men. Having low income would constitute a

potential threat to the masculine image as capable and self-sufficient (Collinson & Hearn, 2005; Strier, Sigad, Eisikovits, & Buchbinder, 2014). For low-income men, education may provide one means of safeguarding against such threats to masculinity in that it can be conveyed publicly (e.g., through speech: Kraus, Torrez, et al., 2019). Moreover, unlike other sources of status (e.g., income), education is more stable. Accordingly, having an advanced degree may make one's income less important when it comes to conferring status. For example, if men are low on income, they may nonetheless perceive and present themselves as high in status by downplaying their wealth in favor of education (see Crocker & Wolfe, 2001). Indeed, this strategy is readily endorsed among some second-generation immigrant groups (Urban, 2012).

If ranking highly on any dimension of SES is sufficient evidence of masculinity, then wealthy and/or highly educated men should consider themselves as part of a common high-status ingroup—men, and may ultimately respond positively toward all other high-status individuals (e.g., the rich). In such a scenario, it is unclear whether having high levels of both income and education might further boost implicit preferences for the wealthy in men relative to women. In any case, the two accounts presented here propose opposing but not mutually exclusive effects of education for low-income men and women. Whereas the cultural account suggests that low income and increasing levels of education may attenuate women's implicit positive associations with the rich, the precarious manhood account suggests that this same combination could enhance men's implicit positive associations with the rich.

Predictions

Based on the literature reviewed above and findings from our pilot data, we offer the following hypotheses. First, we predict that income will increase implicit pro-wealthy bias more for men (H1A) than women (H1B). Our second key prediction (H2) is that both income and

education will interact in shaping men's (vs. women's) implicit pro-wealthy bias (Gender × Income × Education interaction). At low income levels, men relative to women will show greater pro-wealthy bias as a function of increasing education (Gender × Education interaction). The cultural fit account specifically predicts that greater increasing education will lead to an attenuation of pro-wealthy bias among low-income women (H2A). Not exclusive of the cultural fit account, the precarious manhood account predicts that increasing education will have the opposite effect among low-income men (H2B). At high income levels, increasing education may further enhance pro-wealthy bias in both women (H2C) and men (H2D), but it is not clear whether this increase will also depend on gender.

Method

All data for the pilot and confirmatory analyses for this report come from the *Attitudes, Identities, and Individual Differences (AIID) Study* (Hussey et al., 2019). Collected entirely through Project implicit (https://implicit.harvard.edu/implicit/) between 2004 and 2007, AIID consists of a large ($N \simeq 200{,}000$ samples with complete data) and diverse sample of participants who were randomly assigned to complete implicit (Implicit Association Test—IAT: Greenwald, McGhee, & Schwartz, 1998) and self-report measures of attitudes, identity, motivations, and cultural perceptions. Our analyses focus on the rich–poor evaluative IAT. The rich–poor dimension was one of 95 different IAT dimensions used in the AIID Study. Because each participant completed measures relevant to a single IAT dimension, the final sample for confirmatory analysis is substantially smaller than the total AIID dataset ($1{,}100 < n < 1{,}200$). In addition to the IAT, participants also completed a subset of individual difference measures and a standard set of demographics items, including gender, income level, and education level. In this registered report, we present all measures relevant to our confirmatory and exploratory analyses

(see Measures). However, a full overview of the AIID Study and its measures is available at https://osf.io/af5dn/.

Participants voluntarily completed the AIID Study, and no inducements or incentives from the AIID research team were utilized. At the start of the study, participants viewed an introduction screen that included a consent agreement stating that participation was voluntary, non-compensated, and with minimal risk. The consent agreement also stated that responses were confidential and anonymous, protected and analyzed in the aggregate, and with information of who to contact in the event of any queries. Also included was the Project Implicit privacy policy. Consent was presumed through participation in the study. Data were collected in accordance with University of Virginia IRB protocol #2003017300.

The present analyses are in response to a recent call from the AIID team intended to promote the use of registered reports using data from the AIID Study. To preserve experimenter blinding and promote the use of registered reports (Nosek, Ebersole, DeHaven, & Mellor, 2018), Hussey and colleagues (2019) split the AIID dataset into stratified subsamples, with approximately 15% of the data available for pilot analyses and approximately 85% of the data withheld for confirmatory analyses. For this stage-1 registered report, we only have access to the smaller sample intended for pilot analyses (see https://osf.io/entbj/ for confirmation letter from C. R. Ebersole). We provide details on the sample used for pilot analyses and our estimate of the sample size to be used for confirmatory analyses (anticipated to be released upon provisional stage-1 acceptance). Data from the pilot analyses will not be included in the confirmatory analyses.

Inclusion and Exclusion Criteria

Using the pilot AIID dataset, we initially included all participants who completed the rich–poor evaluative IAT (n_{pilot} = 274; $n_{confirmatory}$ = TBD). We will use the same inclusion criteria in our confirmatory analyses. Similarly, all exclusion criteria reported below apply to both our presently reported pilot analyses and future confirmatory analyses.

Trial-level exclusions in scoring of the IAT. *D* scores for the IAT were computed in line with existing recommendations, which require some trial-level exclusions (see Table 4 from Greenwald, Nosek, & Banaji, 2003). Namely, trials exceeding 10,000 ms were eliminated prior to analysis. Additionally, latencies for incorrect responses were replaced with the block mean RT plus 600 ms. These exclusions were incorporated into the *D* scores already provided in the AIID dataset. Relative to the original IAT scoring algorithm (Greenwald et al., 1998), the updated scoring algorithm (even with the above exclusions) includes more trials by including practice block data and by transforming rather than excluding error latencies. This has been shown to reduce noise from excessively slow responses and from previous experience with the IAT, ultimately increasing the power to detect relationships between the IAT and individual difference measures such as those used in the present study (Greenwald et al., 2003). Thorough details on the IAT scoring algorithm and relevant trial-level exclusions are provided in the section below that describes the IAT measure and its scoring.

Participant-level exclusions. Individuals were required to make site-wide user IDs and passwords prior to participating in any study on Project Implicit, with the result that users could be followed across studies, including repeat participations in this study. Although no repeated IDs were uncovered in our pilot dataset, should any arise in the confirmatory dataset, we will keep only the ID associated with the earliest time and date. Participant-level exclusions will be the same for the confirmatory dataset as they were for the pilot dataset.

After IAT *D* scores were computed (see section below describing IAT scoring), we implemented participant-level exclusions, following the guidelines offered by Project Implicit (see footnote 4 of Nosek et al., 2007). The implementation of the criterion listed below aim to reduce the incidence of careless responding in the final dataset. Specifically, we excluded any participants who met any of the following criteria (n_{pilot} = 46; $n_{confirmatory}$ = TBD):

(1) Greater than or equal to 35% of responses under 300 ms in any one practice block.

(2) Greater than or equal to 25% of responses under 300 ms in any one critical block.

(3) Greater than or equal to 10% of responses under 300 ms in critical blocks.

(4) Greater than or equal to a 50% error rate in any one practice block.

(5) Greater than or equal to a 40% error rate in practice blocks.

(6) Greater than or equal to a 40% error rate in any one critical block.

(7) Greater than or equal to a 30% error rate in critical blocks.

(8) In addition to implementing criteria 1–7 that were used by Nosek and colleagues (2007), we also adopted a stricter exclusion criterion, removing any participant with greater than or equal to 10% responses over 10,000 ms in IAT critical blocks. This additional criterion was implemented to exclude participants who were potentially insufficiently attentive during, or confused by, the IAT.

Finally, we excluded all rich–poor evaluative IAT participants who failed to complete all demographic items in our reported analyses in the following order: income level (n_{pilot} = 49; $n_{confirmatory}$ = TBD), education level (n_{pilot} = 4; $n_{confirmatory}$ = TBD), and gender (n_{pilot} = 0; $n_{confirmatory}$ = TBD).

Participants

Pilot dataset. After all exclusions (see above), our final sample consisted of 175 participants (n_{women} = 115, M_{age} = 32.4 years, $Range_{age}$ = 14–66 years). The sample was predominantly White (71.4%), but more than a quarter of participants were minorities: Mixed Race or Other/Unknown (9.7%), Hispanic (8.0%), Asian/Pacific Islander (6.3%), and Black (2.3%). Four participants (2.3%) did not provide their racial/ethnic demographic information. For a breakdown of our sample by gender, income level, and education level, see Table 1.

Table 1. Distribution of participants from the pilot dataset (n = 175) by gender, income level, and education level.

Gender	Income Level (USD)	No High School Diploma	High School Graduate	Associate's Degree or Some College	Bachelor's Degree	Graduate Degree or Education
Women	< $25,000	0	0	12	6	1
	$25,000 - $49,999	0	5	8	17	3
	$50,000 - $74,999	0	1	7	10	8
	$75,000 - $149,999	0	2	7	8	7
	> $150,000	1	1	2	4	5
Men	< $25,000	0	0	8	7	1
	$25,000 - $49,999	1	0	2	6	0
	$50,000 - $74,999	0	2	3	4	4
	$75,000 - $149,999	0	0	4	9	5
	> $150,000	0	0	1	0	3

Note. Numbers within each cell indicate the sum total of pilot participants in that condition.

Confirmatory dataset. Although it is impossible to know in advance the exact number of participants that will meet our exclusion criteria, we estimate based on the pilot dataset that our final sample prior to exclusions will be approximately 1,827 participants. This estimate assumes that our pilot dataset prior to exclusions consists of a stratified sample of 15% the AIID dataset. Because these samples are stratified by Hussey and colleagues, we anticipate a similar demographic makeup in our confirmatory dataset. Based on the 36% exclusion rate in our pilot sample (see exclusions above), we anticipate a final sample of approximately 1,167 participants.

