

Implications of Research Staff Demographics for Psychological Science

Serena Does
University of California, Los Angeles

Naomi Ellemers
Utrecht University

John F. Dovidio
Yale University

Jasmine B. Norman
University of Utah

Avital Mentovich
University of Haifa

Romy van der Lee
VU University of Amsterdam

Phillip Atiba Goff
John Jay College of Criminal Justice

Long-standing research traditions in psychology have established the fundamental impact of social categories, such as race and gender, on people's perceptions of themselves and others, as well as on general human cognition and behavior. However, there is a general tendency to ignore research staff demographics (e.g., researchers' race and gender) in research development and research reports. Variation in research staff demographics can exert systematic and scientifically informative influences on results from psychological research. Consequently, research staff demographics need to be considered, studied, and/or reported, along with how these demographics were allowed to vary across participants or conditions (e.g., random assignment, matched with participant demographics, or included as a factor in the experimental design). In addition to providing an overview of multidisciplinary evidence of research staff demographics effects, it is discussed how research staff demographics might influence research findings through (a) ingroup versus outgroup effects, (b) stereotype and (implicit) bias effects, and (c) priming and social tuning effects. Finally, an overview of recommended considerations is included (see [Appendix](#)) to help illustrate how to systematically incorporate relevant research staff demographics in psychological science.

Keywords: experimenter effects, intergroup processes, demographics, validity, generalizability

Social categories, such as gender and race, can have important psychological implications. Social categories shape how individuals see themselves and others, and they

influence cognitions and behavior (e.g., Barreto, Spears, Ellemers, & Shahinper, 2003; Ellemers, 2012)—often in ways outside of conscious awareness (Handley, Brown, Moss-Racusin, & Smith, 2015; Scheepers, Saguy, Dovidio, & Gaertner, 2014). There is growing evidence that even under highly controlled conditions and after extensive training of research staff, demographics of research staff (e.g., researcher race or gender) can influence scientific results (Schiebinger et al., 2011–2015). However, there is a general tendency to ignore research staff demographics in research development and reports. Nevertheless, variation in research staff demographics can exert systematic and scientifically informative influences on results of psychological research.

Why Research Staff Demographics Matter

It has been established that participants are highly sensitive to the cues emitted by researchers in a study (Rosenthal & Fode, 1963) and are often motivated to cooperate with a

This article was published Online First March 1, 2018.

Serena Does, UCLA Anderson School of Management, University of California, Los Angeles; Naomi Ellemers, Department of Psychology, Utrecht University; John F. Dovidio, Department of Psychology, Yale University; Jasmine B. Norman, Department of Psychology, University of Utah; Avital Mentovich, School of Criminology, University of Haifa; Romy van der Lee, Department of Organization Sciences, VU University Amsterdam; Phillip Atiba Goff, Department of Psychology and Center for Policing Equity, John Jay College of Criminal Justice.

We gratefully acknowledge the valuable contributions of Kai van Eekelen, Tatev Papikyan, Hans IJzerman, Diana Sanchez, Judi Mesman, Carel ten Cate, Christine Wells, Jon Peck, Joni Ricks-Oddie, and Henry J. Does.

Correspondence concerning this article should be addressed to Serena Does, UCLA Anderson School of Management, Collins Center A4.16, 110 Westwood Plaza, Box 951481, Los Angeles, CA. E-mail: doessrs@ucla.edu

researcher to confirm what they perceive to be the researcher's hypothesis (Orne, 1962). Indeed, there is ample evidence demonstrating that there are "experimenter effects" in psychological research (Kintz, Delprato, Mettee, Persons, & Schappe, 1965; Rosenthal & Fode, 1963). However, even without the deliberate intent of the researchers or knowledge of the hypotheses among research staff members, the demographics of research staff can constitute psychologically meaningful parts of the experimental environment that can systematically alter the results.

This issue has long been recognized in psychology. For example, classic work on the phenomenon of perceptual defense (see Eriksen, 1960) has implicated the potential role of research staff demographics. McGinnies (1949),¹ consistent with his hypothesis that people may be less likely to perceive stimuli that may be threatening in a particular context, found that participants took longer to identify a sexually offensive than a neutral word. Although McGinnies did not systematically vary the gender of the research assistants (see Rumenik, Capasso, & Hendrick, 1977), the fact that the research assistant was a woman may have made sexually offensive words particularly socially threatening in that context. These effects of research staff demographics have been documented across disciplines and across research stages, from participant recruitment to data coding and analysis. As discussed in more detail later, variations in research staff demographics can significantly influence participants' psychological, behavioral, and physiological responses (Cooper et al., 2003; Danso & Esses, 2001; Levav & Argo, 2010; Moorman, Newman, Millikan, Tse, & Sandler, 1999; Schiebinger et al., 2011–2015; Vrana & Rollock, 1998).

Although many of the examples of research staff demographics effects outlined in this contribution come from research specifically designed to test such effects, there are also studies that serendipitously revealed systematic effects associated with research staff demographics. Moreover, even seemingly unrelated research staff characteristics can unintentionally affect results of studies with identical designs. For example, there is work showing that expertise of a given research team predicts replication success (Bench, Rivera, Schlegel, Hicks, & Lench, 2017) and that the height of an experimenter influences people's estimates of a stimulus person's height (Rump & Delin, 1973). The observation by Rump and Delin (1973) that experimenter height influenced participants' perceptions was unexpected and not part of their original hypothesis. Another example comes from a study on animal behavior, in which laboratory personnel anecdotally reported changes in animal pain behavior as a function of the presence of researchers (Sorge et al., 2014). Although it was not their research focus initially, they discovered that male, but not female, researchers elicited physiological and behavioral stress responses in mice through their scent. The authors concluded that these findings "strongly suggest that standard laboratory practice

should account for experimenter sex when investigating any phenomenon possibly affected by stress" (Sorge et al., 2014, p. 632). Thus, in some cases, effects of research staff demographics have been observed even when they were not the original focus of the research.

The current work considers the potential systematic impact of research staff demographics in psychological research. Omitting research staff demographic information from research reports or articles can lead to misinterpretations of results, failed replication efforts, and potentially misleading claims regarding validity and generalizability. It also precludes meta-analysis of the impact of research staff demographics. More explicit consideration and reporting of the demographics of research staff can help bring new insights into basic psychological processes, illuminate why it is that some findings seem to exist under certain conditions and not others (Van Bavel, Mende-Siedlecki, Brady, & Reinero, 2016a, 2016b), and stimulate new research on the ways diversity in research environments affects human behavior and cognition. The purpose of the current work is to expand existing calls to consider sex, gender, and race-ethnicity of participants to also consider the demographics of research staff—such as experimenters, confederates, interviewers, participant recruiters, and data coders.

Although research staff demographics are important factors to consider, these effects do not imply that they are inherently more influential than other contextual factors (Wang, 2016). The importance of contextual factors, including research staff demographics, will depend on the particulars of the research in question. Recently, research by Van Bavel et al. (2016a) illustrated how contextual sensitivity—the extent to which findings were deemed sensitive to the contextual influences of time, culture, location, and population—can significantly predict replication success. Van Bavel and colleagues (2016a) concluded that "the lesson here is not that context is too hard to study but rather that context is too important to ignore" (p. 6458).

