Journal of Personality and Social Psychology

Perceiver and Target Characteristics Contribute to Impression Formation Differently Across Race and Gender

Sally Y. Xie, Jessica K. Flake, and Eric Hehman

Online First Publication, December 13, 2018. http://dx.doi.org/10.1037/pspi0000160

American Psychological Association

Journal of Personality and Social Psychology:
Interpersonal Relations and Group Processes

0022-3514/18/$12.00

http://dx.doi.org/10.1037/pspi0000160

Perceiver and Target Characteristics Contribute to Impression Formation Differently Across Race and Gender

Sally Y. Xie, Jessica K. Flake, and Eric Hehman
McGill University

Social impressions arise from characteristics of both perceivers and targets. However, empirical research in the domain of impression formation has yet to quantify the extent to which perceiver and target characteristics uniquely contribute to impressions across group boundaries (e.g., race, gender). To what extent does an impression arise from "our mind" versus "a target's face", and does this process differ for impressions across race and gender? We explored this question by estimating intraclass correlation coefficients (ICCs) from cross-classified multilevel models of 188,472 face ratings from 2,230 participants (Study 1) and 219,658 ratings from 2,984 participants (Study 2). We partitioned the total variance in ratings on a trait dimension (trustworthiness, dominance, youthful/attractiveness) into variance explained by perceivers versus targets, and compared these ICCs among different groups (e.g., ratings of own- vs. other-group targets). Overarching results reveal (a) target appearance matters more for women than men, (b) target appearance matters more for impressions on youthful/attractiveness than trustworthiness or dominance dimensions, (c) differences in perceiver/target influences across race did not consistently replicate, and (d) these differences are absent in minimal groups, supporting the role of racial and gender stereotypes in driving these effects. Overall, perceiver characteristics contribute more to impressions than target appearance. Our findings disentangle the contributions of perceiver and targets to impressions and illustrate that the process of impression formation is not equal across various group boundaries.

Keywords: impression formation, person perception, face perception, multilevel modeling

Supplemental materials: http://dx.doi.org/10.1037/pspi0000160.supp

How subjective are our impressions of others, and under what circumstances might these impressions become more or less subjective? For instance, is beauty in the eye of the beholder, or in the face of the target? Although it seems intuitive that the physical appearance of a person influences our perception of them, philosophers have long contemplated the contributions of the perceiver, with social scientists in recent decades empirically demonstrating that perceivers play a critical—and independent—role in driving social impressions. More than just the passive processing of sensory stimuli, perception is an active mental interpretation of the external world, during which perceivers ascribe meaning to targets. Accordingly, that social impressions arise from both perceivers and targets has become central to modern social cognition, with many influential models demonstrating how characteristics of the target (e.g., facial cues, race, gender) and characteristics of the perceiver (e.g., motivation, cognition) together shape impressions (Brewer, 1988; Bruce & Young, 1986; Brunswik, 1952; Correll, Hudson, Guillermo, & Earls, 2016; Fiske & Neuberg, 1990; Freeman & Ambady, 2011; Haxby, Hoffman, & Gobbini, 2000; Kenny & Albright, 1987; Kunda et al., 1996; Neuberg & Fiske, 1987; West & Kenny, 2011).

These models emphasize the joint influence of perceiver and target characteristics in impression formation, particularly in the domains of social attributions and face perception (Freeman & Ambady, 2011; Rhodes, 2006; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015; Webster & Macleod, 2011; Zebrowitz & Montepare, 2005). Despite general agreement that both perceiver and target factors guide impressions, research on the extent to which they independently contribute to any given impression remains relatively scarce and has only recently been quantified in impressions from faces (Hehman, Sutherland, Flake, & Slepian, 2017; Hönekopp, 2006). Parsing the unique contributions of perceiver and target factors to impressions is critical to understanding the nature and process of how we form impressions. For instance, some trait impressions may be particularly target-driven, such that perceivers who are quite different consistently arrive at similar impressions of the same target (Hehman et al., 2017). In contrast, some trait impressions may be particularly perceiver-driven, such that differences between perceivers are largely responsible for variation in ratings

Sally Y. Xie, Jessica K. Flake, and Eric Hehman, Department of Psychology, McGill University.

All authors designed the analyses. Sally Y. Xie aggregated the data. All authors analyzed the data. All authors wrote the manuscript.

We thank Lina Kanawati and Geovanni Badiola for assistance in materials development and data collection. This research was partially supported by a SSHRC Institutional Grant and SSHRC Insight Development Grant (430-2016-00094) to Eric Hehman.

Correspondence concerning this article should be addressed to Sally Y. Xie, Department of Psychology, McGill University, 2001 McGill College Avenue, Montreal, QC, Canada H3A 1G1. E-mail: sally.xie@mail.mcgill.ca

rather than the targets themselves (Hehman et al., 2017). Finally, some impressions may rely upon the unique interplay between target and perceiver, such that we must consider characteristics of both to understand the formed impression. Critically, the processes by which different impressions arise may be quite different: Without knowing the extent to which perceiver- and target-level factors interact to contribute to social impressions, the process of impression formation itself remains obscure.

To that end, the present research aims to advance our understanding of how perceiver and target characteristics contribute to first impressions. We build on recent research quantifying the extent to which perceiver versus target factors uniquely contribute to social impressions (Hehman et al., 2017), focusing instead on how these perceiver-target contributions vary across group boundaries such as gender and race. We first briefly review target, perceiver, and Target × Perceiver contributions to social impressions.

Perceiver and Target Contributions to Social Impressions

Examples of first impressions guided by perceiver characteristics, target characteristics, and the interplay between the two abound throughout the social–cognitive and intergroup literatures (for review, see Todorov et al., 2015; Zebrowitz & Montepare, 2008). Facial impressions are informed by both morphological features and social identity cues (e.g., gender and race; Freeman & Ambady, 2011; Freeman, Pauker, Apfelbaum, & Ambady, 2010; Hehman, Ingbretsen, & Freeman, 2014; Kubota & Ito, 2007). For instance, facial features resembling emotional expressions (e.g., an upturned mouth resembling a smile) are generalized to stable trait inferences congruent with these emotional cues (e.g., a friendly person; Adams, Nelson, Soto, Hess, & Kleck, 2012; Oosterhof & Todorov, 2009; Said, Sebe, & Todorov, 2009; Secord, 1958; Secord, Dukes, & Bevan, 1954; Zebrowitz, Kikuchi, & Fellous, 2010; Zebrowitz & Montepare, 2008). Similar results have also been found for static morphological features: faces with higher facial width-to-height ratio elicit perceptions of aggressiveness, physical strength, and dominance (Carré, McCormick, & Mondloch, 2009; Carré, Morrissey, Mondloch, & McCormick, 2010; Hehman, Leitner, Deegan, & Gaertner, 2015; Hehman, Leitner, & Gaertner, 2013).

However, perceivers are not passive observers forming objective impressions. A panoply of research indicates that individual characteristics varying across different perceivers may additionally provide top-down influences on perception. These perceiver characteristics might include personality traits or the context in which a stimulus is perceived. For instance, whether a target appears threatening depends on whether the perceiver is a parent (Fessler, Holbrook, Pollack, & Hahn-Holbrook, 2014) or in a group (Fessler & Holbrook, 2013). The race and prejudice of the perceiver can also influence impressions: Black male targets are perceived as larger and more threatening by non-Black and high-prejudice perceivers (Wilson, Hugenberg, & Rule, 2017). Furthermore, whether perceivers think about targets in a more categorical or individuated manner may depend on their level of prejudice (for review, see Macrae & Bodenhausen, 2000).

Finally, characteristics of both perceivers and targets can interact to uniquely influence impression formation. For example, one perceiver may find brunettes particularly attractive, but not blondes; another might feel the opposite. Here, ratings of attractiveness depend on both the characteristics of the target (being blond or brunette) and individual differences between perceivers (preferring blondes or brunettes). Interactions between perceivers and targets are some of the most interesting to intergroup and social–cognitive researchers and are commonly examined throughout the field. For instance, the racial prejudice of perceivers facilitates the perception of hostile (but not happy) racially ambiguous faces as Black (Hugenberg & Bodenhausen, 2004; Hutchings & Haddock, 2008). In this example, both the prejudice of the perceiver and the emotional expression of the target together inform the categorization of targets as Black. The interplay between perceiver and target in determining impressions is the focus of the current research. Specifically, we focus on cross-group perceptions, when the target and perceiver belong to different groups (e.g., race, gender).

Intergroup Perception

The importance of group membership has long been recognized (Allport, 1954). People readily categorize themselves as members of groups (Tajfel & Turner, 1986; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987), and group identity impacts how we perceive, evaluate, and behave toward members of other groups (Cikara & Van Bavel, 2014; Gaertner, Dovidio, Anastasio, Bachman, & Rust, 1993; Kubota & Ito, 2017; Macrae & Bodenhausen, 2000; Ratner, Dotsch, Wigboldus, van Knippenberg, & Amodio, 2014; Turner et al., 1987). For instance, an extensive literature examining the initial perception and categorization of own- and other-group targets find that the early visual processing of targets as either own- or other-group influences subsequent stereotyping and evaluation of those targets (for review, see Kawakami, Amodio, & Hugenberg, 2017). Although there is abundant evidence that the product of own- versus other-group perceptions are different, there is less regarding how the process of impression formation varies across group boundaries.

Modern models of impression formation implicitly support this prediction. For instance, the dynamic interactive model of person perception suggests that the processing of bottom-up facial features is dynamically constrained by top-down cognitive states (e.g., motivation, arousal) and stereotype activation (Freeman & Ambady, 2011). To the extent that stereotype activation is involved in the processing of individuals from different social categories, we might expect the process of impression formation to vary. Stereotypes (or other top-down cognitive influences) would contribute more to impressions of one target than to another.

Individuals tend to have stereotypes or expectations about members of other groups and how they might behave (Fiske, Cuddy, Glick, & Xu, 2002; Hehman, Volpert, & Simons, 2014; Jussim, Coleman, & Lerch, 1987). These stereotypes act as templates upon which perceivers rely when forming impressions. Importantly, individuals rely less on stereotypes, and process individuals to a greater extent, when evaluating members of their own group (Cikara & Van Bavel, 2014; Neuberg & Fiske, 1987; Rogers & Biesanz, 2014). For instance, research examining the cross-race effect within the face-recognition literature has long demonstrated superior recognition for members of one's own group relative to other groups (Bernstein, Young, & Hugenberg, 2007; Hehman,

Mania, & Gaertner, 2010; Meissner & Brigham, 2001; Sporer, 2001). This is thought to be driven, in part, by greater processing and individuation of ingroup members, whereas outgroup members are processed more categorically and are subject to top-down influences like motivation and stereotype content (Brewer, 1988; Freeman & Ambady, 2011; Hugenberg, Young, Bernstein, & Sacco, 2010; Neuberg & Fiske, 1987; Rogers & Biesanz, 2014).

Social neuroscience has provided key evidence of different processes of impression formation for targets in different social categories. For example, event-related potentials diverged as early as 120 ms when viewing targets of different races and genders, regardless of whether perceivers attended to social category or individuating information (Ito & Urland, 2003, 2005; Kubota & Ito, 2007, 2017). Numerous fMRI studies have implicated a network of brain regions, which differentially activate to own- and other-group faces, which most commonly include the amygdala, fusiform face area, dorsolateral prefrontal cortex (dlPFC), and the anterior cingulate cortex (ACC; Amodio, 2014; Cikara & Van Bavel, 2014; Cunningham et al., 2004; Golby, Gabrieli, Chiao, & Eberhardt, 2001; Hehman, Ingbretsen, et al., 2014; Kubota, Banaji, & Phelps, 2012; Stanley et al., 2012; Stolier & Freeman, 2016; Van Bavel, Packer, & Cunningham, 2008). For instance, the amygdala is critical for the acquisition, storage, and expression of fear conditioning (Phelps & Ledoux, 2005), and its consistent differential BOLD response in own- and other-group targets highlights the potential role of threat in cross-group perceptions. In contrast, the dlPFC and ACC have been implicated in social cognition generally, and research indicates they play a larger role in accessing and suppressing stereotypic content (Amodio & Frith, 2006).

This body of work has been invaluable in detailing how fundamental the processing of social group membership is in human perception, as well as the contributions of different factors such as threat perception and stereotypes. What this research has not done, however, is quantify to what extent perceiver and target characteristics differentially contribute to a final impression, and how these contributions differentially inform the process of impression formation across group boundaries. By precisely estimating the percentage of variance in an impression that arises from "the perceiver's mind" versus "the target's face," we can examine whether this breakdown varies across group boundaries such as race and gender. Research that does not estimate the potentially different contributions of perceiver and target characteristics to own- and other-group impressions is making a functional assumption that their contributions are equal. Extant evidence indicates this is not the case, and we do not believe that most researchers would believe this to be the case. Thus, our aim was to use statistical models to estimate potentially different perceiver and target contributions to impression formation, quantifying to what extent these processes vary across groups.

Cross-Classified Multilevel Models to Quantify Perceiver and Target Effects

Relatively recent advances in statistical modeling can now address this issue, allowing researchers to estimate the unique influence of perceiver and target characteristics. Partitioning variance in impressions between perceivers and targets has a rich history in social psychology, primarily in the work of Kenny and colleagues modeling impressions from in-person dyadic interactions with the social relations model (Kenny & Albright, 1987; West & Kenny, 2011; for review, see Kenny, 1994). Recent derivations of this approach, such as cross-classified multilevel models, can decompose and quantify the variance in impressions originating at the target and perceiver levels (Judd, Westfall, & Kenny, 2012; Raudenbush & Bryk, 2002).