Measures

Participants in the AIID Study completed a number of measures in addition to the rich–poor IAT. We present first all measures involved in confirmatory analyses, followed by additional measures that may be used for supplemental exploratory analyses.

Measures for confirmatory analysis. Key independent variables used for confirmatory analyses include participant gender, income level, and education level. The IAT *D* score served as our key dependent measure of implicit pro-rich bias. These measures are described in greater detail below.

Gender. Participants were asked to report their sex. Their only response options were male or female. Any participant failing to respond to this item was excluded from all analyses.

Income level. Annual income level was assessed on a five-point scale: (1) < $25,000, (2) $25,000–$49,000, (3) $50,000–$74,999, (4) $75,000–$149,999, (5) > $150,000. All dollar amounts were in U.S. dollars. Participants were also allowed to select "I don't know" for this item. Participants choosing this option or failing to provide a response on this measure were excluded from all analyses. This variable was *z*-transformed prior to all analyses.

Education level. Education level was re-coded by the AIID study coordinators in both the pilot and confirmatory datasets to a five-point scale based on two older items in the AIID study that provided a wider range of values but ultimately resulted in a small number of cases at the scale extremes (see codebook for details at https://osf.io/3sg5e/). The final five-point single-item scale consisted of the following values: (1) Not a high school graduate, (2) High school graduate, (3) Some college or associate's degree, (4) Bachelor's degree, and (5) Graduate degree or graduate education. Participants who failed to provide a response on this measure were excluded from all analyses. This variable was *z*-transformed prior to all analyses.

IAT. The IAT (Greenwald et al., 1998) was used to implicitly measure evaluative associations for the rich and the poor. As for other contemporaneous Project Implicit data (Nosek, Banaji, & Greenwald, 2002; Nosek, Greenwald, & Banaji, 2005), the IAT was presented via the Internet using Java and CGI technology. This software used the respondent's computer resources to present stimuli and record response latencies, thereby reducing noise that would be caused by variable connection speeds. As noted by Nosek and colleagues (2002, 2005), latency recording is limited by the local system's clock rate, with error windows of 16–60 ms. However, the resulting noise is not systematic, and the IAT tends to elicit large effects that are stable due to averaging data across many trials (Nosek et al., 2005).

Introductory blocks and stimuli. Participants began the IAT learning the dimension of interest: rich people versus poor people. In the first block, participants categorized words indicative of the rich and poor using the "a" (e.g., for rich) and ";" (e.g., for poor) keys. Words representing the rich people anchor included: Wealthy, Affluent, Prosperous, and Well Off. Words representing the poor people anchor included: Poor, Impoverished, Broke, and Bankrupt. In the second block, participants then learned the attribute dimension using one of three randomly assigned sets of anchor terms: good versus bad, positive versus negative, or pleasant versus unpleasant. (All analyses collapsed across the three pairs of anchor terms.) As in the first block, participants categorized word stimuli as good/positive/pleasant or bad/negative/unpleasant using the "a" and ";" keys, respectively. For most participants (n_{pilot} = 157; $n_{confirmatory}$ = TBD), words representing the good/positive/pleasant anchor included: Appealing, Delight, Excitement, Glee, Laughing, and Splendid. For this majority of participants, words representing the bad/negative/unpleasant anchor included: Animosity, Dirty, Gross, Evil, Neglected, and Rotten. Due to an apparent bug in stimulus assignment, a small subset of participants (n_{pilot} = 18;

$n_{confirmatory}$ = TBD) completed the rich–poor IAT with a different set of attribute words. Good/positive/pleasant words included: Love, Cheer, Friend, Pleasure, Paradise, and Splendid. Bad/negative/unpleasant words included: Abuse, Grief, Poison, Sadness, Pain, and Bomb. All analyses collapsed across the two attribute word sets.

Dual-categorization blocks and transition block. Next, participants completed the third block. Here, the two preceding tasks were combined such that each key represented two possible categorizations. For example, the "a" key was assigned to both rich people and good/positive/pleasant, and the ";" key was assigned to both poor people and bad/negative/unpleasant. Respondents then categorized both kinds of word stimuli (i.e., those denoting wealth and valence), which alternated throughout this 20-trial practice block. After a brief rest, participants then completed the fourth block (i.e., the first critical block). Here, participants simply repeated the same task as in the third block, but over 40 trials. In the fifth block (i.e., transition block), participants only categorized stimuli along the wealth dimension. But this time, the keys were reversed such that "a" was assigned to poor people and ";" was assigned to rich people. Having completed the fifth block, participants then completed a 20-trial practice block where the "a" key was assigned to both poor people and good/positive/pleasant, and the ";" key was assigned to both rich people and bad/negative/unpleasant. Finally, participants completed a final critical block of 40 trials that was otherwise identical to the preceding 20-trial practice block. As in previous work (Nosek et al., 2005, 2007), the ordering of the third/fourth and sixth/seventh blocks was counterbalanced across participants to minimize the impact of block order effects.

Scoring. The IAT effect is calculated using latency data from the two critical blocks and their preceding practice trials. Categorizing stimuli faster when poor people share the same key

as bad attributes (and when rich people are paired with good attributes) than vice-versa indicates a stronger association strength between poor people and bad (and rich people and good) relative to the opposite key mapping. In other words, this would reflect an implicit pro-rich bias. The IAT *D* scores provided in the AIID dataset were computed according to recommended guidelines, which include some trial-level exclusions (see Table 4 from Greenwald et al., 2003). Specifically, the scoring algorithm for *D* scores in the AIID pilot and confirmatory datasets is as follows:

(1) Use all trials from blocks 3, 4, 6, and 7.

(2) Eliminate any trials with latencies exceeding 10,000 ms.

(3) Compute the mean of correct latencies for each block.

(4) Compute one pooled standard deviation for all trials in blocks 3 and 6 and another for all trials in blocks 4 and 7.

(5) Replace each error latency with its respective block's mean RT, plus 600 ms.

(6) Average the resulting values ofr each of the four blocks.

(7) Compute two difference scores: Block 6 – block 3, block 7 – block 4.

(8) Divide each difference by its associated pooled-trials standard deviation (step 4).

(9) Average the two quotients from step 8.

In summary, the *D* score reflects an average of two sub-scores: (1) average response times between the two practice blocks (i.e., blocks 3 & 6) divided by the standard deviation of all response times for both blocks, and (2) average response times between the two critical blocks (i.e., blocks 4 & 7) divided by the standard deviation of all response times for both blocks.

Due to its similarity to Cohen's *d*, the IAT *D* score is thought to reflect the magnitude of implicit associations (Nosek et al., 2005). The IAT *D* score served as our key dependent measure

of implicit pro-rich bias, with larger positive *D* scores reflecting a greater positive association for the rich relative to the poor. *D* scores were standardized into *Z* scores prior to analysis, as were all continuous independent variables. This was done to facilitate comparisons of effect sizes across analyses and studies.

Measures for exploratory analyses. In addition to the aforementioned measures for confirmatory analysis, the AIID Study includes a rich set of additional variables that may be used for exploratory post-hoc analyses. These additional measures include standard demographic variables and explicit measures tapping into thoughts and feelings about the rich–poor dimension that was assessed by the rich–poor evaluative IAT (i.e., the dependent variable used for confirmatory analysis).

Demographic items. Demographics were assessed prior the completion of other measures in this study during the participant's initial site-wide sign-up. In addition to gender, income level, and education level (i.e., the independent variables for confirmatory analysis), demographic items included participants' age, citizenship, country of residence, social class, English language fluency, race/ethnicity, occupation, political identity (liberalism vs. conservatism), religion, religiosity, and ZIP code.

Explicit measures. In addition to the IAT, participants also completed 27–29 self-report items regarding their own and others' attitudes toward rich people and poor people. These items were pulled from a pool of 76 items, randomized with some constraints (see codebook for details at https://osf.io/3sg5e/). Self-report items were grouped by the AIID coordinators into 18 different measures. Although these measures do not form the main focus of the present report, we provide descriptive statistics and some limited exploratory analyses of these measures to better contextualize our central analyses of implicit bias on the IAT. The AIID study team

counterbalanced the experiment ordering such that individual difference measures preceded the IAT for some participants. Measures 1–14 were included in descriptive and correlational analyses. Measures and 1–8 were additionally included in extended parallel regressions.

(1) Personal evaluations. Participants separately rated the rich and the poor on one of three closely related dimensions: positivity (e.g., "How positive or negative do you feel towards the rich?"), warmth (e.g., "How warm or cold do you feel towards the rich?"), or likeability (e.g., How much do you like or dislike the rich?"), responding on a scale from 1 ("Strongly negative", "Cold", or "Strongly dislike") to 10 ("Strongly positive", "Warm", or "Strongly like"). The explicit evaluations measure consists of the difference in evaluations for the poor minus the rich. This measure was used in the exploratory parallel regressions as an index of the self's explicit evaluations of the poor versus rich.

(2) Others' evaluations. Participants completed three to four items from a set of six items asking about the degree to which others (e.g., friends, family, people in general) prefer the poor over the rich. Participants responded on a scale from -3 ("Strongly prefer the rich to the poor") to 3 ("Strongly prefer the poor to the rich"). This measure was used in exploratory parallel regressions as a measure of others' evaluations of the poor versus rich.