Although researchers focused on addressing a specific hypothesis may not be sensitive to the contextual factors in their empirical investigation, these factors can still function as moderators. For example, research that includes physiological measures such as cortisol would benefit from considering and reporting the time of day that data were collected, because time of day is a known moderator in cortisol research (for a meta-analysis see Dickerson & Kemeny, 2004). Similarly, research on economic attitudes in the United States might consider and report time, in terms of pre- and postrecession, because economic climate can moderate economic attitudes as well as the dynamics underlying these attitudes (Van Bavel et al., 2016a). Whereas some

¹ We thank an anonymous reviewer for pointing this work out to us.

contextual factors such as time of day will be relevant in some studies (e.g., research involving certain physiological measurements), long-standing research traditions in the area of intergroup relations suggests that research staff demographics can influence a broad range of studies that are common in psychological science. As such, research staff demographics merit a special focus when it comes to designing and reporting research.

Which Research Staff Demographics Are Important to Consider and Report?

Specific research staff characteristics, such as physical attractiveness or socioeconomic status, can have systematic effects, depending on the research populations and phenomena under investigation. However, the current work highlights examples based on race, ethnicity, and gender, which have particularly well documented, potent psychological effects on individuals' behaviors, cognitions, and perceptions. This large body of research on race and gender offers a solid theoretical framework and a promising starting point for an analysis of research staff effects. To illustrate how researchers, journal editors, and reviewers can identify, and systematically incorporate, those research staff demographics that are most relevant in their own work, an overview of potential considerations is included in the [Appendix](#).

Research Staff Demographics Effects

The study of research staff effects requires more than simply documenting the race, ethnicity, and gender of research staff and reporting their representation across conditions; it also involves recording this information for each participant (or other given unit level in a study) and directly analyzing the influence of research staff demographic variables on the outcomes of the study and the processes through which these effects occurred. Even when such effects are not statistically significant and/or are not reported in the main text, it would be valuable to present these results in supplementary materials or retain a record of the effects to be made available to researchers for secondary analyses or meta-analyses. Based on the intergroup relations literature, it becomes apparent that research staff demographics can systematically shape research findings through a variety of psychological processes, including (a) ingroup versus outgroup effects, (b) stereotype and (implicit) bias effects, and (c) priming and social tuning effects. Each of these processes are discussed in more detail in the next sections. These group-level effects can occur independently and simultaneously and can all influence the dynamics between research staff and participants, and thereby research findings.

Ingroup Versus Outgroup

Ingroup refers to an individual's own social group along the lines of a specific social category (e.g., for a Black person, Blacks constitute the ingroup). In contrast, *outgroup* refers to groups other than the individual's own social group (e.g., for a Black person, Asians and Whites constitute outgroups). There is a large body of literature documenting that people's cognitions and behaviors toward ingroup members are generally more positive than toward outgroup members. This phenomenon is described in the literature as ingroup favoritism and/or outgroup derogation (e.g., [Turner, Brown, & Tajfel, 1979](#)). However, beyond general positivity, the process of social categorization and the consequent recognition of another person as an ingroup or an outgroup member shape social perception, affect, cognition, and behavior—each of which can affect the responses of research participants.

For instance, cognitively, people retain more information in a more detailed fashion for ingroup members than for outgroup members ([Park & Rothbart, 1982](#)). Behaviorally, people are more trusting of ([Platow, Foddy, Yamagishi, Lim, & Chow, 2012](#)), and helpful toward, ingroup than toward outgroup members ([Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997](#)); work harder for groups identified as ingroups than as outgroups ([Worchel, Rothgerber, Day, Hart, & Butemeyer, 1998](#)); and show greater readiness to approach ingroup members and avoid outgroup members ([Paladino & Castelli, 2008](#)).

In a related vein, compared to an *intragroup* interaction, the perceived demands of an *intergroup* interaction may be higher, which often makes the latter more anxiety-provoking, threatening, and cognitively depleting than the former ([Mendes, Blascovich, Lickel, & Hunter, 2002](#); [Trawalter, Richeson, & Shelton, 2009](#)). For example, Whites interacting with a Black (vs. White) confederate exhibited cardiovascular reactivity associated with psychological threat ([Mendes et al., 2002](#)). Thus, research staff demographics can affect processes related to psychophysiological arousal, threat, cognitive functioning, and cognitive depletion.

In- versus outgroup effects can also arise whenever research findings are dependent on a researcher's perceptions, interpretation, coding, or other type of assessment. Indeed, recent insights have suggested that research staff demographics can shape a staff's perceptions in profound ways ([Ellemers, 2016](#); [Xiao, Coppin, & Van Bavel, 2016](#)). Namely, people are better at recognizing and processing ingroup faces compared to outgroup faces (e.g., [Kawakami et al., 2014](#); [Rhodes, Hayward, & Winkler, 2006](#)) and more accurately perceive emotions that are expressed by ingroup compared to outgroup members ([Elfenbein & Ambady, 2002](#)). Thus, in studies involving raters who code affect by watching videos of participants' facial expressions, for example, there can be greater inaccuracies in an intergroup compared to an intragroup context.

Stereotypes and (Implicit) Bias

Research staff demographics can exert influence on research findings through the effect of stereotypes on the responses of members belonging to stereotyped groups as well as on others' orientation toward members of stereotyped groups. *Stereotype threat*, which refers to the concern that one's actions will be viewed through the lens of a negative stereotype about one's group (Steele & Aronson, 1995), systematically affects members of social groups in contexts that make their group identities salient. For example, Blacks are stereotyped as being less intelligent than Whites (Aronson, Fried, & Good, 2002). Consequently, Black students taking an intelligence test might become concerned that their test performance is going to confirm this negative stereotype about their racial group. Having such concerns—that is, experiencing stereotype threat—can deplete the cognitive resources students need to perform well on the intelligence test. The experience of stereotype threat can thus cause Black students to underperform on the test (Taylor & Walton, 2011).

Research staff demographics can significantly moderate the experience of stereotype threat and thus influence the behavior of participants in systematic ways. For example, Black students showed performance decrements associated with stereotype threat only in the presence of a White test administrator but not a Black test administrator (Marx & Goff, 2005). Different contexts can also activate different social categories—and the stereotypes related to those social categories—in the same people, thereby influencing research findings. For example, Asians are stereotyped as having superior quantitative skills compared to other racial groups, and women are stereotyped as having inferior quantitative skills compared to men (Shih, Pittinsky, & Ambady, 1999). From an intersectional perspective, it was found that Asian women performed worse on a math test when their gender identity was subtly activated but performed better when their ethnic identity was subtly activated, compared to Asian women who had neither identity activated (Shih et al., 1999). These findings show how different stereotype effects can arise as a function of environmental cues. Because research staff demographics constitute cues that can activate different social categories in participants, and for facilitating the interpretation of research findings for a scientific audience, it would be valuable for researchers to consider and report research staff demographics that may be most appropriate and relevant when studying a given population and/or phenomenon.