When repeated observations (i.e., ratings of different targets) are clustered (i.e., multiple ratings made by a single perceiver), estimates from these models are used to calculate intraclass correlation coefficients (ICCs). These ICCs represent the percentage of variance in a dependent variable explained by different clusters of the multilevel model. In the current context, these clusters (perceiver characteristics, target characteristics, or the interplay between the two) explain the percentage of variance in a trait rating. Recently, this modeling approach has been extended to face perception, calculating the percentage of variance in impressions coming from the perceiver versus the face itself (Hehman et al., 2017; Hönekopp, 2006).

In the current context, the dependent variable is a participant's rating of a target on a specific trait (e.g., trustworthiness), which can be clustered at both the perceiver level (i.e., multiple ratings of various targets made by a single perceiver) and the target level (i.e., multiple ratings of one target made by various perceivers). Thus, the data are clustered in a cross-classified fashion. The ICCs are calculated for the different clusters (i.e., the target and the perceiver) within the model. Here, the *perceiver-ICC* represents the percentage of variance in ratings that stems from between-perceiver variability (i.e., variability in the characteristics of different perceivers), which might be present due to stable perceiver differences (e.g., race, gender, traits) or temporary factors (e.g., arousal). The *target-ICC* represents the percentage of variance in ratings that comes from between-target characteristics (i.e., variability in the appearance of targets). In other contexts (i.e., rating the similarity between the self and a target) these perceiver- and target-ICCs have been referred to as "assimilation" and "consensus" respectively (Kenny, 1994; Kenny & West, 2010). Finally, the *interaction-ICC* represents the percentage of variance that stems from the unique interaction between targets and perceivers, in some research labeled "personal taste" (Hönekopp, 2006). Using our previous example, one perceiver may find brunettes particularly attractive, but not blondes; another perceiver might feel the opposite. The attractiveness judgments in this example arise from interactions between perceiver preferences and target characteristics.

Recent research used ICCs from cross-classified models to estimate the extent to which perceiver- and target-level characteristics contribute to a wide variety of impressions and dimensions commonly examined in social–cognitive research (e.g., trustworthiness, dominance; Hehman et al., 2017). This work revealed a surprising amount of variability in the relative contributions of perceiver and target factors to ratings of different traits and dimensions. For instance, perceiver-level factors contributed 34% to ratings of "intelligence," whereas target-level factors contributed only 13%. In contrast, perceiver-level factors contributed only 17% to ratings of dominance, whereas target-level factors contributed 22%. These results indicate that impressions of intelligence are far more perceiver-driven than impressions of dominance, whereas the characteristics of targets play a larger role in deter-

mining ratings of dominance than of intelligence. Thus, although previous research had made the implicit, yet functional, assumption that perceiver and target contributions were equivalent across traits and dimensions, these results revealed that the process of impression formation for different traits varied—perceiver and target characteristics contributed different amounts.

Although this work was an important demonstration of the variability that exists in the relative contributions of perceivers and targets to different impressions, it was limited in that some of the most important and interesting questions were not examined. Namely, how the process of perception might vary across group boundaries. Previous work with this approach examined perceiver and target contributions to different traits for either only own-group perceptions, or collapsing across different groups (e.g., along gender, race) without differentiating between ingroup and outgroup perceptions. As such, the question of whether perceiver and target characteristics contribute differently to impressions for ingroup and outgroup members, or whether the process of perception is different, remains unanswered.

The Present Research

The current research aimed to answer this fundamental question. Our goals were threefold: (a) to estimate perceiver- and target-level contributions to impressions across race and gender (Study 1); (b) to compare and test broader patterns in these estimates (Study 1B); and (c) to examine whether differences in these estimates are driven by group membership or group-based stereotypes, by comparing estimates of perceiver and target contributions for impressions across minimal groups (Study 2). These goals are addressed in subsequent analyses using both an exploratory and confirmatory dataset.

Study 1: Comparing ICCs Across Race and Gender Groups

In Study 1, we tested our hypothesis that impressions made across group boundaries would differ in the relative contributions of perceiver- and target-level factors. For instance, when evaluating a target from a racial outgroup, perceivers may be relying upon their knowledge of racial stereotypes during the impression formation process, which causes the actual facial appearance of the target being evaluated to have a smaller impact (i.e., smaller target-ICC) relative to perceiver-level factors (i.e., larger perceiver-ICC) in these cross-group impressions.

In light of existing social hierarchies within North America, we expected racial majority and minority status to shape how own- and other-group targets are perceived. Status has been operationalized as social power, group size, wealth, prestige, among other factors (Mattan, Kubota, & Cloutier, 2017). In North America, White people have historically been higher in status than racial minority groups as defined by many of these factors (Marger, 2003; Sachdev & Bourhis, 1991). We anticipated that impressions might be formed differently across majority-minority group boundaries as a result of these different statuses. Specifically, even though majority-group members viewing minorities and minorities viewing majorities are both racial cross-group impressions, we did not anticipate the process of impression formation to be equivalent across both scenarios because of differences in status. Previous research has demonstrated that majority versus minority status alone can shape one's attitudes and preferences toward other groups (Hehman et al., 2012; Tropp & Pettigrew, 2005; Verkuyten, 2005). Thus, we partitioned predictions across racial groups by racial majority or minority status (i.e., majorities viewing minorities, minorities viewing majorities, etc.) to understand how perceiver and target contributions in impressions may vary across these types of perceptions. We refer to this different status throughout as majority or minority racial group status, though the specific elements of status (e.g., group size, power) that are responsible for any effects here are unknown.

Our hypothesis that the process of impression formation varies in cross-group perceptions hinges on the assumption that perceivers use stereotype content as a template to form their impressions. Yet stereotypes are shaped by not only race and social status, but also gender. For instance, stereotypes for Black men and women in North America are very different, and these individuals are perceived and treated in distinct ways (Purdie-Vaughns & Eibach, 2008). Furthermore, race and gender appear to be inextricably linked (Freeman & Ambady, 2011; Johnson, Freeman, & Pauker, 2012; McDonald, Navarrete, & Van Vugt, 2012; Navarrete et al., 2009; Navarrete, McDonald, Molina, & Sidanius, 2010; Van Vugt, De Cremer, & Janssen, 2007). Gender categorization of faces is facilitated when the target's race and gender are stereotypically congruent (e.g., Asian = female, Black = male), and hindered when incongruent (Johnson et al., 2012). Accordingly, to consider either race, social status, or gender independently would mask potential differences in the process of perception for each of these groups, introducing error into the estimates of variance. Thus, we hypothesized that perceiver- and target-ICCs for own- and other-group perceptions will additionally depend on whether targets are female or male.

We have argued that it is necessary to consider race, social status, and gender simultaneously because stereotype content varies as a function of these factors. Yet stereotype content also differs across specific racial groups. For instance, stereotypes regarding Asians and Blacks, though both racial minorities in North America, are quite distinct (Fiske et al., 2002). Accordingly, although we combined perceivers from these two groups along with other minority groups to maximize the number of observations—and thus the stability of the ICC estimates—doing so may have masked important differences in the extent to which perceiver- and target-characteristics influenced the process of perception. To examine this possibility, in a subsequent analysis, we further partitioned the data based on whether the targets and perceivers were Black, Asian, or White. We did so only for these groups as they were the three largest racial groups among both our stimuli and participants.

Study 2: Minimal Groups

Finally, because our rationale for expecting different perceiver- and target-ICCs across groups hinged on stereotypic associations with these racial and gender groups, we tested this possibility directly in Study 2 using a minimal group paradigm (Ratner & Amodio, 2013; Tajfel, Billig, Bundy, & Flament, 1971). Specifically, we created groups for which no stereotype content was available, and then examined perceiver- and target-characteristic contributions to minimal cross-group impressions. If stereotype

content was responsible for differences in perceiver- and target-ICCs in cross-group perceptions, then these differences should at least partially disappear when we compare arbitrary groups from the minimal group paradigm. In contrast, if perceiver- and target-ICC differences are based on group membership alone (absent stereotype information), then any differences observed in Study 1 should persist with minimal groups. We tested these possibilities in Study 2.

Methods

Dimensional Space Underlying Impression Formation

There are innumerable traits to describe and evaluate individuals, and many of these traits—such as "smart" and "intelligent" – are highly correlated. Past researchers have used data-reduction techniques to distill these traits into the foundational dimensions underlying face perception. Most work has found two dimensions, one thought to reflect whether a target's intentions are good or ill: "trustworthiness/intentions," and another to reflect that target's ability to enact those intentions: "dominance/ability" (Fiske, Cuddy, & Glick, 2007; Fiske et al., 2002; Freedman, Leary, Ossorio, & Coffey, 1953; Leary, 1957; Oosterhof & Todorov, 2008; Todorov, Said, Engell, & Oosterhof, 2008). In addition, recent research with a more representative stimulus set has identified a third factor of "youthful/attractive" (Sutherland et al., 2013). We used this literature to inform which traits load onto which dimensions (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Todorov et al., 2008) and focused our analyses on the dimensions of trustworthiness, dominance, and youthfulness/attractiveness when calculating ICCs.

Analytic Approach

We executed a series of multilevel models in R (lme4: Bates, Mächler, Bolker, & Walker, 2015) to estimate the ICC for each model. In each model, participants' ratings of stimuli on the dimension of interest (e.g., trustworthiness) served as the single dependent variable. We partitioned the variance in ratings of that dimension into three components: (a) variance attributed to the target, (b) variance attributed to the perceiver, and (c) residual or error variance. Crucially, these null or intercept-only models are cross-classified: ratings were nested both within perceivers (i.e., each perceiver rating multiple targets) and within targets (i.e., with each target being rated by multiple perceivers), allowing us to calculate an ICC for both perceiver and target characteristics on each dimension (Judd et al., 2012; Raudenbush & Bryk, 2002).

As described in previous work (Hehman et al., 2017), the cross-classified multilevel model can be defined with two equations, one for each level of the model:

$$\text{Level 1}: \; Y_{i(j_1 j_2)} = \pi_{0(j_1 j_2)} + e_{i(j_1 j_2)} \tag{1}$$

$$\text{Level 2}: \; \pi_{0(j_1 j_2)} = \theta_{000} + b_{0j_1 0} + c_{00j_2} + d_{0(j_1 j_2)} \tag{2}$$

In the first level of this model (Equation 1), $Y_{i(j_1 j_2)}$ is our dependent variable of interest: a rating on dimension i (e.g., trustworthiness) of target j_1 by perceiver j_2. The intercept, $\pi_{0(j_1 j_2)}$ is the expected value of this rating for this target by this perceiver, and the error term, $e_{i(j_1 j_2)}$, has its own associated variance, σ^2. In the second level of the model (Equation 2), the intercept $\pi_{0(j_1 j_2)}$ is an outcome that can vary across perceivers and targets, which allows the total variance of the model to be partitioned into that attributable to perceivers and targets. The grand mean, θ_{000}, is the average rating across all perceivers and targets; $b_{0j_1}0$ is the residual of target j_1 (i.e., the difference between the grand mean and the rating of target j_1), which has variance τ_{b00}; and c_{00j_2} is the residual of perceiver j_2 averaged across all targets, which has variance τ_{c00}. Thus, we have three variance terms in the model: variance across targets, τ_{b00}, variance across perceivers, τ_{c00}, and the variance of the level-1 error term, σ^2. Together, these comprise 100% of the variance in ratings on any dimension. By looking at the size of each variance component relative to the total variance, we can estimate the proportions of variance that come from different elements of the model (i.e., perceivers vs. targets).

Specifically, we used these variance components to calculate perceiver- and target-ICCs (Raudenbush & Bryk, 2002). For example, perceiver-ICC is calculated as the proportion of variance attributable to perceiver characteristics (Equation 3):

$$ICC_{perceiver} = \frac{\tau_{b00}}{\tau_{b00} + \tau_{c00} + \sigma^2} \tag{3}$$

A final potential random effect, $d_{0(j_1 j_2)}$, represents the interaction between perceiver and target variance in the model. Earlier, we stated our interest in this Perceiver × Target interaction, which describes how the process of perception might be different depending on both the characteristics of the perceiver (e.g., race and gender) and characteristics of the target (e.g., race and gender). One approach to examining these interactions would be to directly estimate the random effect associated with the Perceiver × Target interactions in a single multilevel model. To do so requires repeated ratings of a single target by each perceiver, to parse the variance associated with this interaction from the residual variance. Without these repeated ratings, the variance of this Perceiver × Target interaction cannot be disentangled from the level-1 error variance (Beretvas, 2008; Hehman et al., 2017; Raudenbush & Bryk, 2002). Because this data structure was not present in our dataset, and indeed is not common in the impression formation literature in general, we were unable to estimate this component and it is fixed to zero. Therefore, in the reported models the Perceiver × Target interaction is inseparable from residual variance, σ^2.

We conducted these cross-classified, multilevel models across two waves of data collection to address the exploratory nature of our research questions. In Study 1A, we explored observed patterns in the data. Study 1B was confirmatory in nature: we specifically targeted and reexamined some of the differences that emerged in the exploratory Study 1A to test whether they would generalize to different participants and stimuli.