(3) Cultural evaluations. Participants indicated the extent to which society at large evaluates the rich and the poor on one of three closely related dimensions: positivity (e.g., "How positive or negative do you feel towards the rich?"), warmth (e.g., "How warm or cold do you feel towards the rich?"), or likeability (e.g., How much do you like or dislike the rich?"), responding on a scale from 1 ("Strongly negative", "Cold", or "Strongly dislike") to 10 ("Strongly positive", "Warm", or "Strongly like"). The cultural evaluations measure consists of

the difference in evaluations for the poor minus the rich. This measure was used in the exploratory parallel regressions as an index of the cultural evaluations of the poor versus rich.

(4) Internal pressure. Separately for the rich and the poor, participants completed one item from a set of four items tapping into the extent to which making positive evaluations and avoiding negative evaluations of the rich/poor is consistent with their personal values. Participants responded on a scale from 1 ("Strongly disagree") to 6 ("Strongly agree"). The internal pressure measure consists of the difference for the poor minus for the rich. This measure was used in exploratory parallel regressions as a measure of internal pressure to adjust one's evaluations based on social status.

(5) Others' pressure. Separately for the rich and the poor, participants completed two items from a set of eight items asking about the extent to which the moderate their attitudes toward the rich/poor in order to gain approval or avoid condemnation from others. Participants responded on a scale from 1 ("Strongly disagree") to 6 ("Strongly agree"). The others' pressure measure consists of the difference for the poor minus for the rich. This measure was used in exploratory parallel regressions as a measure of external cultural pressure to adjust one's evaluations based on social status.

(6) Cultural pressure. Separately for the rich and the poor, participants completed two items from a set of eight items asking about perceptions of the average person's motivations and experience of cultural pressure to evaluate the rich/poor positively. Participants responded on a scale from 1 ("Not at all motivated" or "Strongly disagree") to 6 ("Strongly motivated" or "Strongly agree"). The cultural pressure measure consists of the difference for the poor minus for the rich. This measure was used in exploratory parallel regressions as a measure of external cultural pressure to adjust one's evaluations based on social status.

(7–8) Monopolar evaluations—ambivalence. Separately for each social class, participants responded to two monopolar items: "Thinking of only the positive things and not the negative, how positive are the rich?"; "Thinking of only the negative things and not the positive, how positive are the rich?" Participants responded on a scale from 1 ("Not at all positive/negative") to 6 ("Very positive/negative").

Ambivalence was computed separately for the rich and the poor as a weighted index of the minimum intensity of positive evaluations and negative evaluations divided by the difference in magnitude between positive and negative evaluations. Concretely, the formula was as follows: $\text{Ambivalence}_{\text{Rich}} = \text{minimum}(\text{positive}_{\text{Rich}}, \text{negative}_{\text{Rich}}) / (6 + \text{maximum}(\text{positive}_{\text{Rich}}, \text{negative}_{\text{Rich}}) - \text{minimum}(\text{positive}_{\text{Rich}}, \text{negative}_{\text{Rich}}))$. The index ranges from 0 to 1 (exclusive of 0), with larger scores indicating greater ambivalence, operationalized as similarly intense positive and negative scores. This ambivalence index gives greater weight to participants with equally high positive and negative scores compared to participants with equally low positive and negative scores. As for all other differences in explicit measures in this report, the difference score for ambivalence was computed as $\text{Ambivalence}_{\text{Poor}} - \text{Ambivalence}_{\text{Rich}}$.

(9) Relative personal preference. This measure consisted of a single item, "Which do you prefer, the rich or the poor?" Participants responded on a scale from -3 ("Strongly prefer the rich to the poor") to 3 ("Strongly prefer the poor to the rich").

(10) Gut reactions. This measure assessed gut reactions in separate items for the rich and the poor using a single item. For example, "People's gut reactions about a topic can be different from their feelings after they have had time to think about it. For example, someone who is trying to quit smoking might have a very positive gut reaction, but negative actual feelings toward smoking. Rate your gut reactions and actual feelings toward the topics below: Gut

reactions toward the rich." Participants responded on a scale from 1 ("Strongly negative") to 10 ("Strongly positive"). The gut reactions measure consists of the difference in gut reactions for the poor minus the rich.

(11) Actual feelings. This measure assessed actual feelings in separate items for the rich and the poor using a single item. Actual feelings items began with the same preface used for gut reactions items, but with the prompt to report actual feelings instead of gut reactions. Participants responded on a scale from 1 ("Strongly negative") to 10 ("Strongly positive"). The actual feelings measure consists of the difference in gut reactions for the poor minus the rich.

(12) Polarity. In two separate items, participants indicated the perceived consequences of liking the rich for liking the poor, and vice versa. For example, "Having positive feelings toward the rich implies having negative feelings toward the poor." Participants responded on a scale from 1 ("Strongly disagree") to 6 ("Strongly agree"). The polarity measure consists of the average polarity score for both items.

(13) Identity. Separately for the rich and the poor, participants indicated the extent to which they included these identities in their self-concept. Participants responded on a scale from 1 ("None at all") to 6 ("Very much"). The identity measure consists of the difference in identification with the poor minus identification with the rich.

(14) Self-concept centrality. Separately for the rich and the poor, participants completed one item from a set of four items tapping into the extent to which accepting the rich/poor is an important part of their self-concept. Participants responded on a scale from 1 ("Strongly disagree") to 6 ("Strongly agree"). The self-concept measure consists of the difference for the poor minus for the rich. This measure was used in exploratory parallel regressions.

(15) Attitudinal certainty. These items asked about participants' certainty about their feelings toward the rich and the poor. We did not analyze this measure.

(16) Personal importance. These items asked about the personal importance that participants placed on their feelings for the rich and the poor. We did not analyze this measure.

(17) Affective forecasting. These items asked about participants' expectations that their attitudes toward the rich and the poor might change over time. We did not analyze this measure.

(18) Mindfulness. These items asked participants about how much they think about their feelings toward the rich and the poor. We did not analyze this measure.

Protocol

Upon agreeing to participate in the study, AIID participants were randomly assigned to 1 of 95 different attitude domains. Of interest to the present analyses is the rich–poor domain (for other domains, see the full AIID study overview at https://osf.io/af5dn/). Three out of every four participants in the rich–poor domain completed an evaluative IAT that assessed associations between this domain and evaluative anchors (e.g., good, bad). The remaining participants in the rich–poor domain instead completed an identity IAT (anchors: self, other). Our analyses focus only on participants who completed the evaluative IAT. The order of the rich–poor IAT and the explicit measures on the same dimension (see above) was randomized across participants, such that approximately half of the participants completed the IAT first and the remainder completed the explicit measures first. Finally, participants completed 20 individual difference measures. These measures are not relevant to the present analyses. Because they were assessed after all other measures, we do not list them here, but they are available at https://osf.io/af5dn/.

Data Accessibility

All data and analysis scripts are provided on the Open Science Framework (Mattan & Cloutier, 2020) at https://osf.io/jcgyn/?view_only=23d26f11c46f4314bf42bae4631dea4a. At stage 1, this includes only data and scripts for the pilot analyses. However, at stage 2, data and scripts for the confirmatory analyses will be added.

Power Estimates Based on Analyses of Pilot Data

In this section, we first report on the results from the pilot analyses. We intend to report the results of our confirmatory analyses in the same fashion. After reporting the pilot results, we then provide power estimates via simulation (Lane & Hennes, 2018) for each hypothesis using parameters from our pilot analyses and our projected sample for confirmatory analyses.

Results from pilot analyses. Linear regressions were computed in R (for analysis scripts, see Mattan & Cloutier, 2020) to test for the effects of participant gender, income, and education on IAT D scores. For the pilot data, we conducted an omnibus regression model examining the effects of gender, income, education, and all possible interactions between these factors. We then followed up on this model, focusing on the Gender × Income interaction at low and high education levels (±1.5 *SD*).

Relationships and multicollinearity between independent variables. The pilot data provided no evidence of a difference between men and women in terms of income or education, $|t| < .62$, $p > .54$. However, we observed a significant correlation between our standardized predictors for income and education $r(173) = .22$, $p = .004$. Because the variance inflation factors (VIF) for all model terms in the omnibus model were below 1.5, our pilot data were not sufficiently impacted by multicollinearity to warrant orthogonalization of income and education. If we observe multicollinearity in our confirmatory dataset (i.e., if any VIF > 10), we will

orthogonalize the terms contributing to variance inflation. Because they tend to correlate, we anticipate this being most likely for income and education. In that case, we will orthogonalize income and education in all analyses by extracting the residuals for income in a model predicting education as a function of income. The residual variance for income, independent of its relationship with education, would be used in all subsequent analyses.

H1. Men (vs. women) will show greater pro-wealthy bias with increasing income.

Evidence constituting support for this hypothesis would be a significant Gender × Income interaction in the omnibus model (i.e., at the sample's mean education level). In our pilot analysis, this interaction was significant (Figure 1B), b = 0.132, SE = 0.055, $CI_{95\%}$ = [0.024, 0.240], $t(167)$ = 2.413, p = .017. Tests of simple effects indicated that men showed increasing pro-wealthy implicit bias with increasing income levels (H1A), b = 0.106, SE = 0.044, $CI_{95\%}$ = [0.019, 0.193], $t(167)$ = 2.402, p = .017. This effect of income was non-significant for women (H1B), b = -0.026, SE = 0.032, $CI_{95\%}$ = [-0.090, 0.038], $t(167)$ = -0.804, p = .423. Additional simple effects revealed that high-income men showed greater implicit pro-wealthy bias than high-income women, b = 0.282, SE = 0.101, $CI_{95\%}$ = [0.083, 0.482], $t(167)$ = 2.794, p = .006. This simple effect of gender was non-significant at mean income, b = 0.084, SE = 0.055, $CI_{95\%}$ = [-0.025, 0.193], $t(167)$ = 1.523, p = .130, and appeared to reverse at low income, b = -0.114, SE = 0.097, $CI_{95\%}$ = [-0.305, 0.077], $t(167)$ = -1.176, p = .241, but this reversal was non-significant.