Furthermore, research staff demographics can influence results through (implicit) bias processes among individuals interacting with members of stereotyped and/or stigmatized groups. Explicit biases are conscious and often expressed openly; implicit biases are automatically activated negative responses, which often occur without conscious awareness

or intention (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Much research has shown that there are negative explicit and implicit biases against certain groups, such as racial minorities (Pieterse, Todd, Neville, & Carter, 2012), women (Borrell et al., 2010), gay people (Dasgupta, & Rivera, 2006; Meyer, 2003), overweight people (e.g., Sabin, Marini, & Nosek, 2012), and low-status countries (Does & Mentovich, 2016). Moreover, the negative impact of perceived racism, sexism, and other forms of prejudice on the mental and physical well-being of those targeted has been well documented (e.g., Borrell et al., 2010; Meyer, 2003; Pieterse et al., 2012). For example, women and Blacks are less likely to receive certain referrals from physicians than are men and Whites, respectively (Schulman et al., 1999). To the extent that people behave less positively (Dasgupta & Rivera, 2006) and feel more negative emotions (Miller, Smith, & Mackie, 2004) toward groups for whom they harbor prejudice, participants may behave differently depending on a research staff member's race and/or gender and vice versa.

Furthermore, implicit biases can affect how participants respond to research staff and to information that is critical to a specific study. People who are more implicitly biased toward a given group respond less favorably and are less responsive nonverbally to members of that group (Dovidio, Kawakami, & Gaertner, 2002). Implicit bias effects are often as strong as, and occur over and above the influence of, explicit bias effects (Greenwald et al., 2009). Moreover, even when participants do not harbor implicit biases themselves, subtle negative emotional expressions by other participants based on a research staff member's race or gender can trigger, unconsciously and unintentionally, negative responses to the staff member (Weisbuch, Pauker, & Ambady, 2009). Hence, negative explicit and implicit biases toward research staff can have cascading, but often overlooked, influences on the dynamics of the behavior being studied.

Data concerning how these forms of bias among research staff might influence research findings are currently limited. Nevertheless, the evidence that does exist has suggested that biases associated with research staff demographics can shape research findings in important ways. For example, White Americans' implicit (but not explicit) racial prejudice is associated with a greater readiness to perceive anger—that is, threatening affect—in Black but not White faces (Hugenberg & Bodenhausen, 2003). Subsequent work replicated this finding and showed it to be a more general intergroup process, in which angry faces are more likely to be categorized as outgroup faces compared to ingroup faces (Dunham, 2011). Moreover, there is work showing that White coders significantly rate Black families as having poorer family management skills than do White families, even after coders received training. This discrepancy was not found among Black coders (Yasui & Dishion, 2008). The latter finding hints at the hypothesis that, in some

cases, racial minority research staff might be less prone to exhibit explicit bias against racial minority groups than are White research staff. In a related vein, it has been argued that psychotherapists' unexamined classist assumptions pose barriers for inclusion and effective treatment of poor clients (Smith, 2005). Considering prejudices against certain populations and how these forms of prejudice might influence research staff's treatment, evaluation, or inclusion of these populations in their research can thus inform the interpretation of research findings and their implications for theory and application.

Priming and Social Tuning Effects

Mere exposure to a person whose social category is salient can also activate (i.e., prime) thoughts and behaviors associated with that social category. Research staff demographics—which often constitute immediately visible and highly salient elements of the social context—can represent one such prime. As such, priming effects can influence participants' responses to research staff and vice versa, thereby influencing research findings. For example, the presence of a Black confederate might prime a White participant's stereotypes about Black people, thereby potentially shaping the participant's responses and/or behavior during the research. Priming effects have been widely documented. For example, participants who were covertly primed with stereotypes of Blacks (which includes the trait *hostility*) were more likely than nonprimed participants to express hostility after provocation. In the same research, participants who were covertly primed with stereotypes of the elderly (and thus the trait *slowness*) were more likely than nonprimed participants to later walk slowly down a hallway (Bargh, Chen, & Burrows, 1996).

In addition, research staff demographics can influence results through social tuning—which refers to the process of people emulating others' behaviors or attitudes due to affiliative motives. Even relatively subtle cues can cause these effects. For instance, greater liking of a researcher predicted lower anti-Black or pro-White automatic prejudice when the researcher was wearing an antiracism shirt compared to a blank shirt (Sinclair, Lowery, Hardin, & Colangelo, 2005). Similarly, heterosexual participants exhibited less implicit and explicit bias toward lesbians and gay men when the researcher wore an antidiscrimination shirt compared to a blank shirt (Skorinko et al., 2015). These findings suggest that affiliative motives can shape participants' behaviors and (implicit) attitudes as a function of research staff characteristics.

Existing Guidelines: Demographics of Research Participants

Reporting guidelines do exist for the demographics of research participants in contrast to research staff demo-

graphics. The American Psychological Association (APA) has explicit guidelines for the reporting of participant demographics, including, race and gender, in its publication manual (American Psychological Association, 2010). Moreover, policies for scientific research, such as the National Institutes of Health's (NIH, 1993) Health Revitalization Act, have been established to promote inclusion of participants across gender and racial and ethnic backgrounds. A growing number of peer-reviewed journals have implemented editorial guidelines requiring the reporting of information regarding sex, gender, and race-ethnicity of research participants (Schiebinger et al., 2011–2015). In addition, for any NIH-funded pre-clinical research, researchers are required to account for sex as a biological variable (Clayton, 2016). Such guidelines and policies are aimed at promoting research practices that are informed by sex, gender, and racial/ethnic differences—thereby increasing validity and generalizability of scientific findings

Many of the research staff demographics effects discussed in the current work are expected to arise as a function of the interaction between research staff and participants' demographics. Thus, it is only when participant demographics as well as research staff demographics are reported that such effects can be assessed—and ultimately, controlled or systematically varied, depending on the objectives of the research project. However, although there is increasing awareness of the importance of participant demographics in terms of reporting practices (e.g., Schiebinger et al., 2011–2015) as well as in terms of research development (Henrich, Heine, & Norenzayan, 2010), there has been almost no attention focused on research staff demographics. In contrast to the guidelines for participants' demographics outlined previously, no guidelines or policies currently exist regarding reporting of research staff demographics. The next section discusses an assessment of current reporting practices regarding research staff demographics in psychological science.

Current Reporting Practices of Research Staff Demographics

The Open Science Collaboration (OSC, 2015) is a large-scale collaboration of hundreds of researchers who attempted to replicate 99 original experiments reported in articles that were published in three high-ranking psychology journals in 2008. The OSC research replicated about one third to one half of the original effects. Since then, it has been demonstrated that contextual sensitivity—that is, the extent to which findings were deemed sensitive to the contextual influences of time, culture, location, and population—significantly predicted replication success of the studies that were selected by the OSC (Van Bavel et al., 2016a; see also Inbar, 2016; Van Bavel et al., 2016b). To shed light on reporting practices regarding research staff demographics, and following procedures similar to those in Van Bavel and colleagues (2016a), we coded how many of

the original studies examined by the OSC reported demographics of participants and research staff. Specifically, for each of the OSC's 99 original studies, a research intern² coded whether the research report included mention of (a) the study location (i.e., laboratory, field, or online), (b) demographics of participants, (c) involvement of researchers and other research staff, and (d) demographics of researchers and other research staff (see [Table 1](#)).