To summarize our analysis pipeline for Study 1, we (a) created subsets of our data according to the race, gender, and majority/minority group status of perceivers and targets, such that each dataset represented a particular social impression (e.g., White participants' ratings of White male targets); (b) ran a separate cross-classified null model for *each* of these subsets, estimating perceiver- and target-ICCs from each model; and (c) quantitatively compared ICCs across different models and across broader race/gender categories by bootstrapping 95% CIs around the ICCs of

each group, and examining the differences in these CIs across broader categories (e.g., ratings of all female vs. all male targets, averaging across target race). Finally, using the full dataset, we also (d) built larger conditional models and specified each social category variable (e.g., target gender) as a moderator with fixed and random effects in the model, in order to quantify how much variance in impressions was explained by a specific predictor (e.g., target gender). At each stage, we sought to replicate findings with the confirmatory dataset. We describe these steps in detail in subsequent sections.

Comparing ICCs from subset null models. To estimate and compare ICCs for each group, we initially created subsets of our data according to the race, gender, and majority-minority status of perceivers and targets. In each model, ratings on a particular dimension (trustworthiness, dominance, or youthful/attractiveness) served as the dependent variable. We built separate null models for each social impression (e.g., racial-majority perceivers evaluating racial-minority males on dominance) and estimated perceiver and target ICCs for each model. See the online supplementary materials for sample R code and calculations.

Although our aims were primarily descriptive in nature, for selected comparisons we tested whether ICCs for one group differed from one another by bootstrapping 95% confidence intervals (CIs) around their ICCs and examining the overlap in intervals. These intervals allowed us to make quantitative comparisons between groups.[1] R code is available at [osf.io/anwx2] and [hehmanlab.org/toolbox]. Although there were many potential comparisons to explore, we were interested in trends on a broader level. For instance, instead of comparing ratings of Asian female versus Asian male targets specifically, we also wanted to compare ratings of female versus male targets more generally. To compare CIs across these broader social categories (e.g., all ratings of males vs. all ratings of females), we averaged the ICC intervals of groups belonging to some relevant category (e.g., target gender) while collapsing across other categories (i.e., target race). This allowed us to compare ICCs for all ratings of, for example, male versus female targets. See the online supplementary materials for greater detail.

Variance explained in conditional models. Although informative, comparing ICCs does not provide information about the amount of variance in impressions explained by our predictors of interest (e.g., race, gender, majority-minority status). For instance, after comparing ICCs by target gender, we may want to quantify how much variance target gender explains in our models. To formally test for differences among our groups, we turned to a different analytic approach in which we build and test highly specified, conditional models.

We did so by returning to our complete, nonsubsetted dataset. This analysis models all responses together as one large, conditional model with predictors, rather than separate null models. Instead of creating subsets for each social impression, we include all perceiver and target groups as moderators in the model, each with a random slope (nested in either perceiver or target effects). Specifically, we contrast-coded four predictors of interest: perceiver gender, perceiver racial majority/minority status, target gender, and target racial majority/minority status. We then built one maximally identified model with all moderators included as fixed and random effects to estimate the variance explained when *all* perceiver/target characteristics were in the model. See the online supplementary materials for code and greater detail.

Source of the Data

Study 1A: Exploratory dataset. To create precise and generalizable estimates of perceiver- and target-ICCs, a large amount of data was required. To that end, data for the exploratory dataset were aggregated from data collected previously for other purposes and new data collected specifically for the current project. Previously existing data were included if they consisted of social perception ratings of face stimuli and had relevant participant demographic information (i.e., race, gender). All data were either collected online (using Amazon Mechanical Turk and Qualtrics) or in labs. Across all data sets, participants rated photos of facial stimuli on various social perceptions using a 7-point Likert scale (e.g., "How trustworthy is this person?") ranging from 1 (*not at all*) to 7 (*very much*"). Targets were presented in random order, and participants rated each target on only one trait such that all ratings were between-subjects. Online participants were drawn from the MTurk Worker pool across North America, included workers with an approval rating above 90%, and received monetary compensation through MTurk. Lab participants were recruited from a psychology subject pool for course credit. Data collection spanned 2011 to 2018 across 32 studies, in projects both published and unpublished.

These criteria resulted in 188,472 ratings of trait impressions (e.g., friendly) across 2,230 participants and 820 stimuli. Participant race categories included White, Black, Asian, American Indian/First Nation, Pacific Islander, Hispanic, Mixed Race, and Other. Participants who opted not to disclose their race were excluded. Participants were aged 18 to 85 (M_{age} = 36.14, SD_{age} = 12.38, 64% female) and were 74.4% non-Hispanic White, 10.1% Black, 7.3% Asian, and 8.2% other ethnic minorities. Participant ratings of trait impressions included aggressive (n = 6,232), assertive (n = 4,471), attractive (n = 34,365), caring (n = 4,296), competent (n = 15,907), dominant (n = 14,823), friendly (n = 14,360), healthy (n = 4,185), intelligent (n = 16,429), likable (n = 5,401), physically strong (n = 18,701), smart (n = 3,827), trustworthy (n = 10,690), warm (n = 14,373), and youthful (n = 20,412).

These trait ratings were used to create the dimensions underlying impression formation (Oosterhof & Todorov, 2008; Sutherland et al., 2013). Specific traits included (a) trustworthiness (aggressive, trustworthy, warm, friendly, likable, and caring); (b) dominance (dominant, competent, physically strong, assertive, intelligent, and smart); and (c) youthful/attractive (healthy, attractive, and youthful). We drew from the literature to determine which traits loaded on which dimensions (Oosterhof & Todorov, 2008; Sutherland et al., 2013), identical to previous work (Hehman et al., 2017). See the online supplementary materials for a correlation matrix of the relationships between all traits. Across dimensions,

[1] We caution researchers that bootstrapped confidence intervals around a correlation r will only have a symmetrical distribution when $r = 0$ but become progressively less symmetrical as r approaches ± 1. And we observe this in the present data, as across all boostrapped ICCs, we find a very small but positive average amount of bias (nonsymmetrical positive bias = .00978). This is a small issue here due to the smaller ICCs but would become a larger issue as ICCs approach 1.

55,352 ratings, 690 participants, and 645 stimuli formed the trustworthiness dataset (78.1% from controlled stimuli sets); 74,158 ratings, 969 participants, and 733 stimuli formed dominance (57.6% controlled stimuli); and 58,962 ratings, 667 participants, and 654 stimuli formed youthful/attractive (44.9% controlled stimuli). These stimuli were unique identities of faces within each dimension but were not unique across dimensions.

Study 1B: Confirmatory Dataset. To confirm any results from this exploratory study, we also collected an entirely new dataset that surpasses the sample size of the original for purposes of confirmation. Data for the replication study were collected on MTurk (total $n = 4{,}864$), in a sample roughly double the size of the exploratory dataset. Data were cleaned in a manner consistent with Study 1A, in accordance with our preregistered data cleaning process based on response time and frequency of repeated ratings [https://osf.io/65tpb/].[2] Participants who had relevant demographic information (i.e., race, gender) were included. These criteria resulted in 219,658 ratings of trait impressions across 2,984 participants and 873 stimuli. Participants were aged 18 to 85 ($M_{age} = 35.47$, $SD_{age} = 11.55$, 66.4% female) and were 69.4% non-Hispanic White, 10.5% Black, 6.1% Asian, and 14.0% other ethnic minorities. Participant ratings of trait impressions included: aggressive ($n = 16{,}171$), assertive ($n = 17{,}929$), attractive ($n = 13{,}360$), caring ($n = 13{,}947$), competent ($n = 15{,}739$), dominant ($n = 15{,}710$), friendly ($n = 13{,}800$), healthy ($n = 15{,}642$), intelligent ($n = 16{,}775$), physically strong ($n = 15{,}698$), smart ($n = 18{,}623$), trustworthy ($n = 14{,}756$), warm ($n = 14{,}856$), and youthful ($n = 16{,}652$).

Across dimensions, 73,530 ratings, 1,034 participants, and 773 stimuli formed trustworthiness; 100,474 ratings, 1,296 participants, and 873 stimuli formed dominance; and 45,654 ratings, 654 participants, and 773 stimuli formed youthful/attractive dimensions, created in a manner identical to Study 1A. All of the stimuli for the confirmatory dataset were sourced from controlled databases.

Stability of estimates. Throughout, we partitioned the data in various ways to allow for different statistical comparisons (e.g., majority perceivers viewing male minority targets). Partitioning by a greater number of factors correspondingly decreases the sample size in each model, creating the possibility that estimates of ICCs from smaller samples might be highly unstable and therefore unlikely to generalize to other samples. To ensure our ICC estimates were stable, we adapted a sequential sampling "corridor of stability" approach (Schönbrodt & Perugini, 2013) to examine how many ratings were necessary before estimated ICCs were stable. We randomly sampled (with replacement) observations from our data and calculated ICCs from those observations for each sample. We did so with an increasing sample size, until we reached the "point of stability"—defined as the sample size at which 95% of the estimated ICCs fell within an acceptable corridor and did not again exceed the boundaries of this corridor, consistent with previous research (Schönbrodt & Perugini, 2013). R code is available at [hehmanlab.org/toolbox]. We report this approach fully in the online supplementary materials.

Across all dimensions, ICCs were stable by 6,000 observations, even using a conservative corridor of stability. Therefore, all ICCs reported in the present work are stable in accordance with this approach, with the exception of when minority perceivers evaluate other minority targets. Because readers may be interested in ICCs from several other models that were below the threshold needed for stability, we report all estimates but urge caution when considering the generalizability of these estimates.

Stimuli

An important consideration when estimating perceiver-versus-target variance is the intrinsic variance of different sets of stimuli (Hehman et al., 2017; Hönekopp, 2006). For example, participants rating a trait with low variance across a highly selected target sample (e.g., attractiveness of fashion models) results in artificially inflated perceiver-ICC for that trait (Hönekopp, 2006). In Study 1A, we followed recommendations from previous research on data-driven approaches with heterogeneous ambient stimuli, and incorporated large stimuli samples that were diverse in their presentation and in their representation of different traits (Burton, Kramer, Ritchie, & Jenkins, 2016; Hehman et al., 2017; Jenkins, White, Van Montfort, & Burton, 2011; Sutherland et al., 2013). In Study 1A, the stimuli included are heterogeneous, broadly representative of facial variation, and curated from a wide variety of sources, ranging from tightly controlled faces with hair and background features removed to real, naturalistic photos taken by professional photographers or the targets themselves. A number of stimuli also originate from standardized databases (e.g., Chicago Face Database; Ma, Correll, & Wittenbrink, 2015), which offer a balance between these extremes.

To ensure that results were not an artifact of these noisier ambient stimuli, in the Study 1B confirmatory dataset, stimuli were sourced entirely from standardized databases. These databases include the Chicago Face Database (Ma et al., 2015), the MR2 (Strohminger et al., 2016), the Radboud Faces Database (Langner et al., 2010), the NimStim Face Stimulus Set (Tottenham et al., 2009), the Center for Vital Longevity Face Database (Minear & Park, 2004), the OSLO Face Database (Chelnokova et al., 2014), the Eberhardt Face Database (Eberhardt, n.d.), the Face Research Lab London Set (DeBruine & Jones, 2017), and the CUHK Face Sketch database (Wang & Tang, 2009). In total, there were 596 new stimuli introduced in the confirmatory dataset. We wanted to ensure that our results were generalizable to other stimuli and were not a result of a specific stimulus set. However, because it was difficult to find an appropriate amount of stimuli for certain social categories (e.g., Asian women), there was necessarily some stimuli overlap with the exploratory dataset.

Overall, these stimuli are representative of those used in impression formation research, providing a medley of tightly controlled stimuli and natural images of individuals that one might encounter in real life (e.g., politicians, undergraduate volunteers, computer-generated models, mugshots, selfies, Facebook profiles, fraternities, standardized databases; see Supplementary Figure 1). This heterogeneity of stimuli helps to ensure the generalizability of our estimates to other samples (Burton et al., 2016; Jenkins et al., 2011).

[2] Because Study 1A data was aggregated from other studies collected over many years, and this preregistration completed only relatively recently, not all data cleaning in Study 1A can be said to have been preregistered. But data across both studies was cleaned in the same manner.

Study 1: Social Perceptions Across Race, Gender, and Majority-Minority Status

We first examined the roles of perceiver- and target-level factors in evaluating targets across racial and gender group boundaries. We partitioned the data according to (a) target gender, (b) target racial majority/minority status, and (c) perceiver racial majority/minority status. In a subsequent analysis, we partitioned our data according to target and perceiver race (i.e., White, Black, Asian) instead of racial majority/minority status to explore differences among the racial minority groups.

We did not partition our data according to perceiver gender for several reasons. First, sample sizes became very small when we partitioned by both the gender of the perceiver and the target, below the threshold of *N* ratings required for stability according to our resampling approach (see the online supplementary materials). Second, when we estimated ICCs from subsets in which we did additionally partition by perceiver gender, the results did not reveal any meaningful systematic differences between male and female perceivers. Therefore, we collapsed across this factor to increase sample sizes and stability.

ICC Estimates From Null Models

Overview. Given the numerous categories created by the race, majority/minority status, and gender of perceivers and targets, many statistical comparisons are possible, increasing the likelihood that our results capitalize on chance. Therefore, our approach was to initially test select patterns or interesting comparisons in an exploratory manner and reexamine them with a confirmatory dataset. We first present a descriptive, bird's-eye view of all ICC estimates from the exploratory (see Figure 1) and confirmatory (see Figure 2) data sets. On each dimension, perceivers (of racial majority or minority status) evaluate female or male targets (of racial majority or minority status). See Table 1 for a full report of all ICCs and their intervals from the exploratory and confirmatory data sets.