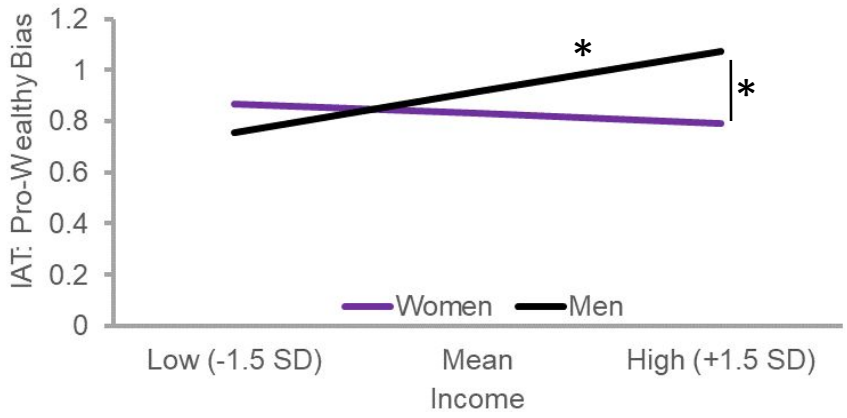

Figure 1. Pro-wealthy implicit bias as a function of the participant's gender and income in the pilot data. Analyses revealed a significant interaction such that men (black line) but not women (purple line) showed greater implicit pro-wealthy bias with increasing income. All significant simple effects are indicated with asterisks, $p < .05$.

H2. The Gender × Income × Education interaction. In our pilot data, the Gender × Income × Education interaction was non-significant, $b = -0.089$, $SE = 0.059$, $CI_{95\%} = [-0.205, 0.027]$, $t(167) = -1.521$, $p = .130$. Although this three-way interaction was non-significant, we nonetheless decomposed the interaction for the purpose of estimating the power of our follow-up models (see H2A and H2B). We focus on theoretically motivated simple effects in the text, but all significant simple effects are indicated in Figure 2. To facilitate interpretation of these exploratory decompositions, we note that the main effect of education level in the omnibus model for the pilot data was significant, with greater education being associated with greater implicit pro-wealthy bias, $b = 0.111$, $SE = 0.029$, $CI_{95\%} = [0.055, 0.168]$, $t(167) = 3.897$, $p < .001$.

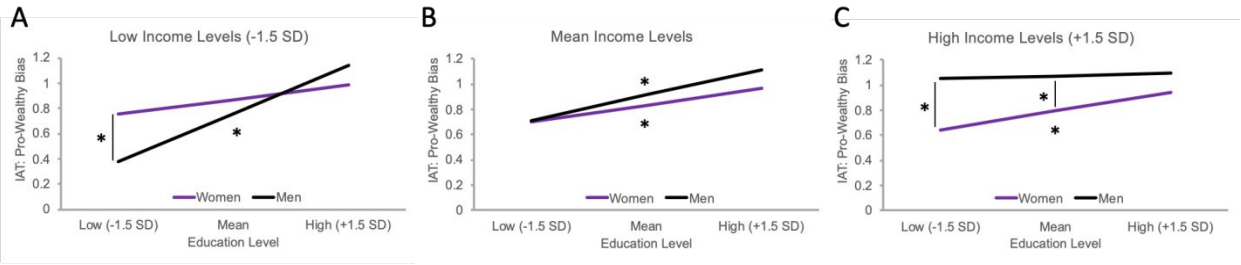

Figure 2. Pro-wealthy implicit bias as a function of the participant's gender, education, and income in the pilot data. All significant simple effects are indicated with asterisks, $p < .05$.

H2A & H2B. At low income levels, men (vs. women) will show greater pro-wealthy bias with increasing education. Initial evidence constituting support for this hypothesis would be a significant Gender × Education interaction in our first follow-up model (i.e., at 1.5 *SD* below the sample's mean income level). In our pilot analysis, this interaction was non-significant (Figure 2A), $b = 0.180$, $SE = 0.113$, $CI_{95\%} = [-0.043, 0.403]$, $t(167) = 1.591$, $p = .113$. We nonetheless followed up on this non-significant interaction with specific theoretically informed contrasts.

The gender roles and cultural fit accounts specifically predict that greater increasing education will lead to an attenuation of pro-wealthy bias among low-income women who face unique challenges in advancing their status through educational prestige (H2A). The strongest support for this prediction would take the form of a negative slope of education for low-income women. However, the simple effect of education on implicit pro-rich bias in low-income women trended in the opposite direction, although non-significant, $b = 0.075$, $SE = 0.066$, $CI_{95\%} = [-0.054, 0.205]$, $t(167) = 1.145$, $p = .254$. Nonetheless, given that increasing education led to a general increase in implicit pro-rich bias in the overall sample, it is noteworthy that low-income women (and high-income men—discussed below) were the only exceptions to this pattern. We

consider this partial support for the prediction that increasing education may attenuate implicit pro-rich bias among low-income women.

Not exclusive of the gender roles and cultural fit accounts, the precarious manhood account predicts that increasing education will have the opposite effect among low-income men, who may view education as a viable means of addressing chronic status insecurity en lieu of income (H2B). The strongest support for this prediction would take the form of a positive slope of education for low-income men. Indeed, we observed that low-income men showed a significant increase in implicit pro-rich bias as a function of education, $b = 0.255$, $SE = 0.092$, $CI_{95\%} = [0.073, 0.436]$, $t(167) = 2.771$, $p = .006$.

H2C & H2D. At high income levels, increasing education will further enhance implicit pro-wealthy bias. Evidence constituting support for this hypothesis would be a significant main effect of education in our second follow-up model (i.e., at 1.5 *SD* above the sample's mean income level). Although we do not have strong a priori predictions of a gender difference among high-income individuals, this predicted main effect may be subsumed by a Gender × Education interaction in the event that education increases implicit pro-wealthy to different degrees for women (H2C) and men (H2D).

In our pilot analysis, the effects of education and the Gender × Education interaction were both non-significant, $b = 0.058$, $SE = 0.048$, $CI_{95\%} = [-0.038, 0.153]$, $t(167) = 1.193$, $p = .235$, and $b = -0.089$, $SE = 0.097$, $CI_{95\%} = [-0.279, 0.102]$, $t(167) = -0.917$, $p = .360$, respectively. To parallel the contrasts reported above for low-income individuals, we tested for effects of education separately for men and women (see Figure 2C). We found that increasing education increased implicit pro-rich bias for women, $b = 0.102$, $SE = 0.048$, $CI_{95\%} = [0.008,$

0.196], $t(167) = 2.136$, $p = .034$, but not for men, $b = 0.013$, $SE = 0.084$, $CI_{95\%} = [-0.153, 0.179]$, $t(167) = 0.158$, $p = .874$.

Power estimates. To estimate power for all predicted effects, we ran simulations on our pilot sample in R (R Core Team, 2019) based on scripts that we adapted from Lane and Hennes (2018). In summary, we generated 1,000 simulated participant-level datasets based on the sample characteristics (e.g., gender ratio, means and standard deviations for income and education) as well as the betas and standard error from the omnibus model used on the pilot data. These simulated datasets included a correlation of .22 between income and education. Each simulated dataset contained 1,000 participants, which is well below our anticipated final sample size. We opted for a smaller sample size to provide a more conservative projected estimate of power. In addition to our main simulations of the pilot data, we also computed parallel sensitivity analyses to provide a sense of how small the Gender × Income × Education interaction could be before falling below 80% power. All other parameters in these parallel simulations were the same as in the main simulations of the pilot data. Lastly, we will also report actual power estimates and sensitivity analyses based on the final sample size after conducting our planned analyses on the confirmatory dataset.

For each simulated dataset, we ran the omnibus model and follow-up models (see above) to test our predicted effects. Power for each test is defined as the proportion of simulations that resulted in a significant effect for that test. Analysis scripts for these simulations are available on the Open Science Framework (Mattan & Cloutier, 2020).

For H1, our projected power estimate for the Gender × Income interaction was 100%. For men (H1A), the simple effect of income was powered at 100%. For women (H1B), the simple effect of income was powered at 48.3%.

For H2, our projected power estimate for the Gender × Income × Education interaction was 98.1%. Because pilot samples may overestimate the magnitude of effects (Albers & Lakens, 2018), particularly for higher order interactions (Simonsohn, 2014), we conducted additional sensitivity simulations for this three-way interaction in order to determine the smallest detectable effect size that could be detected 80% of the time. With all other effect size parameters being equal, these additional simulations indicated that a conservative sample size of 1,000 participants would afford sufficient power (i.e., 80.5%) to detect a significant three-way interaction that is as much as 30% smaller than the effect size observed in our pilot sample.

Our projected power estimate for the Gender × Education interaction at 1.5 *SD* below the sample's mean income level was 99.1%. For low-income women (H2A), the simple effect of education was powered at 88.9%. For low-income men (H2B), the simple effect of education was powered at 100%.