Surprisingly, given that there are existing guidelines for reporting participant demographics—including those of the APA—it was observed that only 46% of the 99 studies reported participant gender and a mere 13% of 99 studies reported participant race or ethnicity. For 16 of the 99 original studies of the [OSC \(2015\)](#), it was not reported, nor could it be inferred from the text, whether they were conducted online or in a laboratory or field setting, thus providing insufficient information to determine whether, how, and what type of research staff were involved in data collection. Of the remaining 83 original studies, 75 studies were explicitly reported as having been conducted either in the laboratory or in the field, thus likely requiring the presence of at least one researcher. However, of these 75 studies, only 9% reported research staff gender and only 5% reported research staff race–ethnicity. These findings suggest that demographics of research staff are indeed commonly ignored in research reports.

Although the limited sample size restricts the ability to draw any firm conclusions, it is worth noting that of the 10 original studies for which the research topic was directly related to race and/or gender, only six studies reported participants' race, eight studies reported participants' gender, and only three studies reported research staffs' race and/or gender. Thus, even for those studies that were specifically about race and/or gender—in which salience of social categories was heightened even more ([Oakes, 1987](#))—demographics of participants and research staff were not always reported. This makes it impossible to assess whether and how in- versus outgroup dynamics may have influenced results. Furthermore, a closer look at the eight original studies that did report research staff gender and/or race (see [Table 1](#)) revealed this information was not systematically considered in replication efforts. In theory, future research might be able to empirically consider the effects of research staff demographics on the process and outcomes of research by contacting the authors of the 99 OSC studies for demographic information about the research staff and testing meta-analytically whether studies with a more diverse research staff produce more replicable results.³ Original experiments that systematically vary research staff demographics would also be particularly valuable for directly testing the potential causal role of research staff demographics on research findings and factors (e.g., research topic, subdiscipline) that could moderate these effects. In addition, an experimental approach limits the

effects of potential confounding influences that would be difficult to isolate in a meta-analysis in which study characteristics vary substantially.

Implications of Research Staff Demographics

There are different ways in which demographics of research staff can constitute active—rather than passive or neutral—elements of the research process. What follows is a point-by-point discussion of potential implications of effects of research staff demographics for research and policy.

Recruitment

Psychology is often faced with the issue of low participation rates, particularly of racial and ethnic minorities ([Murthy, Krumholz, & Gross, 2004](#)). Demographics of recruiters (i.e., race and gender) can pose barriers to individuals' decision to participate in research ([Moorman et al., 1999](#); [Senn & Desmarais, 2001](#)). For example, women were more likely to participate in a breast cancer study when recruited by a same-race interviewer (i.e., an ingroup member; [Moorman et al., 1999](#)). In addition, men and women were more likely to sign up for a study on sexual experiences when recruited by a man compared to a woman ([Senn & Desmarais, 2001](#)). More generally, it could be the case that lower participation rates of racial or ethnic minorities ([Murthy et al., 2004](#)) are because these individuals are far more likely to be approached by a recruiter from a racial outgroup (i.e., White recruiters) than White individuals are.

Recruiter demographics also play an important role in recruitment at the institutional level. Namely, primary care practices with racial minority physicians were more likely to be successfully recruited for an intervention study when there was greater commonality between recruiters and key practice members ([Hudson, Harris-Haywood, Stange, Orzano, & Crabtree, 2006](#)). Thus, the demographics of recruiters can have important implications for enrollment of participants as well as institutions in scientific studies—and future work should take this into consideration.

Data Collection

Most laboratory and field studies entail some form of contact between participants and research staff, for example, for recruitment, reception, or instruction purposes. Even in the case of online studies—such as studies conducted through Amazon Mechanical Turk—the consent form usually includes at least the name of the principal investigator, from which participants can infer the gender and in some cases the race or ethnicity. Whenever there is direct

² The research intern was a White man and psychology major who was unaware of the research aims.

³ We thank an anonymous reviewer for this suggestion.

Table 1
Reporting Frequency of Participant and Research Staff Demographics in 99 Original Experiments Replicated by the Open Science Collaboration (2015)

Variable	Participants		Experimenters		Other research staff ^a	
	Yes	No	Yes	No	Yes	No
Involvement reported	99	0	21	78	6	93
Total <i>N</i> reported	99	0	4	95	3	96
Gender reported ^b	46	53	6	93	1	98
Race or ethnicity reported	13 ^c	86	3	96	1	98

^a Refers to recruiters, coders, and confederates. ^b When gender of experimenters or other research staff could be inferred from the text based on gender pronouns, this was counted as having reported gender. ^c In seven studies, authors solely reported the (native) language of participants, which was counted as not having reported race or ethnicity.

contact with participants and/or whenever research staff demographics are known, research staff demographics can influence participants' psychological, behavioral, and/or physiological responses (Cooper et al., 2003; Danso & Esses, 2001; Levav & Argo, 2010; Moorman et al., 1999; Senn & Desmarais, 2001; Vrana & Rollock, 1998). Such effects can occur even when contact with participants is brief. For example, White participants exhibited increased heart rate levels in the presence of a Black compared to a White same-gender researcher who took their pulse for 30 seconds (Vrana & Rollock, 1998).

The majority of studies that have examined the effects of research staff demographics have focused on the data collection stage of the research process. Results from such work have generally shown that variations in research staff demographics can have far-reaching implications for the phenomena under investigation (however, for an exception see Rotundo & Sackett, 1999). Effects of research staff demographics have been observed in terms of, but not limited to, participants' academic test performance (Danso & Esses, 2001; Marx & Goff, 2005), heart rate (Larkin, Ciano-Federoff, & Hammel, 1998; Vrana & Rollock, 1998), anxiety level (Lamarche, Gammage, & Gabriel, 2011), financial risk-taking (Levav & Argo, 2010), pain reporting (Weisse, Foster, & Fisher, 2005), and autonomic and subjective responses to pain stimuli (Aslaksen, Myrbakk, Høifødt, & Flaten, 2007). For example, men with higher social fear levels exhibited higher heart rates to a math task than did men with lower social fear levels, but only if the present confederate was a woman, not a man (Larkin et al., 1998). As mentioned earlier, research has shown how patients can benefit from race concordance with their physician (Cooper et al., 2003; Cooper-Patrick et al., 1999) as well as with their therapist (Thompson & Alexander, 2006). Similarly, Chinese and Euro-Canadian participants reported more affective pain and increased nonverbal pain displays when the researcher was from an ethnic ingroup compared to an ethnic outgroup (Hsieh, Tripp, & Ji, 2011).

The overall pattern of findings offers substantial evidence for why research staff demographics need to be

considered and, at the very least, reported in research involving humans. As can be seen from these examples, the potential impact of research staff demographics on research findings is not restricted to those research areas that specifically study processes related to race and gender but rather stretches across a wide spectrum of research topics and areas. A greater understanding of the phenomena under consideration might be achieved when the ways in which research staff demographics might interact with participants' demographics are considered. Reporting of relevant research staff demographics will facilitate post hoc interpretations or meta-analysis of data examining the potential effects of those demographics. While being mindful of the fact that resource constraints often dictate the breadth of research designs, researchers can consider these options: (a) holding constant or randomizing the demographic variation across conditions and participants; (b) controlling for research staff demographics in statistical analysis; or, most rigorously, (c) including research staff demographics in the research design (Danso & Esses, 2001; Vrana & Rollock, 1998). Studies in which research staff demographics are systematically varied might also consider specific theoretically guided moderating influences, such as the relevance of research staff characteristics (e.g., weight) to the topic being studied (e.g., weight stigma) or the type of response examined.