Beyond racial majority/minority status, we also estimated ICCs for own-race and other-race impressions more specifically. That is, the racial minority group includes both Black and Asian perceivers and Black and Asian targets and does not differentiate between the two. In Figure 3, we present perceivers' (White, Black, or Asian) impressions of female or male White, Black, or Asian targets. However, because certain impressions (e.g., Asian viewing Black) had small sample sizes and ICCs were below our stability thresholds, these estimates are instead presented in online supplementary materials. See Table 2 for a report of these ICCs and their intervals from the exploratory and confirmatory data sets.

Perceiver versus target contributions. First, we wanted to explore the relative contributions of perceiver versus target characteristics more generally, in line with previous work (Hehman et al., 2017). We averaged the CIs of all perceiver-ICCs and compared that to the averaged CIs of all target-ICCs. Results indicated that across all dimensions, impressions were more perceiver- than target-driven (Figure 4A). On average, between-perceiver ICCs (95% CI [.193, .264]) were larger than between-target ICCs (95% CI [.112, .188]). This indicates that between-perceiver differences comprised a larger percent of the total variance than between-target differences.

Further, to examine whether the relative breakdown of perceiver- versus target-ICCs varied across dimensions, we also averaged and compared perceiver- and target-ICCs on each dimension. Results revealed that ratings on the youthful/attractiveness dimension (95% CI [.171, .270]) were more target-driven than ratings on both the dominance and trustworthiness dimensions (95% CI [.082, .148]), consistent with previous work (Hehman et al., 2017; Hönekopp, 2006).

Figure 1. Relative contributions of perceiver-level, target-level, and residual variance to impressions on (a) trustworthiness, (b) dominance, and (c) youthful/attractiveness dimensions for male and female targets according to racial group status. Results from Study 1: Exploratory dataset. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

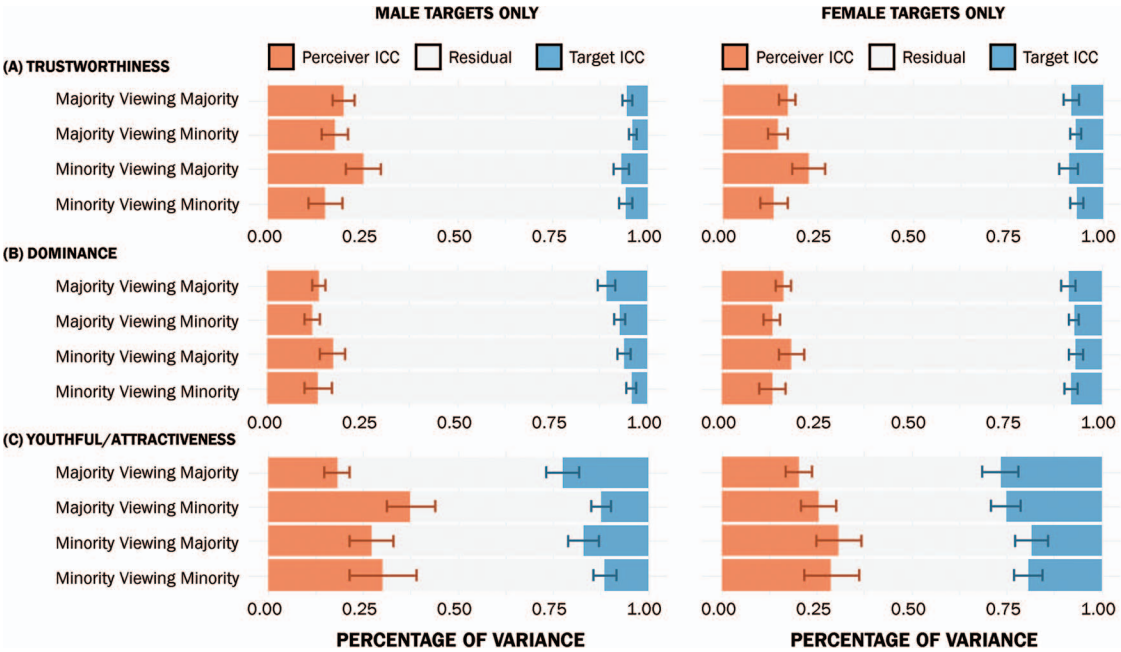

Figure 2. Relative contributions of perceiver-, target-, and residual variance to impressions on (a) trustworthiness, (b) dominance, and (c) youthful/attractiveness dimensions for male and female targets according to racial group status. Results from Study 1: Confirmatory dataset. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

These patterns were replicated in the confirmatory dataset. Larger perceiver-ICCs (95% CI [.162, .241]) than target-ICCs (95% CI [.086, .116]) were present across all dimensions. Furthermore, ratings on youthful/attractiveness were again more target-driven than ratings on trustworthiness and dominance. This effect was larger in the confirmatory dataset—target-ICCs were greater for youthful-attractiveness (95% CI [.152, .229]) compared to both trustworthiness and dominance (95% CI [.053, .087]). Unlike the exploratory dataset, however, this discrepancy also extended to perceiver-ICCs, which were larger for youthful/attractiveness (95% CI [.217, .331]) than for the other two dimensions (95% CI [.134, .196]).

Target gender. Our primary theoretical interest was in examining whether perceptions might differ across group boundaries. We found support for this possibility in multiple ways. First, to explore gender differences, we averaged ICC intervals for all impressions of female targets (e.g., minority viewing majority, majority viewing minority) and compared that to the averaged intervals of all male targets reported in Table 1. Systematic differences emerged in the extent to which perceiver and target characteristics contributed to impressions for male and female targets (Figure 4B). Across all dimensions, target-ICCs were larger for ratings of women (95% CI [.132, .236]) than for men (95% CI [.091, .140]). This pattern replicated in the confirmatory dataset, with larger target-ICCs for ratings of female (95% CI [.100, .152]) compared to male targets (95% CI [.071, .116]). These findings suggests that appearance drives impressions of women to a greater extent than it does for men across all dimensions.

Majority versus minority racial group status. We next examined the effect of racial majority-minority status. Specifically, we examined whether the way in which majorities formed impressions of other majorities was different than how they formed impressions of minorities, and whether the way in which minorities formed impressions of other minorities differed from how they formed impressions of majority group members (Figure 4C). In the exploratory analysis, there was some evidence that perceiver characteristics drove ratings to a greater extent when minority-group perceivers were evaluating majority-group targets (95% CI [.208, .322]) compared to the reverse (95% CI [.137, .212]). Follow-up analyses revealed that this effect was specific to minority perceivers' ratings of majority male targets being particularly perceiver-driven (95% CI [.267, .374]) relative to the reverse (95% CI [.190, .244]) but did not generalize to female targets.

Yet this pattern of results did not replicate in the confirmatory dataset. Specifically, there were no differences in perceiver-ICCs between minority perceivers' ratings of majority male targets (95% CI [.188, .279]) and the reverse (95% CI [.146, .200]). Thus, although minority ratings of majority-male targets were more perceiver-driven in our exploratory dataset, this pattern did not replicate in our confirmatory dataset.

Partitioning Asian and Black categories. Our theoretical framework highlights the potential for stereotype information to act as a template for perceivers, differentially shaping the process of impression formation for cross-group impressions. In the analyses above, we partitioned the data by racial majority and minority group status for theoretical reasons. Yet the minority groups subsumed under this label (e.g., Black, Asian) have very different stereotypical associations, and very different relationships with majority group Whites (Fiske et al., 2002; Johnson et al., 2012; Verkuyten, 2005). These different relationships and stereotypic

Table 1
Intraclass Correlations (ICCs) and 95% Confidence Intervals (CIs) by Target Gender and Racial Group Status

	Study 1: Exploratory				Study 2: Confirmatory			
Target	Perceiver ICC	[95% CI]	Target ICC	[95% CI]	Perceiver ICC	[95% CI]	Target ICC	[95% CI]
Male targets								
Trustworthiness								
Majority viewing majority	.244	[.215, .273]	.149	[.126, .172]	.201	[.173, .229]	.053	[.039, .067]
Majority viewing minority	.218	[.189, .249]	.069	[.050, .089]	.176	[.143, .211]	.038	[.028, .048]
Minority viewing majority	.343	[.288, .400]	.576	[.037, .078]	.252	[.207, .298]	.069	[.049, .089]
Minority viewing minority	.240	[.181, .301]	.038	[.023, .053]	.152	[.107, .198]	.057	[.039, .075]
Dominance								
Majority viewing majority	.190	[.168, .212]	.122	[.105, .140]	.137	[.118, .155]	.106	[.083, .129]
Majority viewing minority	.161	[.138, .184]	.085	[.064, .106]	.119	[.098, .140]	.072	[.058, .087]
Minority viewing majority	.308	[.261, .357]	.060	[.047, .072]	.173	[.140, .208]	.059	[.042, .076]
Minority viewing minority	.184	[.139, .230]	.044	[.028, .060]	.135	[.100, .171]	.040	[.028, .053]
Youthful/attractiveness								
Majority viewing majority	.218	[.188, .248]	.170	[.146, .193]	.182	[.148, .216]	.223	[.180, .267]
Majority viewing minority	.221	[.185, .258]	.249	[.200, .299]	.374	[.314, .440]	.124	[.096, .150]
Minority viewing majority	.308	[.253, .365]	.140	[.115, .165]	.272	[.216, .331]	.171	[.130, .211]
Minority viewing minority	.164	[.110, .219]	.205	[.155, .257]	.300	[.215, .391]	.114	[.082, .144]
Female targets								
Trustworthiness								
Majority viewing majority	.218	[.171, .266]	.193	[.133, .256]	.170	[.148, .193]	.083	[.063, .103]
Majority viewing minority	.101	[.063, .140]	.238	[.182, .297]	.145	[.120, .170]	.070	[.056, .085]
Minority viewing majority	.205	[.142, .270]	.166	[.109, .225]	.227	[.184, .271]	.090	[.066, .114]
Minority viewing minority	.185	[.069, .311]	.166	[.106, .226]	.134	[.098, .171]	.069	[.051, .086]
Dominance								
Majority viewing majority	.191	[.157, .227]	.077	[.054, .101]	.163	[.143, .184]	.087	[.068, .107]
Majority viewing minority	.121	[.081, .160]	.129	[.093, .166]	.133	[.112, .153]	.072	[.059, .085]
Minority viewing majority	.218	[.159, .279]	.060	[.035, .085]	.184	[.151, .219]	.068	[.049, .087]
Minority viewing minority	.128	[.049, .211]	.176	[.121, .234]	.134	[.100, .168]	.081	[.063, .099]
Youthful/attractiveness								
Majority viewing majority	.184	[.143, .224]	.340	[.266, .419]	.202	[.168, .237]	.265	[.218, .314]
Majority viewing minority	.223	[.168, .279]	.214	[.162, .267]	.255	[.208, .302]	.251	[.214, .290]
Minority viewing majority	.202	[.148, .256]	.301	[.232, .372]	.309	[.250, .368]	.185	[.142, .227]
Minority viewing minority	.304	[.193, .423]	.138	[.089, .188]	.288	[.217, .362]	.193	[.154, .231]

associations might influence the process of impression formation for each group. To investigate to what extent the process of impression formation might vary as a function of these factors, here we focused on the two racial minority groups present in the data: Blacks and Asians. We did not have any specific a priori hypotheses but report the ICCs for these groups separately in Table 2.

We found some qualitative differences in ICCs across racial groups (see Figure 3), but the 95% CIs for these ICCs were large due to the smaller samples—particularly for Black and Asian perceivers—and frequently overlapped to a large extent. The only significant difference involved Black perceivers' impressions of White male targets on dominance, which was particularly perceiver-driven (perceiver-ICC 95% CI [.326, .498]) compared to White perceivers' impressions of Black male targets (perceiver-ICC 95% CI [.199, .263]). Yet this pattern did not replicate in the confirmatory dataset. No other patterns were consistently present across both the exploratory and confirmatory data sets. Because of the smaller sample of Black and Asian perceivers in these data and their correspondingly large CIs, these results are difficult to interpret. We present these findings in Figure 3 but caution against overinterpreting these estimates. We return to this issue in the General Discussion.

In summary, we found evidence that (a) impressions in general were more perceiver- than target-driven, (b) target appearance drove impressions of youthful/attractiveness more so than other dimensions, (c) target appearance drove impressions of women more so than men, and (d) there were no consistent differences in majority versus minority perceptions. Other comparisons with this data are possible, and for purposes of exhaustive reporting we include some additional (but nonsignificant) analyses in the online supplementary materials.

Conditional Models

Above, we compared ICCs from null models. Here, we turn to a different analytic approach to test the extent to which different factors explain variance in impressions, building conditional models in which the group-level predictors of interest (e.g., perceiver gender, target gender, perceiver racial majority/minority status, and target racial majority/minority status) are included as moderators.