Our projected power estimate for the main effect of education at 1.5 *SD* above the sample's mean income level was 80.1%. Our projected power estimate for the Gender × Education interaction for high-income participants was 57.7%. For high-income women (H2C), the simple effect of education was powered at 98.4%. For high-income men (H2D), the simple effect of education was powered at 5.7%.

Contextual Analyses of Individual Differences in the Pilot Data

Although we did not have control over the experimental design for the AIID dataset, the dataset contains many explicit measures of personal and perceived cultural attitudes toward the rich and the poor. These measures can help provide greater context for the sample's implicit pro-

wealthy bias and also allow for comparisons between self-report and implicit measures of pro-wealthy bias. In the final confirmatory dataset, we will conduct and report the same analyses.

Sample descriptives. In the pilot dataset, we found little evidence of perceived polarity in attitudes about the rich and poor; participants tended to disagree that evaluating the poor negatively necessarily meant having positive attitudes toward the rich (and vice versa). This is consistent with findings from Horwitz and Dovidio (2017) showing that pro-rich attitudes were not predicted by anti-poor attitudes, and vice versa. With respect to individuals' attitudes, we observed a pro-rich bias for virtually all measures tapping into personal beliefs about the rich and the poor, consistent with findings from a more recent dataset (Varnum, 2013). Interestingly, for measures that probed perceptions of others' and cultural attitudes toward the rich and poor, participants (1) believed that society at large generally favors the poor over the rich, and (2) felt pressure from others to moderate their explicit attitudes about the poor more so than their attitudes about the rich. In sum, participants tended to value the rich at the individual level but evidently believed their attitudes contrasted with those of society at large. For full descriptive statistics, see Table 2.

Table 2. Descriptive Statistics for Select Individual Difference Measures.

Measures (Poor - Rich)	*M*	*CI 95%*	*t*	*df*	*p*
1. Personal Evaluations	-0.706	[-1.078, -0.333]	-3.736	162	**< .001**
2. Others' Evaluations	1.022	[0.870, 1.174]	13.246	158	**< .001**
3. Cultural Evaluations	2.988	[2.474, 3.502]	11.475	162	**< .001**
4. Internal Pressure	-0.691	[-0.953, -0.430]	-5.214	161	**< .001**
5. Others' Pressure	0.141	[0.000, 0.282]	1.975	162	**.050**
6. Cultural Pressure	-0.031	[-0.206, 0.145]	-0.346	162	.730
7/8. Ambivalence	-0.034	[-0.099, 0.031]	-1.033	63	.306
9. Relative Personal Preference	0.019	[-0.198, 0.236]	0.171	159	.865
10. Gut Reactions	-0.327	[-0.778, 0.124]	-1.434	164	.154
11. Actual Feelings	-0.594	[-1.001, -0.187]	-2.881	164	**.004**
12. Overall Polarity (Rich & Poor) *	2.68	[2.290, 3.071]	-4.198	60	**< .001**
13. Identity	-0.051	[-0.467, 0.364]	-0.246	77	.807
14. Self-Concept Centrality	-0.945	[-1.199, -0.691]	-7.348	162	**< .001**

Note. Except for overall polarity, all t-tests are one-sample t-tests against zero (i.e., no difference between the poor and the rich).
* One-sample t-test against 3.5 (neither agree nor disagree)

Correlations with IAT. To determine whether there were any relationships between implicit pro-rich bias and explicit measures of attitudes about the poor and rich, we conducted correlations between the IAT and difference scores (poor – rich) on a host of explicit measures (see Table 3). These correlations revealed that greater pro-rich bias on the IAT was predicted by *less* explicit pro-rich bias and internal pressure to regulate pro-rich bias. This was mostly the case for explicit measures tapping into individual attitudes rather than perceptions about others' or cultural attitudes.

Table 3. Correlations with IAT D Scores (Pro-Rich Bias).

Measures (Poor - Rich)	*r*	*n*	*p*
1. Personal Evaluations	**.26**	**163**	**.001**
2. Others' Evaluations	**.16**	**159**	**.039**
3. Cultural Evaluations	.02	163	.798
4. Internal Pressure	**.16**	**162**	**.036**
5. Others' Pressure	-.06	163	.436
6. Cultural Pressure	.03	163	.683
7/8. Ambivalence	.10	64	.440
9. Relative Personal Preference	**.19**	**160**	**.017**
10. Gut Reactions	**.22**	**165**	**.004**
11. Actual Feelings	**.24**	**165**	**.002**
12. Overall Polarity (Rich & Poor)	-.02	61	.897
13. Identity	**.24**	**78**	**.033**
14. Self-Concept Centrality	.11	163	.175

Note . Polarity reflects average perceived polarity for rich and poor.

Although these correlations may seem puzzling at first glance, we are not the first to find divergent patterns of evaluation based on wealth using implicit versus explicit measures. Previous research using the IAT has found that poor (Rudman et al., 2002) and middle-class (Horwitz & Dovidio, 2017) participants prefer their respective ingroups relative to the rich on explicit measures. However, both groups show a clear pro-rich bias on the rich-poor and rich-middle-class IATs, respectively. Using both correlational and categorical analytic approaches, Rudman and colleagues (2002) observed a status-based gradient in how much people showed a divergence in their implicit and explicit ingroup biases. Generally, members of powerful high-status groups (e.g., rich/White people) were consistent in preferring their own group (vs. poor/Asian people) on both implicit and explicit measures. Low-power groups (e.g., the poor, minorities) tended showed increasingly large discrepancies on implicit and explicit measures. The least powerful group (viz., poor people) showed both a sizeable explicit ingroup bias and the strongest outgroup bias on the IAT relative to all other low-power groups. In fact, poor

participants' implicit preference for the rich was comparable in magnitude to implicit ingroup biases from participants belonging to powerful high-status groups. Taken together, results from Rudman and colleagues suggest that this negative relationship between implicit and explicit measures may be due to low participant status. Exploratory analyses of our pilot dataset suggest the opposite relationship. Higher status participants (defined in terms of income or education) tended to show greater divergence in their implicit and explicit attitudes, showing higher pro-rich bias on the IAT and lower pro-rich bias on explicit measures.

In summary, divergences in explicit measures and IAT scores are sometimes observed. Based on the MODE model (Fazio & Olson, 2003), it is thought that these differences emerge when spontaneous group associations (assessed with the IAT) receive further elaboration as participants formulate their self-reported evaluation of a given group (e.g., the rich). The degree to which self-reported evaluations of the rich and poor are determined by initial spontaneous associations depends on whether the participants (1) are motivated to adjust their self-reported evaluations, and (2) have the time and mental resources to do so. Given that participants in the present study (1) reported cultural norms about evaluating the poor more favorably, and (2) were not constrained by time limits or additional cognitive load when they completed the explicit measures, it would seem that the conditions could increase the chances of observing diverging patterns of implicit and explicit bias. We refrain from lengthy speculation about the mechanism underlying this unexpected divergence, pending replication. Despite occasional divergences between implicit and explicit measures, previous work has found that implicit measures can be better predictors of real-world behaviors than are explicit measures, particularly when participants are motivated and able to modulate their explicit evaluations (Greenwald, Banaji, & Nosek, 2015; Greenwald, Poehlman, Uhlmann, & Banaji, 2009). For example, the rich-middle-

class IAT (rather than parallel self-report measures) predicted leniency on a rich driver who was responsible for a car accident (Horwitz & Dovidio, 2017).

Parallel regressions. At the encouragement of an expert reviewer, we also conducted additional parallel regressions meant to predict scores on a selection of the aforementioned difference scores on several self-report items. Where possible, we focused on multi-item rather than single-item measures of attitudes toward the rich and the poor (see exploratory measures for descriptions). Separately for the self, generic "others", and society at large we analyzed (1) differences in explicit evaluations (poor minus rich), and (2) differences in felt pressure to adjust one's evaluations (poor minus rich). This resulted in a total of six parallel models. We also included two additional models examining: (1) the difference in the degree to which accepting the poor (vs. rich) is important to the participant's self-concept, and (2) the difference in ambivalence felt toward the rich (vs. poor). Only two models showed any significant effects.

Personal evaluations. In the model examining the participant's own attitudes toward the poor versus rich, we observed a significant Income × Education interaction, b = -0.5926, SE = 0.2159, $CI_{95\%}$ = [-1.019, -0.166], $t(155)$ = -2.745, p = .007. At low and average education levels, increasing income was associated with reduced explicit pro-rich bias, b = 1.185, SE = 0.399, $CI_{95\%}$ = [0.397, 1.972], $t(155)$ = 2.972, p = .003. At high education levels, this pattern appeared to reverse, but this was non-significant, b = -0.593, SE = 0.367, $CI_{95\%}$ = [-1.318, 0.132], $t(155)$ = -1.616, p = .108. At low income levels, increasing education was associated with reduced pro-rich bias, b = 1.172, SE = 0.406, $CI_{95\%}$ = [0.371, 1.973], $t(155)$ = 2.891, p = .004. At high income levels, this pattern appeared to reverse, but this was non-significant, b = -0.606, SE = 0.359, $CI_{95\%}$ = [-1.316, 0.105], $t(155)$ = -1.684, p = .094. In summary, the greatest pro-rich explicit bias was observed for individuals with the lowest income and education, which is the opposite of

what we found in our analyses of IAT data. In addition to this Income × Education interaction, we also observed a non-significant Gender × Education interaction, $b = 0.791$, $SE = 0.410$, $CI_{95\%} = [-0.018, 1.601]$, $t(155) = 1.931$, $p = .055$. This non-significant interaction was characterized by an apparent decrease in pro-rich bias with increasing education levels, but only for men, $b = 0.679$, $SE = 0.335$, $CI_{95\%} = [0.017, 1.341]$, $t(155) = 2.025$, $p = .045$, and not women, $b = -0.112$, $SE = 0.236$, $CI_{95\%} = [-0.578, 0.353]$, $t(155) = -0.476$, $p = .635$. All other effects in this analysis of the participant's self-reported evaluations of the poor versus rich were non-significant, $p > .07$.