It is worth noting that the value and effectiveness of randomization will partly depend on the diversity of research staff—for example, if eight out of 10 researchers are White women, randomization will not be an appropriate approach to account for researcher race or gender. Moreover, examining, rather than controlling for, effects of research staff demographics and their interactions with other factors—such as other contextual factors or participants' demographics—will expand knowledge of psychological mechanisms, inform boundary conditions, help build theory, and improve the effectiveness of applications. The Appendix presents an overview of considerations to help illustrate how research staff demographics can be system-

atically considered and incorporated in research. These considerations are not solely for those who study intergroup processes but rather for all those whose research involves participants. As noted in the rationale for giving explicit attention to research staff demographics, there have also been studies in other areas (e.g., perception of height, animal stress responses) that have serendipitously revealed systematic effects associated with research staff demographics. Thus, effects of research staff demographics can arise even when they are not the original focus of the research. The list of considerations in the [Appendix](#) is meant to be illustrative rather than exhaustive and may be tailored to address particular topic-related considerations.

Data Coding and Analysis

When research staff are not visible to, or do not come into contact with, participants, research staff demographics can still influence research findings. Demographics of “nonvisible” research staff (such as coders and observers) can shape staff’s observations, interpretations, and/or coding of data, thereby impacting research results. As mentioned earlier, there is evidence of ethnocentric bias in White coders’ ratings of Blacks’ family management skills (Yasui & Dishion, 2008). In addition, White Australian and Chinese participants were more accurate at detecting experimentally manipulated changes in own-race than in other-race faces (Rhodes et al., 2006). Recent research established that important intergroup biases, such as face recognition of ingroup versus outgroup members, are partly driven by a preferential attention to the eyes of racial ingroup members (Kawakami et al., 2014). Moreover, when coding interviewees’ emotions, men have been found to focus more on aggressiveness (anger) and women more on “deeper” feelings (despair, worry; Abrilian, Devillers, & Martin, 2006). Gender differences have also been documented when it comes to observing pain in others. Namely, women are perceived as having more pain than men have, and men underestimate pain in others to a greater degree than women do (Robinson & Wise, 2003). Together, these findings suggest that even if coders and observers are unaware of the research aims, conditions, or hypotheses, there may still be systematic observer effects due to research staff demographics. Similar strategies as those outlined for the data collection stage can be used to account for research staff demographics in observations and rating or coding of data.

Policy Implications

There are substantial benefits when policy is informed by science that considers the role of staff demographics. To illustrate such effects, this section outlines evidence from large cross-sectional and field studies showing the importance of research staff demographics in the context of de-

veloping policy and guidelines. In the medical context, nationally representative data of hospitalized Medicare beneficiaries were analyzed, and it was found that in the same hospital, patients who were treated by physicians who were men had significantly higher mortality rates and readmission rates compared to patients who were treated by physicians who were women (Tsugawa et al., 2017). Clearly, uncovering physician gender effects on critical patient outcomes has major implications for the development of guidelines, practices, and policy regarding medical practices. In the context of education, several field studies have found that incorporating staff demographics can help combat racial and gender achievement gaps. For example, a 1-year assignment to an own-race teacher significantly increased math and reading performance of Black and White students by 3–4 percentile points (Dee, 2004). Another field study showed that the academic gender gap between men and women in STEM fields (i.e., science, technology, engineering, and mathematics) is eradicated when women are taught by professors who are women rather than men (Carrell, Page, & West, 2010). In light of the Black–White gap in vocabulary test scores in the General Social Survey (GSS), Black respondents perform better on vocabulary testing when tested by a Black rather than a White interviewer. Because most Black respondents are tested by White interviewers in the GSS, these findings suggest that the racial gap in vocabulary performance is overestimated (Huang, 2009). Taken together, findings from these field studies illustrate the “real world” importance of reporting and incorporating research staff demographics in psychological science to accurately, and more effectively, inform policy. The research practices just outlined can help promote more accurate and effective translations of research findings to interventions and policies.

General Discussion

Gender and racial or ethnic differences are not new phenomena in science, but currently these differences are primarily considered through one of two lenses. The first lens focuses on the gender and race or ethnicity of those being studied (i.e., the research population). As previously mentioned, a growing number of journals currently have guidelines requiring authors to report participants’ sex, gender, and race–ethnicity (Schiebinger et al., 2011–2015), and the NIH requires researchers to account for participants’ sex as a biological variable (Clayton, 2016). Furthermore, recent work has shown that findings based on populations from WEIRD societies (i.e., Western, Educated, Industrialized, Rich, and Democratic) are very limited in their generalizability to the larger population internationally (Henrich et al., 2010). Such research has further highlighted the importance of systematically considering, reporting, and diversifying participants’ demographics. Yet, as was revealed by

our coding of the reporting of demographics in the studies from the OSC (2015) project, these demographics are currently not being taken into account systematically.

The focus of the second lens is on those doing the research (i.e., researchers). From this perspective, research staff demographics are considered in terms of (addressing) the underrepresentation of women and racial-ethnic minorities, or the overrepresentation of men and Whites, in the scientific community (Moss-Racusin et al., 2014). In a related vein, a study on NIH research awards found that, after controlling for merit-based predictors such as educational background and publication records, Black applicants had lower award probability than did White applicants (Ginther et al., 2011). Similarly, an examination of the Dutch equivalent of the National Science Foundation showed that women had a lower probability of receiving research funding compared to men with equal ratings of application quality (Van der Lee & Ellemers, 2015). These studies add to a growing body of evidence suggesting that principal investigators' race and gender influence their opportunities and outcomes in academia.

Notwithstanding the importance of both of these lenses, the present work proposes a third lens: Namely, a focus on the interplay between those being studied and those who are conducting the research. Although generally ignored in research reports (see Table 1), demographic variation among research staff can exert systematic—and scientifically informative—influences on participants and phenomena under investigation. Specifically, research staff demographics can influence findings—through (a) in- versus outgroup effects, (b) stereotype and (implicit) bias effects, and (c) priming and social tuning effects—illustrating the importance of incorporating research staff demographics across research stages. Certainly, the effects of research staff demographics may be one of many possible influences in a study and may not appear directly relevant to the focal hypothesis. However, by reporting the demographics of research staff and testing for their potential effects, researchers can offer their scientific audience a more comprehensive understanding of the elements shaping the results and their interpretation. The Appendix includes practical considerations for researchers, journal editors, and reviewers to help guide efforts to consider and incorporate research staff demographics in psychological science. This information about staff demographics and their effects can be made available to a journal's audience in a variety of ways, such as in online supplementary materials, as determined by journal editors. Regardless of how this information is made accessible, its availability may make both authors and readers more aware of the potential systematic influence of these factors and may stimulate new lines of research to understand both the extent of such effects and the mechanisms that may underlie them.