As previously described, these analyses were conducted using a large dataset in which all responses were modeled together. Because of convergence issues, we were unable to build a maximally identified model in which all four moderators (which we included as fixed and random effects) fully interacted with each other at each level of the model. Thus, we followed recommendations for modifying the maximal model (Brauer & Curtin, 2018) until we produced one that converged consistently across different dimen-

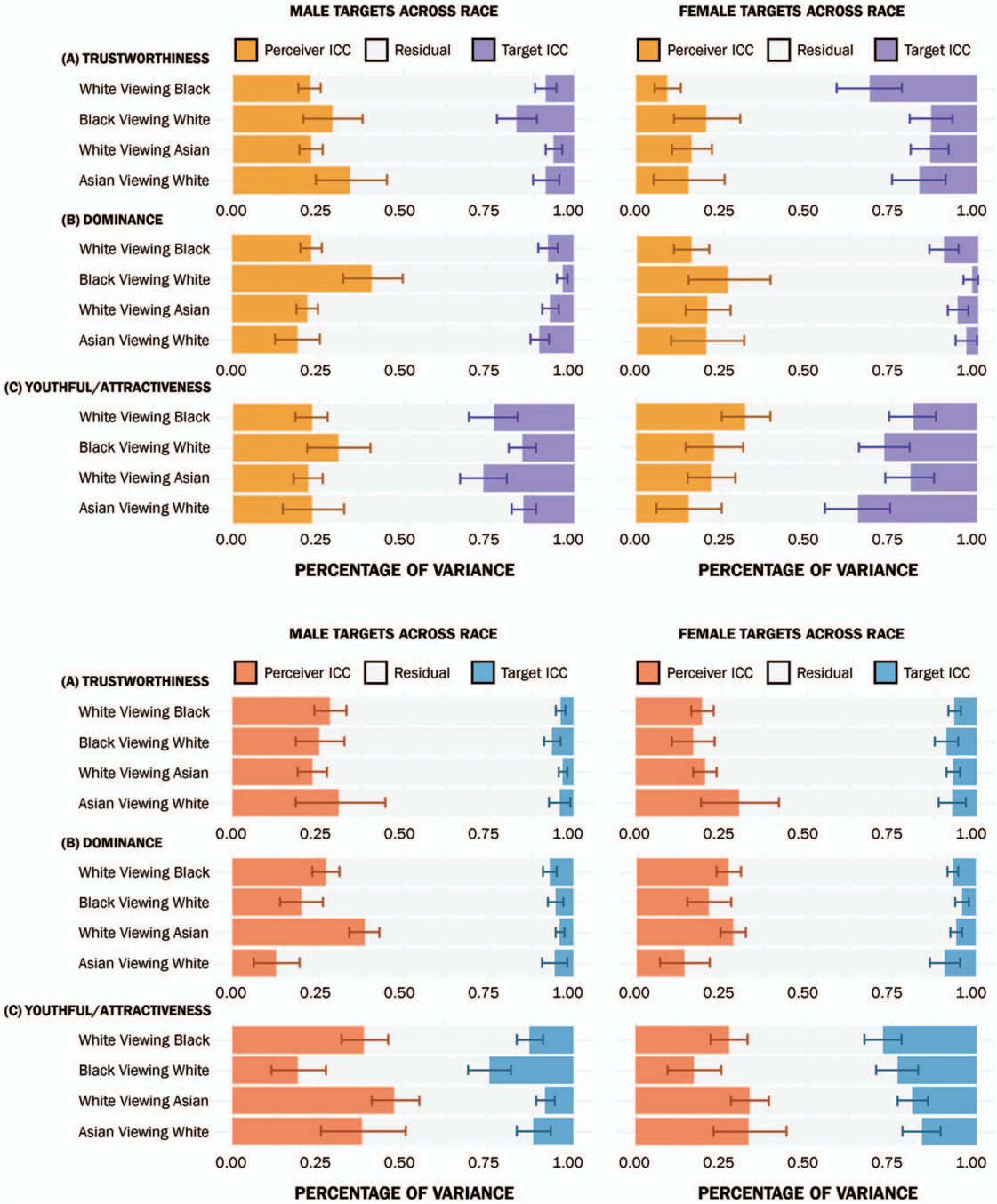

Figure 3. Relative contributions of perceiver-, target-, and residual variance to impressions for male and female targets according to racial group status. Results from Study 1: exploratory and confirmatory data sets. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

sions. See the online supplementary materials for a full description of this process. Moving forward, we refer to this modified maximal model as the maximal model unless otherwise specified.

All fixed and random effect estimates for each predictor are reported in Table 3. Fixed effects represent the mean difference in ratings as a function of our predictors, such as whether the average rating of trustworthiness was different for female versus male targets. Random effects represent the overall variability in these ratings, such as whether the degree to which female targets were rated differently from male targets varied across perceivers. The estimated fixed effects and their significance varies across the exploratory and confirmatory data sets. This is most likely due to the different types of stimuli involved in each dataset. Of primary interest to understanding how gender and

Table 2
Intraclass Correlations (ICCs) and 95% Confidence Intervals (CIs) by Target Gender and Race for White, Black, and Asian Targets

	Study 1: Exploratory				Study 2: Confirmatory			
Target	Perceiver ICC	[95% CI]	Target ICC	[95% CI]	Perceiver ICC	[95% CI]	Target ICC	[95% CI]
Male targets								
Trustworthiness								
White viewing Black	.224	[.192, .258]	.083	[.052, .114]	.287	[.240, .335]	.039	[.024, .053]
Black viewing White	.291	[.206, .381]	.168	[.110, .227]	.249	[.184, .329]	.062	[.038, .087]
White viewing Asian	.228	[.195, .263]	.062	[.036, .085]	.235	[.192, .277]	.032	[.019, .044]
Asian viewing White	.343	[.242, .452]	.083	[.045, .120]	.311	[.186, .448]	.042	[.011, .073]
Dominance								
White viewing Black	.231	[.199, .263]	.076	[.048, .105]	.275	[.235, .314]	.070	[.050, .090]
Black viewing White	.409	[.326, .498]	.034	[.019, .048]	.202	[.140, .266]	.053	[.031, .076]
White viewing Asian	.220	[.189, .251]	.069	[.044, .093]	.387	[.343, .432]	.040	[.027, .053]
Asian viewing White	.190	[.127, .257]	.100	[.073, .127]	.128	[.063, .197]	.055	[.019, .092]
Youthful/attractiveness								
White viewing Black	.231	[.184, .277]	.236	[.167, .309]	.389	[.319, .457]	.131	[.091, .167]
Black viewing White	.307	[.216, .402]	.152	[.112, .192]	.192	[.114, .275]	.248	[.185, .310]
White viewing Asian	.220	[.178, .262]	.265	[.197, .336]	.478	[.408, .547]	.083	[.056, 111]
Asian viewing White	.232	[.147, .324]	.148	[.112, .183]	.378	[.260, .509]	.119	[.068, .168]
Female targets								
Trustworthiness								
White viewing Black	.092	[.053, .131]	.314	[.221, .411]	.195	[.162, .230]	.067	[.048, .085]
Black viewing White	.204	[.111, .306]	.135	[.072, .198]	.168	[.107, .230]	.090	[.057, .124]
White viewing Asian	.164	[.105, .222]	.139	[.084, .196]	.202	[.168, .237]	.069	[.049, .089]
Asian viewing White	.153	[.052, 259]	.169	[.091, .249]	.302	[.192, .420]	.073	[.034, .112]
Dominance								
White viewing Black	.160	[.108, .212]	.102	[.059, .144]	.273	[.236, .306]	.069	[.051, .085]
Black viewing White	.267	[.152, .391]	.018	[−.009, .043]	.209	[.151, .280]	.043	[.020, .062]
White viewing Asian	.208	[.142, .275]	.061	[.031, .090]	.286	[.246, .321]	.057	[.042, .075]
Asian viewing White	.204	[.101, .313]	.037	[.005, .068]	.142	[.070, .217]	.093	[.048, .137]
Youthful/attractiveness								
White viewing Black	.321	[.250, .394]	.188	[.121, .257]	.276	[.220, .328]	.275	[.221, .329]
Black viewing White	.230	[.146, .315]	.271	[.198, .345]	.171	[.094, .251]	.234	[.171, .296]
White viewing Asian	.220	[.150, .290]	.195	[.126, .268]	.334	[.279, .392]	.189	[.145, .234]
Asian viewing White	.157	[.060, .251]	.356	[.256, .446]	.331	[.228, .444]	.162	[.106, .217]

racial boundaries influence the variability of these ratings is the estimated random effects and variance explained by the predictors. Accordingly, we focus on these analyses in the subsequent sections.

To determine the percentage of variance uniquely explained by each predictor, we built additional three-predictor models in which a single predictor (e.g., target gender) was removed from the maximal four-predictor model and we examined the change in variance explained to determine how much each variable contributed above and beyond all the others. Because different models were necessarily created for each dimension, results from both exploratory and confirmatory data sets are organized by dimension below.

Trustworthiness. First examining the trustworthiness dimension, the maximal model with all four predictors explained 37.5% of the between-perceiver variance and 15.4% of the between-target variance in the exploratory dataset. Target gender uniquely explained 27.5% of the between-perceiver variance and target racial majority/minority status explained 19.5%. Perceiver gender explained 4.6% of between-target variance, although perceiver racial majority/minority status explained 7.8%, beyond each other respectively.

We sought to confirm these results with the confirmatory dataset. On the trustworthiness dimension, the full model with all predictors explained 66.8% of the between-perceiver variance and only 1.3% of the between-target variance. Compared to the exploratory dataset, these predictors explained more between-perceiver variance and less between-target variance, which corresponds with the results from our ICC comparisons—perceiver-ICCs were overall larger than target-ICCs for both data sets, and this difference was more pronounced in the confirmatory dataset. Target gender uniquely explained 49.7% of the between-perceiver variance, whereas target racial majority/minority status explained 9.1%, beyond one another respectively. Perceiver gender explained 7.9% of the between-target variance, and perceiver racial majority/minority status explained 13.2%. Overall, on the trustworthiness dimension, target gender explained the most variance out of all four predictors, and this pattern replicated across data sets.

Dominance. For dominance, the maximal model with all predictors explained 35.8% of between-perceiver variance and 5.4% of between-target variance in the exploratory dataset. Target gender explained 28.0% and target racial group status explained 17.0% of the between-perceiver variance. Perceiver racial group status explained 5.4% of the between-target variance, although perceiver gender did not explain any unique variance on the dominance dimension.

In the confirmatory dataset, the maximal model with all predictors explained 99.2% of the between-perceiver variance and only 3.5% of the between-target variance. In line with the ICC com-

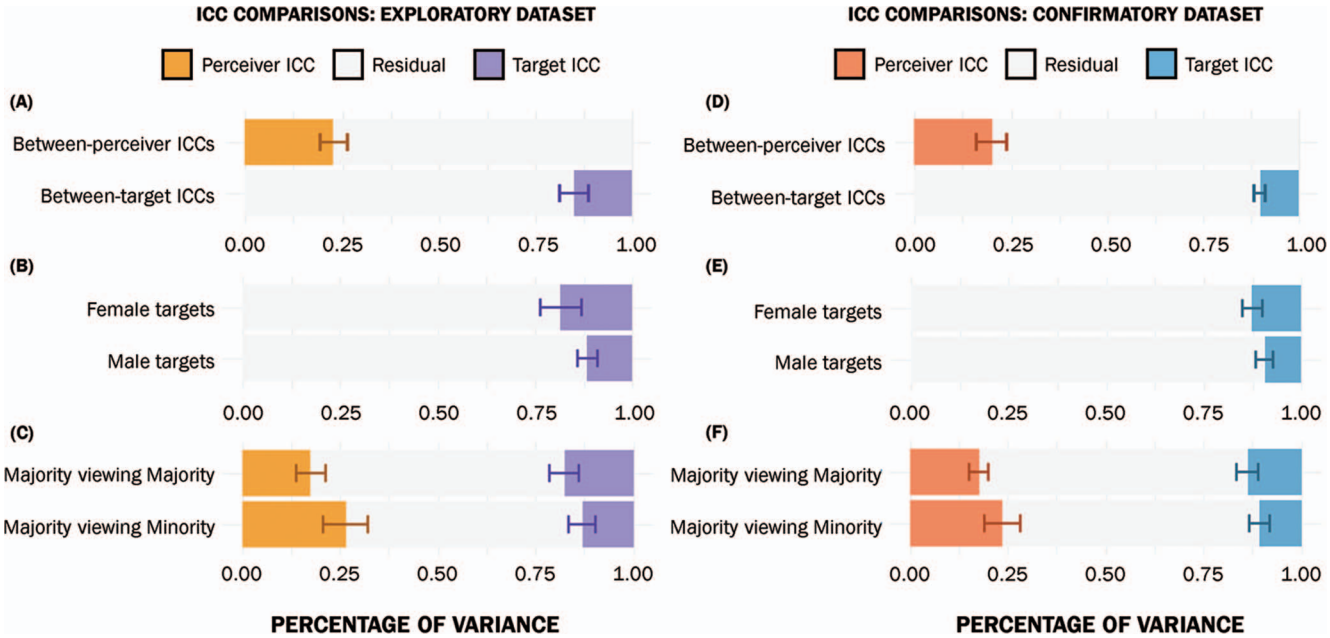

Figure 4. Relative contributions of perceiver characteristics, target characteristics, and the residual across selected comparisons averaged across groups. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

parisons, there was a large discrepancy between perceiver- and target-ICCs. Target gender was again the most important contributor, uniquely explaining 88.0% of the between-perceiver variance, whereas target racial majority/minority status explained 41.9%. Neither perceiver racial majority/minority status nor perceiver gender explained any variance between targets.

Youthful/attractiveness. Finally, for youthful/attractiveness, the full model explained 14.7% of the between-perceiver variance and 4.5% of the between-target variance in the exploratory dataset. Target gender explained almost all of the between-perceiver variance (13.4%). Target racial group status, perceiver racial group status, and perceiver gender each explained less than 1.1% of variance on this dimension.