Ambivalence. In the model examining the participant's ambivalence toward the poor versus rich, we observed a significant main effect of education, $b = -0.084$, $SE = 0.040$, $CI_{95\%} = [-0.163, -0.005]$, $t(56) = -2.118$, $p = .039$. Greater education was associated with a stronger mixture of positive and negative feelings toward the rich compared to the poor. This is intriguing because greater education was also associated with generally greater pro-rich implicit bias. In other words, greater ambivalence may not necessarily correspond to a reduction in pro-rich implicit bias. In addition to this main effect of education, we also observed a significant Gender × Income interaction, $b = 0.176$, $SE = 0.088$, $CI_{95\%} = [0.000, 0.352]$, $t(56) = 2.008$, $p = .050$. This interaction was characterized by a non-significant decline in ambivalence toward the rich (vs. poor) with increasing income for men, $b = 0.146$, $SE = 0.074$, $CI_{95\%} = [-0.002, 0.295]$, $t(56) = 1.972$, $p = .054$. This same relationship was non-significant for women, $b = -0.030$, $SE = 0.047$, $CI_{95\%} = [-0.123, 0.064]$, $t(56) = -0.636$, $p = .527$. This overall pattern is consistent with increasing implicit bias as a function of increasing income in men but not women (see main analyses). All other effects in this analysis of ambivalence toward the rich versus the poor were non-significant, $p > .13$.

Self-concept centrality. Like the other models, the self-concept model showed no significant effects. However, there were two non-significant main effects that are worth comment. Should these main effects become reliable in the larger and therefore better powered confirmatory sample, they may provide additional evidence of internal pressure to modulate status bias in high-status individuals. Both gender (men > women) and increasing income appeared to predict increasing agreement with the notion that accepting the poor (vs. rich) is important to one's self-concept, $b = 0.540$, $SE = 0.286$, $CI_{95\%} = [-0.025, 1.105]$, $t(56) = 1.889$, $p = .061$, and $b = 0.267$, $SE = 0.144$, $CI_{95\%} = [-0.018, 0.551]$, $t(56) = 1.850$, $p = .066$, respectively. Effects from all other models were non-significant, $p > .08$.

Summary of parallel regressions. In conclusion, the parallel regression analyses of individual differences revealed relatively few effects in the pilot data. Among the few effects we observed, it would seem that increasing status is generally associated with diminished explicit pro-rich bias, but only for composite measures tapping into participants' own explicit evaluations of the poor and rich. Although our sample for pilot analyses of attitudinal ambivalence was relatively small, we observed here a potentially important difference with respect to income and education. Whereas increasing education was associated with greater ambivalence toward the rich, increasing income was associated with reduced ambivalence toward the rich among male participants. In contrast with the findings for measures of individual attitudes, we observed no effects of external pressure to modulate class-based attitudes (cf. personal vs. others' and cultural attitudes). At this stage, our best interpretation of these results is that individuals with higher status (viz., through some combination of income, education, and/or gender) tend to downregulate their positive evaluations of the rich, perhaps out of greater ambivalence among the educated and/or a desire to minimize apparent privilege (Kay & Jost, 2003) among the

wealthy. Although we do not have the means to test the precise mechanism for status-based downregulation of pro-rich bias in this dataset, we intend to follow up on these analyses in the main confirmatory dataset, providing the reader with (1) a replication of these piloted findings, (2) an extended discussion of how they may shed further light on the implicit/explicit divergence in pro-rich biases that occurs for high-status participants, and (3) additional discussion of future directions that may help uncover the mechanism underlying status-based downregulation of pro-rich bias.

Authors' Contributions

BDM conducted all analyses, drafted the manuscript, and provided major revisions; JC provided critical feedback and revised the manuscript. Both authors gave final approval for publication and agree to be held accountable for the work performed therein.

Acknowledgements

The authors would like to thank Jennifer T. Kubota for feedback on the analysis plan. The authors would also like to thank the AIID team for facilitating access to the data analyzed in this report. Finally, the authors thank Michael W. Kraus and an anonymous reviewer for their invaluable insights and suggestions.

References

Albers, C., & Lakens, D. (2018). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology*, *74*(September 2017), 187–195. https://doi.org/10.1016/j.jesp.2017.09.004

Andersen, R., & van de Werfhorst, H. G. (2010). Education and occupational status in 14 countries: The role of educational institutions and labour market coordination. *British Journal of Sociology*, *61*, 336–355. https://doi.org/10.1111/j.1468-4446.2010.01315.x

Anderson, C., Willer, R., Kilduff, G. J., & Brown, C. E. (2012). The origins of deference: When do people prefer lower status? *Journal of Personality and Social Psychology*, *102*, 1077–1088. https://doi.org/10.1037/a0027409

Baumeister, R. F., & Sommer, K. L. (1997). What do men want? Gender differences and two spheres of belongingness: Comment on Cross and Madson (1997). *Psychological Bulletin*, *122*, 38–44. https://doi.org/10.1037/0033-2909.122.1.38

Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status characteristics and social interaction. *American Sociological Review*, *37*, 241–255. https://doi.org/10.2307/2093465

Bjornsdottir, R. T., & Rule, N. O. (2017). The visibility of social class from facial cues. *Journal of Personality and Social Psychology*, *113*, 530–546. https://doi.org/10.1037/pspa0000091

Blickenstaff, J. C. (2005). Women and science careers: Leaky pipeline or gender filter? *Gender and Education*, *17*, 369–386. https://doi.org/10.1080/09540250500145072

Boukarras, S., Era, V., Aglioti, S. M., & Candidi, M. (2019). Modulation of preference for abstract stimuli following competence-based social status primes. *Experimental Brain*

Research. https://doi.org/10.1007/s00221-019-05702-z

Brand, J. E., & Xie, Y. (2010). Who benefits most from college? Evidence for negative selection in heterogeneous economic returns to higher education. *American Sociological Review*, *75*, 273–302. https://doi.org/10.1177/0003122410363567

Breen, R., & Jonsson, J. O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*, *31*, 223–243. https://doi.org/10.1146/annurev.soc.31.041304.122232

Buss, D. M. (1981). Sex differences in the evaluation and performance of dominant acts. *Journal of Personality and Social Psychology*, *40*, 147–154. https://doi.org/10.1037/0022-3514.40.1.147

Caswell, T. A., Bosson, J. K., Vandello, J. A., & Sellers, J. G. (2014). Testosterone and men's stress responses to gender threats. *Psychology of Men & Masculinity*, *15*, 4–11. https://doi.org/10.1037/a0031394

Cheng, J. T., & Tracy, J. L. (2013). The impact of wealth on prestige and dominance rank relationships. *Psychological Inquiry*, *24*, 102–108. https://doi.org/10.1080/1047840X.2013.792576

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*, 103–125. https://doi.org/10.1037/a0030398

Cloutier, J., Ambady, N., Meagher, T., & Gabrieli, J. D. E. (2012). The neural substrates of

person perception: Spontaneous use of financial and moral status knowledge. *Neuropsychologia*, *50*, 2371–2376. https://doi.org/10.1016/j.neuropsychologia.2012.06.010

Cloutier, J., & Gyurovski, I. (2014). Ventral medial prefrontal cortex and person evaluation: Forming impressions of others varying in financial and moral status. *NeuroImage*, *100*, 535–543. https://doi.org/10.1016/j.neuroimage.2014.06.024

Collinson, D. L., & Hearn, J. (2005). Men and masculinities in work, organizations and management. In M. S. Kimmel, J. Hearn, & R. W. Connell (Eds.), *Handbook of Studies on Men and Masculinities* (pp. 289–310). Thousand Oaks, CA: Sage Publications.

Connell, R. W. (1987). *Gender and power*. Cambridge, UK: Polity Press.

Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic masculinity: Rethinking the concept. *Gender & Society*, *19*, 829–859. https://doi.org/10.1177/0891243205278639

Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review*, *108*, 593–623. https://doi.org/10.1037//0033-295X.108.3.593

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*, 631–648. https://doi.org/10.1037/0022-3514.92.4.631

Cunningham, W. A., Nezlek, J. B., & Banaji, M. R. (2004). Implicit and explicit ethnocentrism: Revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin*, *30*, 1332–1346. https://doi.org/10.1177/0146167204264654

Desmarais, S., & Curtis, J. (1997). Gender and perceived pay entitlement: Testing for effects of experience with income. *Journal of Personality and Social Psychology*, *72*, 141–150.

https://doi.org/10.1037/0022-3514.72.1.141

Dovidio, J. F., Ellyson, S. L., Keating, C. F., Heltman, K., & Brown, C. E. (1988). The relationship of social power to visual displays of dominance between men and women. *Journal of Personality and Social Psychology*, *54*, 233–242. https://doi.org/10.1037/0022-3514.54.2.233

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Laurence Erlbaum.