The systematic uncovering of research staff demographics effects starts with recording research staff demographics at the unit record level (i.e., for each observation) and a consistent documentation of participant and research staff demographics in research reports, as well as the method in which research staff demographics were allowed to vary across participants or conditions. In addition, where possible, it is valuable to statistically analyze the effects of research staff demographics on the phenomena under investigation and make this information available to readers and for secondary data analysis by other researchers. Doing so may provide new insights and draw attention—conceptually and empirically—to the potentially impactful, but often overlooked, influence of research staff demographics on scientific conclusions.

Determining the magnitude and prevalence of research staff demographics effects in psychological research requires research specifically designed to test for these effects and meta-analytical research of studies that included, or at the very least reported, research staff demographics. Such practices will help to advance validity, replicability, applicability, and generalizability of research findings in psychological science.

References

- Abrilian, S., Devillers, L., & Martin, J.-C. (2006). Annotation of emotions in real-life video interviews: Variability between coders. In *5th international conference on language resources and evaluation* (pp. 2004–2009). Available at <http://www.lrec-conf.org/proceedings/lrec2006/>
- American Psychological Association. (2010). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Aronson, J., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*, 113–125. <http://dx.doi.org/10.1006/jesp.2001.1491>
- Aslaksen, P. M., Myrbakk, I. N., Høifødt, R. S., & Flaten, M. A. (2007). The effect of experimenter gender on autonomic and subjective responses to pain stimuli. *Pain, 129*, 260–268. <http://dx.doi.org/10.1016/j.pain.2006.10.011>
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology, 71*, 230–244. <http://dx.doi.org/10.1037/0022-3514.71.2.230>
- Barreto, M., Spears, R., Ellemers, N., & Shahinper, K. (2003). Who wants to know? The effect of audience on identity expression among minority group members. *British Journal of Social Psychology, 42*, 299–318. <http://dx.doi.org/10.1348/014466603322127265>
- Bench, S. W., Rivera, G. N., Schlegel, R. J., Hicks, J. A., & Lench, H. C. (2017). Does expertise matter in replication? An examination of the reproducibility project: Psychology. *Journal of Experimental Social Psychology, 68*, 181–184. <http://dx.doi.org/10.1016/j.jesp.2016.07.003>
- Borrell, C., Artazcoz, L., Gil-González, D., Pérez, G., Rohlf, I., & Pérez, K. (2010). Perceived sexism as a health determinant in Spain. *Journal of Women's Health, 19*, 741–750. <http://dx.doi.org/10.1089/jwh.2009.1594>
- Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *Quarterly Journal of Economics, 125*, 1101–1144. <http://dx.doi.org/10.1162/qjec.2010.125.3.1101>

- Clayton, J. A. (2016). Studying both sexes: A guiding principle for biomedicine. *Federation of American Societies for Experimental Biology Journal*, 30, 519–524. <http://dx.doi.org/10.1096/fj.15-279554>
- Cooper, L. A., Roter, D. L., Johnson, R. L., Ford, D. E., Steinwachs, D. M., & Powe, N. R. (2003). Patient-centered communication, ratings of care, and concordance of patient and physician race. *Annals of Internal Medicine*, 139, 907–915. <http://dx.doi.org/10.7326/0003-4819-139-11-200312020-00009>
- Cooper-Patrick, L., Gallo, J. J., Gonzales, J. J., Vu, H. T., Powe, N. R., Nelson, C., & Ford, D. E. (1999). Race, gender, and partnership in the patient-physician relationship. *Journal of the American Medical Association*, 282, 583–589. <http://dx.doi.org/10.1001/jama.282.6.583>
- Danso, H. A., & Esses, V. M. (2001). Black experimenters and the intellectual test performance of White participants: The tables are turned. *Journal of Experimental Social Psychology*, 37, 158–165. <http://dx.doi.org/10.1006/jesp.2000.1444>
- Dasgupta, N., & Rivera, L. M. (2006). From automatic antigay prejudice to behavior: The moderating role of conscious beliefs about gender and behavioral control. *Journal of Personality and Social Psychology*, 91, 268–280. <http://dx.doi.org/10.1037/0022-3514.91.2.268>
- Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *The Review of Economics and Statistics*, 86, 195–210.
- Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, 130, 355–391. <http://dx.doi.org/10.1037/0033-2909.130.3.355>
- Does, S., & Mentovich, A. (2016). Rooting for the top dog: How social dominance motives shape group preference in intergroup competition. *Journal of Experimental Social Psychology*, 62, 24–29. <http://dx.doi.org/10.1016/j.jesp.2015.09.002>
- Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82, 62–68. <http://dx.doi.org/10.1037/0022-3514.82.1.62>
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, 33, 510–540. <http://dx.doi.org/10.1006/jesp.1997.1331>
- Dunham, Y. (2011). An angry = outgroup effect. *Journal of Experimental Social Psychology*, 47, 668–671. <http://dx.doi.org/10.1016/j.jesp.2011.01.003>
- Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin*, 128, 203–235. <http://dx.doi.org/10.1037/0033-2909.128.2.203>
- Ellemers, N. (2012, May 18). The group self. *Science*, 336, 848–852. <http://dx.doi.org/10.1126/science.1220987>
- Ellemers, N. (2016). Back to the future: How today's neurocognitive techniques substantiate predictions made 50 years ago. *Psychological Inquiry*, 27, 290–293. <http://dx.doi.org/10.1080/1047840X.2016.1215210>
- Eriksen, C. W. (1960). Discrimination and learning without awareness: A methodological survey and evaluation. *Psychological Review*, 67, 279–300. <http://dx.doi.org/10.1037/h0041622>
- Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L., & Kington, R. (2011). Race, ethnicity, and NIH research awards. *Science*, 333, 1015–1019. <http://dx.doi.org/10.1126/science.1196783>
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the implicit association test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97, 17–41. <http://dx.doi.org/10.1037/a0015575>
- Handley, I. M., Brown, E. R., Moss-Racusin, C. A., & Smith, J. L. (2015). Quality of evidence revealing subtle gender biases in science is in the eye of the beholder. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 13201–13206. <http://dx.doi.org/10.1073/pnas.1510649112>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <http://dx.doi.org/10.1017/S0140525X0999152X>
- Hsieh, A. Y., Tripp, D. A., & Ji, L.-J. (2011). The influence of ethnic concordance and discordance on verbal reports and nonverbal behaviours of pain. *Pain*, 152, 2016–2022. <http://dx.doi.org/10.1016/j.pain.2011.04.023>
- Huang, M.-H. (2009). Race of the interviewer and the Black-White test score gap. *Social Science Research*, 38, 29–38. <http://dx.doi.org/10.1016/j.ssresearch.2008.07.004>
- Hudson, S. V., Harris-Haywood, S., Stange, K. C., Orzano, A. J., & Crabtree, B. F. (2006). Recruiting minority primary care practices into practice-based research. *Medical Care*, 44, 696–700. <http://dx.doi.org/10.1097/01.mlr.0000215818.38729.81>
- Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, 14, 640–643. http://dx.doi.org/10.1046/j.0956-7976.2003.psci_1478.x
- Inbar, Y. (2016). Association between contextual dependence and replicability in psychology may be spurious. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E4933–E4934. <http://dx.doi.org/10.1073/pnas.1608676113>
- Kawakami, K., Williams, A., Sidhu, D., Choma, B. L., Rodriguez-Bailón, R., Cañadas, E., . . . Hugenberg, K. (2014). An eye for the I: Preferential attention to the eyes of ingroup members. *Journal of Personality and Social Psychology*, 107, 1–20. <http://dx.doi.org/10.1037/a0036838>
- Kintz, B. L., Delprato, D. J., Mettee, D. R., Persons, C. E., & Schappe, R. H. (1965). The experimenter effect. *Psychological Bulletin*, 63, 223–232. <http://dx.doi.org/10.1037/h0021718>
- Lamarche, L., Gammage, K. L., & Gabriel, D. A. (2011). The effects of experimenter gender on state social physique anxiety and strength in a testing environment. *Journal of Strength and Conditioning Research*, 25, 533–538. <http://dx.doi.org/10.1519/JSC.0b013e3181c1f7b3>
- Larkin, K. T., Ciano-Federoff, L. M., & Hammel, D. (1998). Effects of gender of observer and fear of negative evaluation on cardiovascular reactivity to mental stress in college men. *International Journal of Psychophysiology*, 29, 311–318. [http://dx.doi.org/10.1016/S0167-8760\(98\)00019-1](http://dx.doi.org/10.1016/S0167-8760(98)00019-1)
- Levav, J., & Argo, J. J. (2010). Physical contact and financial risk taking. *Psychological Science*, 21, 804–810. <http://dx.doi.org/10.1177/0956797610369493>
- Marx, D. M., & Goff, P. A. (2005). Clearing the air: The effect of experimenter race on target's test performance and subjective experience. *British Journal of Social Psychology*, 44, 645–657. <http://dx.doi.org/10.1348/014466604X17948>
- McGinnies, E. (1949). Emotionality and perceptual defense. *Psychological Review*, 56, 244–251. <http://dx.doi.org/10.1037/h0056508>
- Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interactions with White and Black men. *Personality and Social Psychology Bulletin*, 28, 939–952. <http://dx.doi.org/10.1177/01467202028007007>
- Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin*, 129, 674–697. <http://dx.doi.org/10.1037/0033-2909.129.5.674>
- Miller, D. A., Smith, E. R., & Mackie, D. M. (2004). Effects of intergroup contact and political predispositions on prejudice: Role of intergroup emotions. *Group Processes & Intergroup Relations*, 7, 221–237. <http://dx.doi.org/10.1177/13684302040406109>
- Moorman, P. G., Newman, B., Millikan, R. C., Tse, C.-K. J., & Sandler, D. P. (1999). Participation rates in a case-control study: The impact of age, race, and race of interviewer. *Annals of Epidemiology*, 9, 188–195. [http://dx.doi.org/10.1016/S1047-2797\(98\)00057-X](http://dx.doi.org/10.1016/S1047-2797(98)00057-X)
- Moss-Racusin, C. A., van der Toorn, J., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2014). Scientific diversity interventions. *Science*, 343, 615–616. <http://dx.doi.org/10.1126/science.1245936>