These findings replicated in the confirmatory dataset: the full model explained 62.3% of between-perceiver variance and 4.3% of between-target variance. Again, target gender explained most of the variance between perceivers (60.5%), whereas target racial majority/minority status uniquely explained 11.9% of this variance. Perceiver gender explained 3.8% of the variance between targets, whereas perceiver racial majority/minority status did not explain any variance.

Discussion

Across all dimensions, the between-perceiver variance explained by these maximally specified models was much higher in the confirmatory dataset, ranging from 62.3% to 99.2% (vs. 14.7% to 37.5% in the exploratory dataset), though only for variance between perceivers. In contrast, the explained between-target variance was much lower in the confirmatory dataset, ranging from 1.3% to 4.3% (vs. 4.5% to 15.4% in the exploratory dataset). Yet the predictors were consistent across data sets in terms of which explained more or less variance. Overall, analyses of these maximal models indicate that our predictors of interest (i.e., gender and racial group status) explain more variance between perceivers than between targets. These findings correspond with the results of our ICC comparisons from the previous section, which revealed larger perceiver-ICCs and smaller target-ICCs overall—indicating that perceiver characteristics are contributing more to impressions than are target characteristics.

Of all the predictors, target gender explained the most between-perceiver variance on all dimensions, followed by target racial majority/minority status. This pattern replicated across data sets. Thus, of the four predictors in our model, between-perceiver variance in impressions can largely be attributed to the gender of the targets, with some additional variance being explained by whether targets belonged to a racial majority or minority group.

In contrast, the between-target variance in ratings can largely be explained by whether the perceiver belonged to a racial majority or minority group. On the dominance dimension in particular, perceiver racial majority/minority status explained the most between-target variance. Finally, perceiver gender explained the least variance in ICCs on all dimensions.

In summary, analyses of these conditional models reveal that the gender and racial majority/minority status of perceivers and targets contribute to impressions in different ways: (a) these predictors explain more variance between perceivers than between targets, (b) target gender consistently explains the most variance between perceivers, and (c) the racial majority/minority status of targets and perceivers also explains some variance.

Study 2: Minimal Groups

In Study 1, variation in perceiver and target contributions to impressions across racial and gender boundaries is evident. How-

Table 3
Fixed- and Random-Effect Estimates for Perceiver/target Gender and Racial Group Status From Conditional Models for the Trustworthiness, Dominance, and Youthful/Attractiveness Dimensions

	Exploratory dataset			Confirmatory dataset		
	β	*SE*	*p*	β	*SE*	*p*
Fixed effects						
Trustworthiness						
PerRace	−.021	.034	.54	−.004	.023	.86
TarRace	−.037	.036	.30	.030	.023	.17
PerGen	.025	.034	.46	−.046	.023	**<.05***
TarGen	.031	.039	.43	.048	.028	.08
PerGen × TarRace	.055	.024	**<.05***	.023	.017	.17
PerGen × TarGen	.061	.028	**<.05***	.033	.023	.15
PerGen × PerRace × TarRace	.063	.024	**<.01***	.024	.017	.15
PerGen × TarGen × PerRace	.069	.028	**<.05***	.001	.023	.99
Four-way interaction	.050	.023	**<.05***	.023	.017	.17
Dominance						
PerRace	−.023	.028	.40	.017	.018	.36
TarRace	−.012	.029	.69	−.017	.020	.93
PerGen	−.052	.028	.75	.017	.019	.37
TarGen	−.012	.034	.12	−.025	.025	.92
PerGen × TarGen	.051	.025	**<.05***	.024	.019	.20
PerRace × TarRace × TarGen	.021	.018	.25	.031	.012	**<.01***
Youthful/attractiveness						
PerRace	.003	.003	.28	.012	.040	.77
TarRace	.003	.004	.44	−.085	.036	**<.05***
PerGen	.000	.004	.99	−.014	.040	.72
TarGen	.004	.003	.23	.033	.050	.52
TarGen × TarRace	.015	.004	**<.001***	−.059	.036	.09
PerGen × TarRace	.005	.002	**<.01***	.047	.023	**<.05***
TarGen × PerGen	.009	.002	**<.01***	−.015	.040	.71
PerRace × TarRace × PerGen	.004	.002	**<.05***	.016	.023	.47
TarRace × TarGen × PerGen	.004	.002	**<.05***	.014	.023	.55
Random effects	τ_{c00}	τ_{b00}	σ^2	τ_{c00}	τ_{b00}	σ^2
Trustworthiness	.259	.229	1.293	.117	.146	1.697
TarRace	.069	—	—	.103	—	—
TarGen	.083	—	—	.178	—	—
PerRace	—	.004	—	—	.003	—
PerGen	—	.006	—	—	.003	—
Dominance	.231	.200	1.519	.002	.189	1.907
TarRace	.063	—	—	.067	—	—
TarGen	.104	—	—	.278	—	—
PerRace	—	.001	—	—	.002	—
PerGen	—	.000	—	—	.003	—
Youthful/attractiveness	.436	.445	1.285	.255	.504	1.385
TarRace	.044	—	—	.098	—	—
TarGen	.062	—	—	.414	—	—
PerRace	—	.008	—	—	.006	—
PerGen	—	.002	—	—	.003	—

Note. Each conditional model was a multilevel model, cross classified at the perceiver (PerceiverID) and target (StimID) level. The R code for each model is as follows, where the rating on each dimension serves as the DV: lmer(Rating ~ TarRace * TarGen * PerRace * PerGen + (TarRace + TarGen | PerceiverID) + (PerRace + PerGen | StimID)). There were four contrast-coded predictors in the model: TarRace = target racial majority/minority status, PerRace = perceiver racial majority/minority status, TarGen = target gender, and ParGen = perceiver gender. β = estimate of fixed effects; *SE* = standard error; *p* = *p*-value. Significant *p* values are bolded with an asterisk. For estimates of random effects, τ_{c00} = between-perceiver variance; τ_{b00}= between-target variance; σ^2 = residual variance.

ever, multiple factors associated with these salient social categories may elicit these apparent differences. We have hypothesized that racial and gender stereotypes provide templates, acting as preexisting information that perceivers may draw upon (thus using less of the actual appearance of the target) to form impressions. An alternative hypothesis is that these differences do not rely upon stereotypes of the specific groups. Rather, the mere categorization of targets as members of an outgroup elicits differences in how targets of that group are processed (Brewer, 1979; Cikara & Van Bavel, 2014; Neuberg & Fiske, 1987; Van Bavel et al., 2008).

To test this possibility, in a third sample collected for this purpose we induced minimal group membership (Ratner & Amodio, 2013; Tajfel et al., 1971) and had participants form impressions of their own and another group in a manner identical to the data above. Participants were led to believe that they belonged to a particular group. Because these groups were fictional, partici-

pants had no preexisting stereotypes to draw upon when forming impressions of members of these groups. In the absence of stereotype information, if perceiver- and target-ICCs were different across group boundaries, then it would suggest that mere categorization and group identification, rather than racial and gender stereotypes, are driving differences between groups in our original analyses. In contrast, if differences between the perceiver- and target-ICCs of minimal ingroup and outgroup targets does not differ, this would provide some evidence that differences emerge at least partially as a result of stereotypic information specifically associated with the groups involved.

Methods

Participants and design. Participants (n = 323) were recruited from Mechanical Turk in a 2-level (group: overestimator, underestimator) between-subjects design. Fifty-three participants were removed in accordance with our preregistered cleaning process. Because stimuli were White, analyses were restricted to Non-Hispanic White participants only (n = 196) to ensure racial stereotypes were not involved. As a manipulation check, participants failing to correctly identify to which minimal group they had been assigned at the end of the task (n = 22) were additionally removed, leaving 174 for analysis (M_{age} = 39.44, SD_{age} = 13.29, 72% female).

Procedure. Minimal group assignment followed classic procedures (Deegan, Hehman, Gaertner, & Dovidio, 2015; Diehl, 1990; Gramzow, Gaertner, & Sedikides, 2001). Participants were presented with photos of randomized object assortments (e.g., flock of birds, jar of jellybeans). Participants estimated the number of units in each photo using a slider from 1 to 1,000, and after 14 trials, received randomized feedback identifying them as an "overestimator" or "underestimator" (Tajfel et al., 1971) and told that overestimators and underestimators tended to share similar cognitive styles. Each participant then proceeded to rate 80 faces. Faces were randomly labeled as overestimators or understimators by participant. Ratings were made on one of six randomly assigned traits that load most strongly on their respective dimensions (friendly, trustworthy, strong, dominant, attractive, youthful; Sutherland et al., 2013; Todorov et al., 2008), using a 7-point Likert scale ranging from 1 (*not at all*) to 7 (*very much*). Finally, participants were asked to identify whether they were an overestimator or underestimator as a manipulation check.

Stimuli. Participants rated facial stimuli from the Chicago Face Database (Ma et al., 2015), consisting of standardized color photographs of 40 White male and 40 White female faces with neutral expressions. Each face was presented with a label of "overestimator" or "underestimator" in large, 18-pt, bolded font, centered above the photo. The order in which faces were presented and their accompanying labels were randomized by participant.

Results

Analyses included 13,861 ratings of trait impressions across 174 participants and 80 stimuli. We calculated perceiver- and target-ICCs for impressions on the three dimensions for both minimal own- and other-group targets in the same manner as Study 1. We collected ratings on only six traits, and so unlike previous analyses, here dimensions reflect averaged composites of the two traits loading on each dimension.

Across all dimensions, there were no significant differences in perceiver-level variance when viewing minimal ingroup (perceiver-ICC 95% CI [.165, .332]) and outgroup (perceiver-ICC 95% CI [.152, .314]) targets, nor in target-level variance (ingroup target-ICC 95% CI [.115, .226]; outgroup target-ICC 95% CI [.108, .213]). Otherwise, some patterns evident in the data were consistent with the results above. For instance, perceiver-level factors contributed more to impressions than target factors, and ratings along the youthful/attractiveness dimension were the most target-driven (see Figure 5).

Importantly, these results provide greater insight into the results from Study 1. Evidence indicates that the extent to which perceiver- and target-characteristics contribute to impressions is not driven by mere categorization or group identification alone. Rather, it appears differences may have emerged across racial and gender groups because of preexisting stereotypic information about these groups. In the absence of this information, differences in the process of impression formation across groups no longer appear.

General Discussion

Social impressions arise from both perceiver and target characteristics. Although previous research has demonstrated that perceiver and target characteristics vary in the extent to which they contribute to ratings on various trait dimensions (Hehman et al., 2017; Hönekopp, 2006), whether these contributions differ for cross-group impressions remains unknown. This interplay between perceiver and target characteristics is at the theoretical heart of much social–cognitive and intergroup research, yet how much perceiver and target characteristics uniquely contribute to impressions across group boundaries has yet to be quantified. Further, researchers not modeling these potential differences are making a functional assumption that the process is the same across racial and gender group boundaries, an assumption that is generally misaligned with theory (Freeman et al., 2010; West, 2011). Given

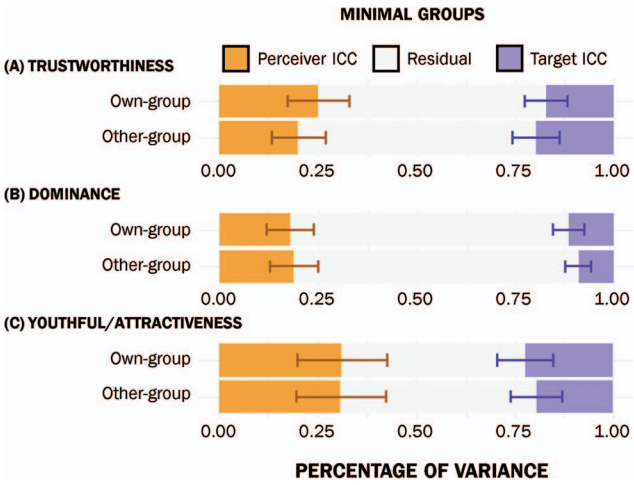

Figure 5. Relative contributions of perceiver characteristics, target characteristics, and the residual across all dimensions for minimal groups. Error bars represent 95% confidence intervals. See the online article for the color version of this figure.

extant evidence that the race and gender of targets and perceivers can interact to influence trait impressions, quantifying this interaction was the focus of our current work. We generally found that the relative influence of perceiver- versus target-level factors varies across gender but not racial boundaries, with perceiver and target characteristics contributing to impressions of female and male targets differently.

To address these questions, we used cross-classified multilevel models to estimate the unique contributions of perceiver- and target-level factors to ratings on various trait dimensions, examining the interplay between perceiver and target characteristics. Similar models have a rich history in partitioning sources of variance in impressions during interactions (Kenny & Albright, 1987; Kenny, Gomes, & Kowal, 2015; Kenny, Horner, Kashy, & Chu, 1992), and here we bring them to bear in a racial and gendered intergroup context. Critically, we demonstrate that (a) impressions in general were more perceiver- than target-driven, (b) target appearance drove impressions of youthful/attractiveness more so than other dimensions, (c) target appearance drove impressions of women more so than men, (d) there were no consistent differences in other-race versus own-race perceptions, and (e) there were no consistent differences in majority versus minority perceptions.

Although researchers have long recognized that social impressions are influenced by individual characteristics, the present study is the first to quantify the extent to which perceiver and target characteristics contribute to impressions across group boundaries on the major dimensions underlying impression formation. We discuss these results in greater detail below.