Eagly, A. H. (2009). The his and hers of prosocial behavior: An examination of the social psychology of gender. *American Psychologist*, *64*, 644–658. https://doi.org/10.1037/0003-066X.64.8.644

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

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*, 735–754. https://doi.org/10.1037/0022-3514.46.4.735

Eagly, A. H., & Wood, W. (1982). Inferred sex differences in status as a determinant of gender stereotypes about social influence. *Journal of Personality and Social Psychology*, *43*, 915–928. https://doi.org/10.1037/0022-3514.43.5.915

Eagly, A. H., & Wood, W. (1991). Explaining sex differences in behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin*, *17*, 306–315. https://doi.org/10.1177/0146167291173011

Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327. https://doi.org/10.1146/annurev.psych.54.101601.145225

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist*, *48*, 621–628. https://doi.org/10.1037/0003-066X.48.6.621

Fiske, S. T. (2010). Interpersonal stratification: Status, power, and subordination. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 941–982). Hoboken, NJ: Wiley.

Fiske, S. T., Dupree, C. H., Nicolas, G., & Swencionis, J. K. (2016). Status, power, and intergroup relations: The personal is the societal. *Current Opinion in Psychology*, *11*, 44–48. https://doi.org/10.1016/j.copsyc.2016.05.012

Furnham, A. (1984). Many sides of the coin: The psychology of money usage. *Personality and Individual Differences*, *5*, 501–509. https://doi.org/10.1016/0191-8869(84)90025-4

Garcia, S. M., Weaver, K., & Chen, P. (2018). The status signals paradox. *Social Psychological and Personality Science*. https://doi.org/10.1177/1948550618783712

Gilmore, D. D. (1990). *Manhood in the Making*. New Haven, CT: Yale University Press.

Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*, 491–512. https://doi.org/10.1037/0022-3514.70.3.491

Glick, P., & Fiske, S. T. (2001). Ambivalent sexism. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 33, pp. 115–188).

https://doi.org/10.4135/9781483384269.n26

Goh, J. X., & Hall, J. A. (2015). Nonverbal and verbal expressions of men's sexism in mixed-gender interactions. *Sex Roles*, *72*, 252–261. https://doi.org/10.1007/s11199-015-0451-7

Gordon, M. D. (1980). The role of referees in scientific communication. In J. Hartley (Ed.), *The Psychology of Written Communication* (pp. 263–275). London, UK: Kogan Page.

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology*, *108*, 553–561. https://doi.org/10.1037/pspa0000016

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*, 197–216. https://doi.org/10.1037/0022-3514.85.2.197

Greenwald, A. G., Poehlman, T. A., Uhlmann, E., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. https://doi.org/10.1037/a0015575

Hall, J. A., Coats, E. J., & LeBeau, L. S. (2005). Nonverbal behavior and the vertical dimension of social relations. *Psychological Bulletin*, *131*, 898–924. https://doi.org/10.1037/0033-2909.131.6.898

Hays, N. A. (2013). Fear and loving in social hierarchy: Sex differences in preferences for power

versus status. *Journal of Experimental Social Psychology*, *49*, 1130–1136. https://doi.org/10.1016/j.jesp.2013.08.007

Henry, P. J., & Hardin, C. D. (2006). The contact hypothesis revisited: Status bias in the reduction of implicit prejudice in the United States and Lebanon. *Psychological Science*, *17*, 862–868. https://doi.org/10.1037/024053

Horwitz, S. R., & Dovidio, J. F. (2017). The rich—love them or hate them? Divergent implicit and explicit attitudes toward the wealthy. *Group Processes & Intergroup Relations*, *20*, 3–31. https://doi.org/10.1177/1368430215596075

Hussey, I., Hughes, S., Lai, C. K., Ebersole, C. R., Axt, J., & Nosek, B. A. (2019). *Attitudes, Identities, and Individual Differences (AIID) Study*. https://doi.org/10.17605/OSF.IO/PCJWF

Jack, A. A. (2019). *The Privileged Poor: How Elite Colleges are Failing Poor Students*. Cambridge, MA: Harvard University Press.

Johnson, S. E., Richeson, J. A., & Finkel, E. J. (2011). Middle class and marginal? Socioeconomic status, stigma, and self-regulation at an elite university. *Journal of Personality and Social Psychology*, *100*, 838–852. https://doi.org/10.1037/a0021956

Jury, M., Aelenei, C., Chen, C., Darnon, C., & Elliot, A. J. (2019). Examining the role of perceived prestige in the link between students' subjective socioeconomic status and sense of belonging. *Group Processes & Intergroup Relations*, *22*, 356–370. https://doi.org/10.1177/1368430219827361

Kraus, M. W., Callaghan, B., & Ondish, P. (2019). Social class as culture. In S. Kitayama & D.

Cohen (Eds.), *Handbook of Cultural Psychology* (2nd ed., pp. 721–747). Retrieved from https://michaelkraus.squarespace.com/s/social-class-as-culture-en34.pdf

Kraus, M. W., Park, J. W., & Tan, J. J. X. (2017). Signs of social class: The experience of economic inequality in everyday life. *Perspectives on Psychological Science*, *12*, 422–435. https://doi.org/10.1177/1745691616673192

Kraus, M. W., Piff, P. K., & Keltner, D. (2011). Social class as culture: The convergence of resources and rank in the social realm. *Current Directions in Psychological Science*, *20*, 246–250. https://doi.org/10.1177/0963721411414654

Kraus, M. W., Torrez, B., Park, J. W., & Ghayebi, F. (2019). Evidence for the reproduction of social class in brief speech. *Proceedings of the National Academy of Sciences*, *116*, 22998–23003. https://doi.org/10.1073/pnas.1900500116

Kubota, J. T., Peiso, J., Marcum, K., & Cloutier, J. (2017). Intergroup contact throughout the lifespan modulates implicit racial biases across perceivers' racial group. *PLoS One*, *12*(7), e0180440. https://doi.org/10.1371/journal.pone.0180440

Kunstman, J. W., Plant, E. A., & Deska, J. C. (2016). White ≠ poor: Whites distance, derogate, and deny low-status ingroup members. *Personality and Social Psychology Bulletin*, *42*, 230–243. https://doi.org/10.1177/0146167215623270

Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationships research. *Journal of Social and Personal Relationships*, *35*, 7–31. https://doi.org/10.1177/0265407517710342

Lockheed, M. E., & Hall, K. P. (1976). Conceptualizing sex as a status characteristic:

Applications to leadership training strategies. *Journal of Social Issues*, *32*, 111–124. https://doi.org/10.1111/j.1540-4560.1976.tb02600.x

Lott, B. (2012). The social psychology of class and classism. *American Psychologist*, *67*, 650–658. https://doi.org/10.1037/a0029369

Mahalik, J. R., Locke, B. D., Ludlow, L. H., Diemer, M. A., Scott, R. P. J., Gottfried, M., & Freitas, G. (2003). Development of the Conformity to Masculine Norms Inventory. *Psychology of Men & Masculinity*, *4*, 3–25. https://doi.org/10.1037/1524-9220.4.1.3

Martin, T., & Kirkcaldy, B. (1998). Gender differences on the EPQ-R and attitudes to work. *Personality and Individual Differences*, *24*, 1–5. https://doi.org/10.1016/S0191-8869(97)00143-8

Masser, B. M., & Abrams, D. (2004). Reinforcing the glass ceiling: The consequences of hostile sexism for female managerial candidates. *Sex Roles*, *51*, 609–615. https://doi.org/10.1007/s11199-004-5470-8

Masser, B. M., Viki, G. T., & Power, C. (2006). Hostile sexism and rape proclivity amongst men. *Sex Roles*, *54*, 565–574. https://doi.org/10.1007/s11199-006-9022-2

Mattan, B. D., & Cloutier, J. (2020). *Pro-rich implicit bias as a function of gender, income, and education*. Retrieved from osf.io/jcgyn

Mattan, B. D., Kubota, J. T., & Cloutier, J. (2017). How social status shapes person perception and evaluation: A social neuroscience perspective. *Perspectives on Psychological Science*, *12*, 468–507. https://doi.org/10.1177/1745691616677828

Mattan, B. D., Kubota, J. T., Li, T., Venezia, S. A., & Cloutier, J. (2019). Implicit evaluative

biases toward targets varying in race and socioeconomic status. *Personality and Social Psychology Bulletin*, *45*, 1512–1527. https://doi.org/10.1177/0146167219835230

Mattan, B. D., Wei, K. Y., Cloutier, J., & Kubota, J. T. (2018). The social neuroscience of race-based and status-based prejudice. *Current Opinion in Psychology*, *24*, 27–34. https://doi.org/10.1016/j.copsyc.2018.04.010

Mendoza-Denton, R., Downey, G., Purdie, V. J., Davis, A., & Pietrzak, J. (2002). Sensitivity to status-based rejection: Implications for African American students' college experience. *Journal of Personality and Social Psychology*, *83*, 896–918. https://doi.org/10.1037/0022-3514.83.4.896

Moore-Berg, S. L., Karpinski, A., & Plant, E. A. (2017). Quick to the draw: How suspect race and socioeconomic status influences shooting decisions. *Journal of Applied Social Psychology*, *47*, 482–491. https://doi.org/10.1111/jasp.12454

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*, 16474–16479. https://doi.org/10.1073/pnas.1211286109

Moss-Racusin, C. A., Phelan, J. E., & Rudman, L. A. (2010). When men break the gender rules: Status incongruity and backlash against modest men. *Psychology of Men & Masculinity*, *11*, 140–151. https://doi.org/10.1037/a0018093

Mullen, A. L., & Baker, J. (2015). Participation without parity in U.S. higher education: Gender, fields of study, and institutional selectivity. *NASPA Journal About Women in Higher Education*, *8*, 172–188. https://doi.org/10.1080/19407882.2015.1057167

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, *22*, 103–122. https://doi.org/10.1002/ejsp.2420220202

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*, 101–115. https://doi.org/10.1037//1089-2699.6.1.101

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*, 2600–2606. https://doi.org/10.1073/pnas.1708274114

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II . Method variables and construct validity. *Personality and Social Psychology Bulletin*, *31*, 166–180. https://doi.org/10.1177/0146167204271418

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., … Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*, 36–88. https://doi.org/10.1080/10463280701489053

Oakes, J. M., & Rossi, P. H. (2003). The measurement of SES in health research: Current practice and steps toward a new approach. *Social Science and Medicine*, *56*, 769–784. https://doi.org/10.1016/S0277-9536(02)00073-4

Oh, D., Shafir, E., & Todorov, A. (2019). Economic status cues from clothes affect perceived competence from faces. *Nature Human Behaviour*.