- Murthy, V. H., Krumholz, H. M., & Gross, C. P. (2004). Participation in cancer clinical trials: Race-, sex-, and age-based disparities. *Journal of the American Medical Association*, *291*, 2720–2726. <http://dx.doi.org/10.1001/jama.291.22.2720>
- National Institutes of Health. (1993). *National Institutes of Health Revitalization Act of 1993: Clinical research equity regarding women and minorities*. Retrieved from <https://orwh.od.nih.gov/resources/pdf/NIH-Revitalization-Act-1993.pdf>
- Oakes, P. J. (1987). The salience of social categories. In J. C. Turner, M. A. Hogg, P. J. Oakes, S. D. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group* (pp. 117–141). Oxford, United Kingdom: Blackwell.
- Open Science Collaboration (OSC). (2015, August 28). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. <http://dx.doi.org/10.1126/science.aac4716>
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*, 776–783. <http://dx.doi.org/10.1037/h0043424>
- Paladino, M. P., & Castelli, L. (2008). On the immediate consequences of intergroup categorization: Activation of approach and avoidance motor behavior toward ingroup and outgroup members. *Personality and Social Psychology Bulletin*, *34*, 755–768. <http://dx.doi.org/10.1177/0146161720831515>
- Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, *42*, 1051–1068. <http://dx.doi.org/10.1037/0022-3514.42.6.1051>
- Pieterse, A. L., Todd, N. R., Neville, H. A., & Carter, R. T. (2012). Perceived racism and mental health among Black American adults: A meta-analytic review. *Journal of Counseling Psychology*, *59*, 1–9. <http://dx.doi.org/10.1037/a0026208>
- Platow, M. J., Foddy, M., Yamagishi, T., Lim, L., & Chow, A. (2012). Two experimental tests of trust in in-group strangers: The moderating role of common knowledge of group membership. *European Journal of Social Psychology*, *42*, 30–35. <http://dx.doi.org/10.1002/ejsp.852>
- Rhodes, G., Hayward, W. G., & Winkler, C. (2006). Expert face coding: Configural and component coding of own-race and other-race faces. *Psychonomic Bulletin & Review*, *13*, 499–505. <http://dx.doi.org/10.3758/BF03193876>
- Robinson, M. E., & Wise, E. A. (2003). Gender bias in the observation of experimental pain. *Pain*, *104*, 259–264. [http://dx.doi.org/10.1016/S0304-3959\(03\)00014-9](http://dx.doi.org/10.1016/S0304-3959(03)00014-9)
- Rosenthal, R., & Fode, K. L. (1963). Psychology of the scientist: V. Three experiments in experimenter bias. *Psychological Reports*, *12*, 491–511. <http://dx.doi.org/10.2466/pr0.1963.12.2.491>
- Rotundo, M., & Sackett, P. R. (1999). Effect of rater race on conclusions regarding differential prediction in cognitive ability tests. *Journal of Applied Psychology*, *84*, 815–822. <http://dx.doi.org/10.1037/0021-9010.84.5.815>
- Rumenik, D. K., Capasso, D. R., & Hendrick, C. (1977). Experimenter sex effects in behavioral research. *Psychological Bulletin*, *84*, 852–877. <http://dx.doi.org/10.1037/0033-2909.84.5.852>
- Rump, E. E., & Delin, P. S. (1973). Differential accuracy in the status-height phenomenon and an experimenter effect. *Journal of Personality and Social Psychology*, *28*, 343–347. <http://dx.doi.org/10.1037/h0035117>
- Sabin, J. A., Marini, M., & Nosek, B. A. (2012). Implicit and explicit anti-fat bias among a large sample of medical doctors by BMI, race-ethnicity and gender. *PLoS ONE*, *7*(11), e48448. <http://dx.doi.org/10.1371/journal.pone.0048448>
- Scheepers, D., Saguy, T., Dovidio, J. F., & Gaertner, S. A. (2014). A shared dual identity promotes a cardiovascular challenge response during interethnic interactions. *Group Processes & Intergroup Relations*, *17*, 324–341. <http://dx.doi.org/10.1177/1368430213517271>
- Schiebinger, L., Klinge, I., Sánchez de Madariaga, I., Paik, H. Y., Schraudner, M., & Stefanick, M. (Eds.). (2011–2015). *Gendered innovations in science, health & medicine, engineering and environment*. Retrieved from <http://genderedinnovations.stanford.edu>
- Schulman, K. A., Berlin, J. A., Harless, W., Kerner, J. F., Sistrunk, S., Gersh, B. J., . . . Escarce, J. J. (1999). The effect of race and sex on physicians' recommendations for cardiac catheterization. *New England Journal of Medicine*, *340*, 618–626. <http://dx.doi.org/10.1056/NEJM199902253400806>
- Senn, C. Y., & Desmarais, S. (2001). Are our recruitment practices for sex studies working across gender? The effect of topic and gender of recruiter on participation rates of university men and women. *Journal of Sex Research*, *38*, 111–117. <http://dx.doi.org/10.1080/00224490109552078>
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*, 80–83. <http://dx.doi.org/10.1111/1467-9280.00111>
- Sinclair, S., Lowery, B. S., Hardin, C. D., & Colangelo, A. (2005). Social tuning of automatic racial attitudes: The role of affiliative motivation. *Journal of Personality and Social Psychology*, *89*, 583–592. <http://dx.doi.org/10.1037/0022-3514.89.4.583>
- Skorinko, J. L., Lun, J., Sinclair, S., Marotta, S. A., Calanchini, J., & Paris, M. H. (2015). Reducing prejudice across cultures via social tuning. *Social Psychological and Personality Science*, *6*, 363–372. <http://dx.doi.org/10.1177/1948550614561125>
- Smith, L. (2005). Psychotherapy, classism, and the poor: Conspicuous by their absence. *American Psychologist*, *60*, 687–696. <http://dx.doi.org/10.1037/0003-066X.60.7.687>
- Sorge, R. E., Martin, L. J., Isbester, K. A., Sotocinal, S. G., Rosen, S., Tuttle, W. F., . . . Mogil, J. S. (2014). Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nature methods*, *11*, 629–632. <http://dx.doi.org/10.1038/nmeth.2935>
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797–811. <http://dx.doi.org/10.1037/0022-3514.69.5.797>
- Taylor, V. J., & Walton, G. M. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin*, *37*, 1055–1067. <http://dx.doi.org/10.1177/0146167211406506>
- Thompson, V. L. S., & Alexander, H. (2006). Therapists' race and African American clients' reactions to therapy. *Psychotherapy: Theory, Research, Practice*, *43*, 99–110. <http://dx.doi.org/10.1037/0033-3204.43.1.99>
- Trawalter, S., Richeson, J. A., & Shelton, J. N. (2009). Predicting behavior during interracial interactions: A stress and coping approach. *Personality and Social Psychology Review*, *13*, 243–268. <http://dx.doi.org/10.1177/1088868309345850>
- Tsugawa, Y., Jena, A. B., Figueroa, J. F., Orav, E. J., Blumenthal, D. M., & Jha, A. K. (2017). Comparison of hospital mortality and readmission rates for Medicare patients treated by male vs. female physicians. *Journal of the American Medical Association: Internal Medicine*, *177*, 206–213. <http://dx.doi.org/10.1001/jamainternmed.2016.7875>
- Turner, J. C., Brown, R. J., & Tajfel, H. (1979). Social comparison and group interest in in-group favoritism. *European Journal of Social Psychology*, *9*, 187–204. <http://dx.doi.org/10.1002/ejsp.2420090207>
- Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016a). Contextual sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 6454–6459. <http://dx.doi.org/10.1073/pnas.1521897113>
- Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016b). Reply to Inbar: Contextual sensitivity helps explain the reproducibility gap between social and cognitive psychology. *Proceedings of*