Broader Patterns

Though the primary goal of the present research was to examine the potentially differential process of perception across group boundaries, some novel and interesting broader effects emerged that warrant discussion. First, we find that perceivers contribute more to any given impression than do targets. This pattern was perhaps the most consistent across all the data collected in the present research. Although social–cognitive models of person perception have long theorized a role of both the perceiver and target in forming impressions, that the contribution of the perceiver is quantitatively larger than that of the target is informative and may be surprising.

In addition, we find that ratings of youthful/attractiveness are the most target driven of the dimensions. Historically, perceptions of attractiveness have been regarded as comprising both an idiosyncratic component (i.e., "personal taste") and a target-driven component (i.e., "shared taste"; Hönekopp, 2006). Here, impressions of the youthful/attractiveness dimension were more target-driven than the other dimensions, conceptually replicating previous research (Germine et al., 2015; Hehman et al., 2017; Hönekopp, 2006). Importantly, we demonstrate that this pattern does not shift across gender and racial group boundaries.

Perceiver and Target Contributions to Impressions Across Gender and Race

Impressions across gender. Regarding our questions of primary interest, differences in the contributions of target and perceiver to impressions were most evident across gender lines. Results across both exploratory and confirmatory data sets indicate that facial appearance plays a more critical role in guiding impressions of female than male targets. Or alternatively, that the actual appearance of male targets matters less to impressions. More evidence in support of this conclusion comes from analyses of the conditional models, which indicate that among the factors examined in the present research, target gender uniquely explained the most between-perceiver variance in ratings on each dimension.

This conclusion is consistent with past evidence that trait judgments of women are considerably influenced by their facial appearance (Carpinella, Hehman, Freeman, & Johnson, 2015; Hehman, Carpinella, Johnson, Leitner, & Freeman, 2014; Pillemer, Graham, & Burke, 2014). These results thus join a body of research demonstrating that appearance generally plays a larger role in women's than men's impressions and subsequent outcomes, such as in business (Hamermesh & Biddle, 1994; Marlowe, Schneider, & Nelson, 1996; Rule & Ambady, 2009) and politics (Carpinella et al., 2015; Hehman, Carpinella, et al., 2014). Importantly, this result was not specific to attractiveness, but generally was present across all the dimensions underlying person perception. These results provide some insight into why appearance drives the outcomes of women more so than men, specifically indicating that the way in which impressions are formed are different for men and women.

Impressions across majority and minority racial group status. Race is a salient social category, and decades of research have demonstrated how impressions of different targets vary systematically as a function of perceiver and target race (Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fiske et al., 2002; Hugenberg, 2005; Todorov et al., 2015). Yet among the ICCs, no patterns were consistently present in both the exploratory and confirmatory data sets to indicate that the ways in which impressions were formed across racial boundaries are different.

Results from the conditional models were more nuanced. Racial majority or minority status of the target played the largest role. On the trustworthiness and dominance dimensions respectively, the racial majority/minority status of targets uniquely explained 19.5% and 17.0% of the between-perceiver variance in the exploratory dataset, and 9.1% and 41.9% of the between-perceiver variance in the confirmatory dataset. Yet for the youthful/attractiveness dimension, targets' racial majority/minority status mattered to a minimal degree. Similarly, the racial majority/minority status of perceivers did not play a large role in impressions, never explaining more than 8% of the variance on any dimension. Thus, although the race of targets and perceivers do contribute to impressions, primarily on trustworthiness and dominance dimensions, it appears to play a less consistent role than target gender.

Previous research has generally found that other-race targets are processed in a shallower, more categorical manner, relative to greater individuation for own-race targets (Hugenberg et al., 2010; Neuberg & Fiske, 1987), and we originally anticipated differences. There are several possible interpretations for this result. First, of course, is that the way impressions are formed for members of each group is identical. An alternative is that own- and other-race impressions race impressions are driven to the same extent by perceiver and target characteristics, but that the constructs that comprise the degree of contribution are different. Meaning, that the approach adopted here can precisely inform how much per-

ceiver and target level factors contribute to variance in impressions, but it does not provide information as to what psychological factors comprise that variance. Ratings of own-race individuals might be perceiver-driven in some ways (e.g., influenced by distinctive assumed similarity), while impressions of other-race members are perceiver-driven in other ways (e.g., influenced by stereotype activation and perceiver prejudice), even though the degree of perceiver contribution is similar. Thus, although we find that the level of contribution from perceiver characteristics is equal across both types of impressions, the specific factors constituting perceiver-level contributions may vary. Future research adopting the current approach can test this possibility by collecting data on these psychological constructs and examining to what extent they explain variance at different levels of the model.

Another possibility is that varying perceptions across racial group boundaries may occur only as a function of certain perceiver characteristics. In other words, differences across racial lines may primarily emerge in the random effect associated with Perceiver × Target interactions. Both our explanatory and confirmatory data sets had fairly large percentages of residual variance. Given the data structure of the present research, Perceiver × Target effects could not be extricated from this Level 1 error variance (Beretvas, 2008; Hehman et al., 2017; Raudenbush & Bryk, 2002), and effects associated with race may be subsumed within this term. Given the large residuals of these models, and the abundance of social–cognitive literature demonstrating Perceiver × Target effects in forming impressions of own- and other-race targets, we believe this explanation to be most likely. Future researchers interested in modeling these effects should target this interaction, particularly when examining variances across racial boundaries.

The Role of Stereotypes

One possible explanation for any differences in the process of perception across group boundaries is that perceivers are drawing upon stereotypes when forming impressions of different types of targets. Judgments on these dimensions require perceivers to make inferences about a target's character based solely on a picture of their face: perceivers may therefore rely more on their personal knowledge of stereotypes and use categorical information to inform their judgments of targets, as opposed to their actual appearance (Adams et al., 2012; Imhoff, Woelki, Hanke, & Dotsch, 2013).

One alternative hypothesis is that these differences do not rely upon the stereotype content of the specific groups being perceived, but instead arise from a fundamental human tendency to perceive gender or racial outgroup targets differently. Intergroup bias arises under even minimal group conditions, and the minimal group paradigm provides a powerful tool for intergroup research because it controls for preexisting stereotypes and associations (as well as confounding factors such as status and power) and allows participants to be randomly assigned to groups (Tajfel et al., 1971). However, in Study 2 and across all dimensions, there were no significant differences in the relative contributions of perceiver versus target characteristics for ingroup and outgroup perceptions. Because the minimal group manipulation did not elicit any variation in perceiver and target ICCs, we suggest that group membership alone is insufficient for explaining our findings. Instead, stereotype content may meaningfully contribute to differences in how perceivers evaluate targets, drawing upon their personal experiences and stereotypes to make inferences about specific categories of targets.

One limiting possibility is that our minimal group manipulation may have been too weak to elicit these effects. We note, however, that we adopted the same paradigm that previous researchers have used in the past with success (Brewer, 1979; Deegan et al., 2015; Gaertner et al., 1993) and used only participants passing a manipulation check, which mitigates this possibility. Nevertheless, future researchers may want to continue to examine group membership alone as altering the process of impression formation. It is important to note that our results do not imply that there are no differences between mean-level impressions of minimal ingroup and outgroup faces—rather, our results indicate perceiver and target characteristics contribute equally to perceptions across minimal groups, and that the process of perception is similar.

Implications

To summarize, a general finding of the current work is that the extent to which perceiver and target characteristics contribute to the process of impression formation varies systematically across gender and racial boundaries. These results have theoretical and methodological implications. For instance, there are some presumptions that a two-dimensional model of trustworthiness and dominance underlying the way humans form impressions is "universal" or "fundamental" (Abele, Cuddy, Judd, & Yzerbyt, 2008; Fiske et al., 2007). Yet our finding that target appearance matters more for women than men undermines this argument, indicating that the process of impression formation varies. If it varies here, it might also vary elsewhere. This opens the possibility that the process of impression formation is also different for different social groups not examined here, or for different contexts, or for impressions formed under different circumstances. In sum, there may not be a universal model for impression formation, and future research will have to map in what circumstances the process of forming impressions changes or is consistent.

Further, we find that perceivers contribute more to any given impression than does the appearance of targets. We would argue that the large majority of research in impression formation has focused on how characteristics of the target might influence impressions, and accordingly, we now understand this space well. Yet our results indicate that if we wish to model and explain ratings with more precision, identifying characteristics of the perceiver that are major factors in forming impressions will be a necessary and fruitful step.

Finally, it is important to note that while the present research focuses on race and gender in driving explicit impressions, this approach is not limited to these specific characteristics. Similar questions might be examined among different social groups or different contexts. Also worth exploring are implicit impressions of different gender and racial targets, as the extent to which perceiver and target characteristics contribute to such impressions—as well as the factors driving this variance—might dramatically vary. With the statistical approach laid out here, researchers can disentangle the relative contributions of the perceiver and the target in their own lines of work and advance our understanding of how characteristics of perceivers and targets interact to drive social impressions.

Limitations

There are several limitations of the present research. First, although our analyses include a large number of ratings of trait impressions (exploratory dataset $n = 188{,}472$; confirmatory dataset $n = 219{,}658$), ICC comparisons were conducted on various subsets of these data. For some more specific analyses, sample sizes necessarily became smaller (e.g., Black perceivers evaluating other Black targets). Though we report in the main text only analyses which surpass our established threshold, estimates based on larger numbers are correspondingly more stable and more likely to be generalizable. It is important to note, however, that although estimates of perceiver- and target-ICCs are more stable for groups that have a greater number of ratings, the number of ratings is unrelated to the value of the estimate itself. We have additionally provided analyses readers may be interested in that do not meet our established threshold for stability in the online supplementary materials.

Another limitation involves the creation of the trustworthiness, dominance, and youthful/attractiveness dimensions. These dimensions were created by averaging together ratings on multiple traits. To determine which specific traits were averaged to create dimensions we relied heavily on previous research whose explicit goals were to determine which traits were part of which dimensions (Oosterhof & Todorov, 2008; Sutherland et al., 2013). Indeed, these dimensions are generally agreed upon by the majority of the social–cognitive field. Yet which traits load onto which dimension will necessarily influence ICC and variance explained estimates from the current research: to the extent that the trait loadings informed by previous research is incorrect, estimates from the current work will change.

A third limitation concerns the heterogeneity and consistency of the stimuli across data sets. In our exploratory dataset, we opted to use widely heterogeneous stimuli as recommended by other research (Sutherland et al., 2013). These encompass real faces from controlled face databases, uncontrolled "noisy" face images collected online, and computer-generated faces, thus reflecting the type of faces regularly used in social–cognitive research. Some of this heterogeneity may increase perceiver variance, as there is greater variation for individual-differences to connect to and influence impressions. To ensure that the results from our exploratory dataset were not an artifact of these noisier ambient stimuli, all stimuli for the confirmatory data sets were sourced from standardized databases, which mitigates this concern to some extent. Effects consistent across these different types of stimuli increases our confidence in the generality of the results. Furthermore, perceiver variance may also have been increased by subsetting our data by social category when estimating ICCs, as this homogenizes the stimuli on relevant dimensions. In general, to the extent that the stimuli involved in the present research are not representative of the real world, or for any social group, our effects may differ from future research using other stimuli.

Finally, another limitation of the present research lies in the specification of our models. Currently, these models do not have repeated impressions of the same targets, as this data structure of multiple ratings of targets per participant is uncommon in the impression formation literature. Because these repeated ratings are not present, it is impossible to estimate the random effect associated with perceiver × target interactions, as it cannot be extricated from the residual variance. Some of the effects of greatest interest to social–cognitive researchers involve these perceiver × target effects. Given a data structure where each target is rated once by each perceiver, the variance associated with these interactions cannot be uniquely estimated from the residual. Here the residuals are quite large, and it is unknown how much of that residual variance is actually associated with Perceiver × Target effects. In the context of repeated ratings, when we can estimate the interaction variance separately from the residual, we might find that interaction variance to be substantial, and potentially a larger proportion than variance associated with perceiver or target effects. Those interested in examining perceptions across racial group boundaries may specifically wish to target these Perceiver × Target interactions by collecting multiple ratings of each target per participant. Overall, these results lay the foundation for future researchers to use this approach.

Conclusion

In summary, the present work advances the impression formation and intergroup literature by quantifying the extent to which perceiver and target characteristics contribute to impressions across racial and gender boundaries. These results demonstrate that the social categories of both perceiver and target influence how impressions are formed and can inform future work on intergroup perceptions by helping researchers decide which variables (perceiver or target characteristics such as race and gender, or the interaction between them) might interact to produce the largest effects on their outcome of interest. Furthermore, these findings reveal systematic differences in the relative importance of perceiver and target characteristics across these social categories and provide insight into the very nature and process of how we form these impressions.

With cross-classified models, we can partition the total variance in a given impression into variance explained by the perceiver and variance explained by the target. Furthermore, by comparing groups that differ on a specific variable of interest (e.g., female vs. male targets), we can examine the relative contributions of perceiver and target characteristics at the level of that variable. Results tentatively suggest that these systematic differences may arise from gendered stereotype content, rather than from a general effect of content-less intergroup perceptions. Our findings reveal the profound variance that exists in impression formation across groups and demonstrate how an emerging statistical approach can be used to disentangle the unique contributions of perceivers and targets to trait judgments, as well as the interplay between the two.

References

Abele, A. E., Cuddy, A. J. C., Judd, C. M., & Yzerbyt, V. Y. (2008). Fundamental dimensions of social judgment. *European Journal of Social Psychology, 38,* 1063–1065. http://dx.doi.org/10.1002/ejsp.574

Adams, R. B., Jr., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: A mechanism for impression formation? *Cognition and Emotion, 26,* 431–441. http://dx.doi.org/10.1080/02699931.2012.666502

Allport, G. W. (1954). The nature of prejudice. *Canadian Psychology*. Reading, MA: Addison Wesley.