Peters, D. P., & Ceci, S. J. (1982). Peer-review practices of psychological journals: The fate of published articles, submitted again. *Behavioral and Brain Sciences*, *5*, 187–195. https://doi.org/10.1017/S0140525X00011183

Petty, R. E., Briñol, P., Tormala, Z. L., & Jarvis, W. B. G. (2006). Implicit ambivalence from attitude change: An exploration of the PAST model. *Journal of Personality and Social Psychology*, *90*(1), 21–41. https://doi.org/10.1037/0022-3514.90.1.21

Posselt, J. R., & Grodsky, E. (2017). Graduate education and social stratification. *Annual Review of Sociology*, *43*, 353–378. https://doi.org/10.1146/annurev-soc-081715-074324

Prince, M. (1993). Women, men, and money styles. *Journal of Economic Psychology*, *14*, 175–182. https://doi.org/10.1016/0167-4870(93)90045-M

R Core Team. (2019). *R: A language and environment for statistical computing*. Retrieved from https://www.r-project.org/

Ridgeway, C. L. (1991). The social construction of status value: Gender and other nominal characteristics. *Social Forces*, *70*, 367–386. https://doi.org/10.1093/sf/70.2.367

Ridgeway, C. L., & Balkwell, J. W. (1997). Group processes and the diffusion of status beliefs. *Social Psychology Quarterly*, *60*, 14–31. https://doi.org/10.2307/2787009

Rivera, L. A. (2011). Ivies, extracurriculars, and exclusion: Elite employers' use of educational credentials. *Research in Social Stratification and Mobility*, *29*, 71–90. https://doi.org/10.1016/j.rssm.2010.12.001

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., & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, *87*, 157–176. https://doi.org/10.1037/0022-3514.87.2.157

Rudman, L. A., Feinberg, J., & Fairchild, K. (2002). Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. *Social Cognition*, *20*, 294–320. https://doi.org/10.1521/soco.20.4.294.19908

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*, 1004–1009. https://doi.org/10.1037/0022-3514.77.5.1004

Rudman, L. A., & Glick, P. (2008). *The Social Psychology of Gender: How Power and Intimacy Shape Gender Relations*. New York, NY: Guilford Press.

Rudman, L. A., & Kilianski, S. E. (2000). Implicit and explicit attitudes toward female authority. *Personality and Social Psychology Bulletin*, *26*, 1315–1328. https://doi.org/10.1177/0146167200263001

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

Sewell, W. H., Haller, A. O., & Ohlendorf, G. W. (1970). The educational and early

occupational status attainment process: Replication and revision. *American Sociological Review*, *35*, 1014–1027. https://doi.org/10.2307/2093379

Shariff, A. F., & Tracy, J. L. (2009). Knowing who's boss: Implicit perceptions of status from the nonverbal expression of pride. *Emotion*, *9*, 631–639. https://doi.org/10.1037/a0017089

Sheltzer, J. M., & Smith, J. C. (2014). Elite male faculty in the life sciences employ fewer women. *Proceedings of the National Academy of Sciences*, *111*, 10107–10112. https://doi.org/10.1073/pnas.1403334111

Shin, H. (2017). Friendship dynamics of adolescent aggression, prosocial behavior, and social status: The moderating role of gender. *Journal of Youth and Adolescence*, *46*, 2305–2320. https://doi.org/10.1007/s10964-017-0702-8

Simonsohn, U. (2014). No-way interactions. *The Winnower*, *5*, e142559. https://doi.org/10.15200/winn.142559.90552

Stephens, N. M., Townsend, S. S. M., & Dittmann, A. G. (2019). Social-class disparities in higher education and professional workplaces: The role of cultural mismatch. *Current Directions in Psychological Science*, *28*, 67–73. https://doi.org/10.1177/0963721418806506

Stephens, N. M., Townsend, S. S. M., Markus, H. R., & Phillips, L. T. (2012). A cultural mismatch: Independent cultural norms produce greater increases in cortisol and more negative emotions among first-generation college students. *Journal of Experimental Social Psychology*, *48*, 1389–1393. https://doi.org/10.1016/j.jesp.2012.07.008

Strier, R., Sigad, L., Eisikovits, Z., & Buchbinder, E. (2014). Masculinity, poverty and work: The multiple constructions of work among working poor men. *Journal of Social Policy*, *43*,

331–349. https://doi.org/10.1017/S0047279413000949

Swencionis, J. K., Dupree, C. H., & Fiske, S. T. (2017). Warmth-competence tradeoffs in impression management across race and social-class divides. *Journal of Social Issues*, *73*, 175–191. https://doi.org/10.1111/josi.12210

Tang, T. L.-P., & Gilbert, P. R. (1995). Attitudes toward money as related to intrinsic and extrinsic job satisfaction, stress and work-related attitudes. *Personality and Individual Differences*, *19*, 327–332. https://doi.org/10.1016/0191-8869(95)00057-D

Tracy, J. L., Shariff, A. F., Zhao, W., & Henrich, J. (2013). Cross-cultural evidence that the nonverbal expression of pride is an automatic status signal. *Journal of Experimental Psychology: General*, *142*, 163–180. https://doi.org/10.1037/a0028412

Twenge, J. M. (2001). Changes in women's assertiveness in response to status and roles: A cross-temporal meta-analysis, 1931-1993. *Journal of Personality and Social Psychology*, *81*, 133–145. https://doi.org/10.1037/0022-3514.81.1.133

Twenge, J. M. (2009). Status and gender: The paradox of progress in an age of narcissism. *Sex Roles*, *61*, 338–340. https://doi.org/10.1007/s11199-009-9617-5

Urban, S. (2012). University education as a compensation strategy among second-generation immigrants. *International Migration Review*, *46*, 919–940. https://doi.org/10.1111/imre.12005

Van Vugt, M., De Cremer, D., & Janssen, D. P. (2007). Gender differences in cooperation and competition: The male-warrior hypothesis. *Psychological Science*, *18*, 19–23. https://doi.org/10.1111/j.1467-9280.2007.01842.x

van Vugt, M., & Tybur, J. M. (2015). The evolutionary foundations of hierarchy: Status, dominance, prestige, and leadership. In D. M. Buss (Ed.), *Handbook of Evolutionary Psychology* (2nd ed., Vol. 2, pp. 788–809). Hoboken, NJ: John Wiley & Sons.

Vandello, J. A., & Bosson, J. K. (2013). Hard won and easily lost: A review and synthesis of theory and research on precarious manhood. *Psychology of Men and Masculinity*, *14*, 101–113. https://doi.org/10.1037/a0029826

Vandello, J. A., Bosson, J. K., Cohen, D., Burnaford, R. M., & Weaver, J. R. (2008). Precarious manhood. *Journal of Personality and Social Psychology*, *95*, 1325–1339. https://doi.org/10.1037/a0012453

Varnum, M. E. W. (2013). What are lay theories of social class? *PLoS ONE*, *8*(7), e70589. https://doi.org/10.1371/journal.pone.0070589

Vescio, T. K., Gervais, S. J., Snyder, M., & Hoover, A. (2005). Power and the creation of patronizing environments: The stereotype-based behaviors of the powerful and their effects on female performance in masculine domains. *Journal of Personality and Social Psychology*, *88*, 658–672. https://doi.org/10.1037/0022-3514.88.4.658

Weaver, J. R., Vandello, J. A., & Bosson, J. K. (2013). Intrepid, imprudent, or impetuous?: The effects of gender threats on men's financial decisions. *Psychology of Men and Masculinity*, *14*, 184–191. https://doi.org/10.1037/a0027087

Weaver, J. R., Vandello, J. A., Bosson, J. K., & Burnaford, R. M. (2010). The proof is in the punch: Gender differences in perceptions of action and aggression as components of manhood. *Sex Roles*, *62*, 241–251. https://doi.org/10.1007/s11199-009-9713-6

Williams, M. J., Paluck, E. L., & Spencer-Rodgers, J. (2010). The masculinity of money: Automatic stereotypes predict gender differences in estimated salariees. *Psychology of Women Quarterly*, *34*, 7–20. https://doi.org/10.1111/j.1471-6402.2009.01537.x

Xie, Y., Fang, M., & Shauman, K. (2015). STEM education. *Annual Review of Sociology*, *41*, 331–357. https://doi.org/10.1146/annurev-soc-071312-145659

Zhang, S., Schmader, T., & Forbes, C. (2009). The effects of gender stereotypes on women's career choice: Opening the glass door. In M. Barreto, M. K. Ryan, & M. T. Schmitt (Eds.), *Psychology of women book series. The glass ceiling in the 21st century: Understanding barriers to gender equality* (pp. 125–150). https://doi.org/10.1037/11863-006