- the National Academy of Sciences of the United States of America*, 113, E4935–E4936. <http://dx.doi.org/10.1073/pnas.1609700113>
- van der Lee, R., & Ellemers, N. (2015). Gender contributes to personal research funding success in the Netherlands. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 12349–12353. <http://dx.doi.org/10.1073/pnas.1510159112>
- Vrana, S. R., & Rollock, D. (1998). Physiological response to a minimal social encounter: Effects of gender, ethnicity, and social context. *Psychophysiology*, 35, 462–469. <http://dx.doi.org/10.1111/1469-8986.3540462>
- Wang, Q. (2016). Why should we all be cultural psychologists? Lessons from the study of social cognition. *Perspectives on Psychological Science*, 11, 583–596. <http://dx.doi.org/10.1177/1745691616645552>
- Weisbuch, M., Pauker, K., & Ambady, N. (2009, December 18). The subtle transmission of race bias via televised nonverbal behavior. *Science*, 326, 1711–1714. <http://dx.doi.org/10.1126/science.1178358>
- Weisse, C. S., Foster, K. K., & Fisher, E. A. (2005). The influence of experimenter gender and race on pain reporting: Does racial or gender concordance matter? *Pain Medicine*, 6, 80–87. <http://dx.doi.org/10.1111/j.1526-4637.2005.05004.x>
- Worchel, S., Rothgerber, H., Day, E. A., Hart, D., & Butemeyer, J. (1998). Social identity and individual productivity within groups. *British Journal of Social Psychology*, 37, 389–413. <http://www.ncbi.nlm.nih.gov/pubmed/9933910>
- Xiao, Y. J., Coppin, G., & Van Bavel, J. J. (2016). Perceiving the world through group-colored glasses: A perceptual model of intergroup relations. *Psychological Inquiry*, 27, 255–274. <http://dx.doi.org/10.1080/1047840X.2016.1199221>
- Yasui, M., & Dishion, T. J. (2008). Direct observation of family management: Validity and reliability as a function of coder ethnicity and training. *Behavior Therapy*, 39, 336–347. <http://dx.doi.org/10.1016/j.beth.2007.10.001>

Appendix

Recommended Considerations Regarding Research Staff Demographics

Introduction–Hypotheses

- Was the literature consulted (and reported) on whether research staff demographics may or may not influence the phenomenon and/or population(s) under investigation?

Procedure–Method

- Have demographical in- versus outgroups of participants been identified, considered, and reported in the research design? For example, if participants are Black students, then White experimenters would be outgroup members and Black experimenters would be ingroup members.
- Are race, gender, and age (and any other relevant information) of *participants* reported?
- Are race and gender (and any other relevant information) of *research staff* reported?
- If using multiple research staff members, is it reported how they were assigned to condition and/or participants (e.g., whether they were randomly assigned to conditions and participants or whether there was matching in terms of creating race or gender congruence between researchers and participants)?

Reporting–Results

- Can the collected data be disaggregated by participants' and research staff's gender and race to allow for later analysis at the unit record level and/or inclusion in meta-analysis?
- Are any analyses that were done to test or control for the influence of research staff demographics on findings reported?
- If using one research staff member, is it reported how this resulted in race or gender (in)congruence with some participants and how that was addressed and/or controlled for in the analysis?
- Is information on research staff demographics reflected in data that is represented in figures and tables, either in the notes or in the tables or figures themselves?

Discussion

- Is the role of research staff demographics in the current work and/or potential directions regarding research staff demographics for future work discussed?

Received July 11, 2016

Revision received July 24, 2017

Accepted July 26, 2017 ■