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. *Nature Reviews Neuroscience, 15,* 670–682. http://dx.doi.org/10.1038/nrn3800

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7,* 268–277. http://dx.doi.org/10.1038/nrn1884

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Beretvas, S. N. (2008). Cross-classified random effects models. In A. A. O'Connell & D. B. McCoach (Eds.), *Multilevel modeling of educational data* (pp. 161–198). Charlotte, NC: Information Age Publishing.

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The cross-category effect. *Psychological Science, 18,* 706–712.

Brauer, M., & Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods, 23,* 389–411.

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86,* 307–324. http://dx.doi.org/10.1037/0033-2909.86.2.307

Brewer, M. B. (1988). A dual process model of impression formation. In R. S. Wyer Jr. & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). New York, NY: Psychology Press. http://dx.doi.org/10.1521/soco.1993.11.1.150

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77,* 305–327. http://dx.doi.org/10.1111/j.2044-8295.1986.tb02199.x

Brunswik, E. (1952). The conceptual framework of psychology. *Psychological Bulletin. 49,* 654–656.

Burton, A. M., Kramer, R. S. S., Ritchie, K. L., & Jenkins, R. (2016). Identity from variation: Representations of faces derived from multiple instances. *Cognitive Science, 40,* 202–223. http://dx.doi.org/10.1111/cogs.12231

Carpinella, C. M., Hehman, E., Freeman, J. B., & Johnson, K. L. (2015). The gendered face of partisan politics: consequences of facial sex typicality for vote choice. *Political Communication, 33,* 1–18.

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science, 20,* 1194–1198. http://dx.doi.org/10.1111/j.1467-9280.2009.02423.x

Carré, J. M., Morrissey, M. D., Mondloch, C. J., & McCormick, C. M. (2010). Estimating aggression from emotionally neutral faces: Which facial cues are diagnostic? *Perception, 39,* 356–377. http://dx.doi.org/10.1068/p6543

Chelnokova, O., Laeng, B., Eikemo, M., Riegels, J., Løseth, G., Maurud, H., . . . Leknes, S. (2014). Rewards of beauty: The opioid system mediates social motivation in humans. *Molecular Psychiatry, 19,* 746–747. http://dx.doi.org/10.1038/mp.2014.1

Cikara, M., & Van Bavel, J. J. (2014). The neuroscience of intergroup relations: An integrative review. *Perspectives on Psychological Science, 9,* 245–274. http://dx.doi.org/10.1177/1745691614527464

Correll, J., Hudson, S. M., Guillermo, S., & Earls, H. A. (2016). Of Kith and Kin: Perceptual Enrichment, Expectancy, and Reciprocity in Face Perception. *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc*., *21,* 336–360.

Cunningham, W. A., Johnson, M. K., Raye, C. L., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2004). Separable neural components in the processing of black and white faces. *Psychological Science, 15,* 806–813. http://dx.doi.org/10.1111/j.0956-7976.2004.00760.x

DeBruine, L. M., & Jones, B. C. (2017). *Face research lab London setz*. http://dx.doi.org/10.6084/m9.figshare.5047666

Deegan, M. P., Hehman, E., Gaertner, S. L., & Dovidio, J. F. (2015). Positive expectations encourage generalization from a positive intergroup interaction to outgroup attitudes. *Personality and Social Psychology Bulletin, 41,* 52–65. http://dx.doi.org/10.1177/0146167214556240

Diehl, M. (1990). The minimal group paradigm: Theoretical explanations and empirical findings. *European Review of Social Psychology, 1,* 263–292. http://dx.doi.org/10.1080/14792779108401864

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

Eberhardt, J. L. (n.d.). *Eberhardt Face Database*. Retrieved from https://web.stanford.edu/group/mcslab/cgi-bin/wordpress/examine-the-research/

Fessler, D. M. T., & Holbrook, C. (2013). Friends shrink foes: The presence of comrades decreases the envisioned physical formidability of an opponent. *Psychological Science, 24,* 797–802. http://dx.doi.org/10.1177/0956797612461508

Fessler, D. M. T., Holbrook, C., Pollack, J. S., & Hahn-Holbrook, J. (2014). Stranger danger: Parenthood increases the envisioned bodily formidability of menacing men. *Evolution and Human Behavior, 35,* 109–117. http://dx.doi.org/10.1016/j.evolhumbehav.2013.11.004

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11,* 77–83. http://dx.doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82,* 878–902. http://dx.doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23,* 1–74. http://dx.doi.org/10.1016/S0065-2601(08)60317-2

Freedman, M. B., Leary, T. F., Ossorio, A. G., & Coffey, H. S. (1953). The interpersonal dimension of personality. *Journal of Personality, 20,* 143–161. http://dx.doi.org/10.1111/j.1467-6494.1951.tb01518.x

Freeman, J. B., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118,* 247–279. http://dx.doi.org/10.1037/a0022327

Freeman, J. B., Pauker, K., Apfelbaum, E. P., & Ambady, N. (2010). Continuous dynamics in the real-time perception of race. *Journal of Experimental Social Psychology, 46,* 179–185. http://dx.doi.org/10.1016/j.jesp.2009.10.002

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C. (1993). The common ingroup identity model: Recategorization and the reduction of intergroup bias. *European Review of Social Psychology, 4,* 1–26. http://dx.doi.org/10.1080/14792779343000004

Germine, L., Russell, R., Bronstad, P. M., Blokland, G. A. M., Smoller, J. W., Kwok, H., . . . Wilmer, J. B. (2015). Individual aesthetic preferences for faces are shaped mostly by environments, not genes. *Current Biology, 25,* 2684–2689. http://dx.doi.org/10.1016/j.cub.2015.08.048

Golby, A. J., Gabrieli, J. D. E., Chiao, J. Y., & Eberhardt, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature, 4,* 845–850.

Gramzow, R. H., Gaertner, L., & Sedikides, C. (2001). Memory for in-group and out-group information in a minimal group context: The self as an informational base. *Journal of Personality and Social Psychology, 80,* 188–205. http://dx.doi.org/10.1037/0022-3514.80.2.188

Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the labor market. *The American Economic Review, 84,* 1174–1194.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233. http://dx.doi.org/10.1016/S1364-6613(00)01482-0

Hehman, E., Carpinella, C. M., Johnson, K. L., Leitner, J. B., & Freeman, J. B. (2014). Early processing of gendered facial cues predicts the electoral success of female politicians. *Social Psychological and Per-*

sonality Science, 5, 815–824. http://dx.doi.org/10.1177/1948550614534701

Hehman, E., Gaertner, S. L., Dovidio, J. F., Mania, E. W., Guerra, R., Wilson, D. C., & Friel, B. M. (2012). Group status drives majority and minority integration preferences. *Psychological Science, 23,* 46–52. http://dx.doi.org/10.1177/0956797611423547

Hehman, E., Ingbretsen, Z. A., & Freeman, J. B. (2014). The neural basis of stereotypic impact on multiple social categorization. *NeuroImage, 101,* 704–711. http://dx.doi.org/10.1016/j.neuroimage.2014.07.056

Hehman, E., Leitner, J. B., Deegan, M. P., & Gaertner, S. L. (2015). Picking teams: When dominant facial structure is preferred. *Journal of Experimental Social Psychology, 59,* 51–59. http://dx.doi.org/10.1016/j.jesp.2015.03.007

Hehman, E., Leitner, J. B., & Gaertner, S. L. (2013). Enhancing static facial features increases intimidation. *Journal of Experimental Social Psychology, 49,* 747–754. http://dx.doi.org/10.1016/j.jesp.2013.02.015

Hehman, E., Mania, E. W., & Gaertner, S. L. (2010). Where the division lies: Common ingroup identity moderates the cross-race facial-recognition effect. *Journal of Experimental Social Psychology, 46,* 445–448. http://dx.doi.org/10.1016/j.jesp.2009.11.008

Hehman, E., Sutherland, C. A. M., Flake, J. K., & Slepian, M. L. (2017). The unique contributions of perceiver and target characteristics in person perception. *Journal of Personality and Social Psychology*. Advance online publication.

Hehman, E., Volpert, H. I., & Simons, R. F. (2014). The N400 as an index of racial stereotype accessibility. *Social Cognitive and Affective Neuroscience, 9,* 544–552.

Hönekopp, J. (2006). Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 199–209. http://dx.doi.org/10.1037/0096-1523.32.2.199

Hugenberg, K. (2005). Social categorization and the perception of facial affect: Target race moderates the response latency advantage for happy faces. *Emotion, 5,* 267–276. http://dx.doi.org/10.1037/1528-3542.5.3.267

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science, 15,* 342–345. http://dx.doi.org/10.1111/j.0956-7976.2004.00680.x

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117,* 1168–1187. http://dx.doi.org/10.1037/a0020463

Hutchings, P. B., & Haddock, G. (2008). Look Black in anger: The role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology, 44,* 1418–1420. http://dx.doi.org/10.1016/j.jesp.2008.05.002

Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face! Visual encoding of stereotype content. *Frontiers in Psychology, 4,* 386. http://dx.doi.org/10.3389/fpsyg.2013.00386

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical measures of attention to the race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology, 85,* 616–626. http://dx.doi.org/10.1037/0022-3514.85.4.616

Ito, T. A., & Urland, G. R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective & Behavioral Neuroscience, 5,* 21–36. http://dx.doi.org/10.3758/CABN.5.1.21

Jenkins, R., White, D., Van Montfort, X., & Burton, A. M. (2011). Variability in photos of the same face. *Cognition, 121,* 313–323. http://dx.doi.org/10.1016/j.cognition.2011.08.001

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology, 102,* 116–131. http://dx.doi.org/10.1037/a0025335

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103,* 54–69. http://dx.doi.org/10.1037/a0028347

Jussim, L., Coleman, L. M., & Lerch, L. (1987). The nature of stereotypes: A comparison and integration of three theories. *Journal of Personality and Social Psychology, 52,* 536–546. http://dx.doi.org/10.1037/0022-3514.52.3.536

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). Intergroup perception and cognition: An integrative framework for understanding the causes and onsequences of social categorization. *Advances in Experimental Social Psychology, 55,* 1–80. http://dx.doi.org/10.1016/bs.aesp.2016.10.001

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: Guilford Press.

Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: A social relations analysis. *Psychological Bulletin, 102,* 390–402. http://dx.doi.org/10.1037/0033-2909.102.3.390

Kenny, D. A., Gomes, S. B., & Kowal, C. (2015). The Intergroup Social Relations Model: ISRM. *Group Dynamics: Theory, Research, and Practice, 19,* 152–165. http://dx.doi.org/10.1037/gdn0000028

Kenny, D. A., Horner, C., Kashy, D. A., & Chu, L. C. (1992). Consensus at zero acquaintance: Replication, behavioral cues, and stability. *Journal of Personality and Social Psychology, 62,* 88–97. http://dx.doi.org/10.1037/0022-3514.62.1.88

Kenny, D. A., & West, T. V. (2010). Similarity and agreement in self-and other perception: A meta-analysis. *Personality and Social Psychology Review, 14,* 196–213. http://dx.doi.org/10.1177/1088868309353414

Kubota, J. T., Banaji, M. R., & Phelps, E. A. (2012). The neuroscience of race. *Nature Neuroscience, 15,* 940–948. http://dx.doi.org/10.1038/nn.3136

Kubota, J. T., & Ito, T. A. (2007). Multiple cues in social perception: The time course of processing race and facial expression. *Journal of Experimental Social Psychology, 43,* 738–752. http://dx.doi.org/10.1016/j.jesp.2006.10.023

Kubota, J. T., & Ito, T. (2017). Rapid race perception despite individuation and accuracy goals. *Social Neuroscience, 12,* 468–478. http://dx.doi.org/10.1080/17470919.2016.1182585

Kunda, Z., Thagard, P., Kunda, Z., Thagard, P., Kintsch, W., & Ross, M. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103,* 284–308. http://dx.doi.org/10.1037/0033-295X.103.2.284

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & Van Knippenberg, A. D. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24,* 1377–1388. http://dx.doi.org/10.1080/02699930903485076

Leary, T. F. (1957). *The interpersonal diagnosis of personality; a functional theory and methodology for personality evaluation*. Eugene, OR: Resource Publications.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago Face Database: A free stimulus set of faces and norming data. *Behavior Research Methods, 47,* 1122–1135. http://dx.doi.org/10.3758/s13428-014-0532-5

Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51,* 93–120. http://dx.doi.org/10.1146/annurev.psych.51.1.93

Marger, M. (2003). *Race and ethnic relations: American and global perspectives*. Thomson Learning Belmont, CA: Wadsworth.

Marlowe, C. M., Schneider, S. L., & Nelson, C. E. (1996). Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased? *Journal of Applied Psychology, 81,* 11–21. http://dx.doi.org/10.1037/0021-9010.81.1.11

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. http://dx.doi.org/10.1177/1745691616677828

McDonald, M. M., Navarrete, C. D., & Van Vugt, M. (2012). Evolution and the psychology of intergroup conflict: The male warrior hypothesis. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 367,* 670–679. http://dx.doi.org/10.1098/rstb.2011.0301

Meissner, C., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7,* 3–35. http://dx.doi.org/10.1037/1076-8971.7.1.3

Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments & Computers, 36,* 630–633. http://dx.doi.org/10.3758/BF03206543

Navarrete, C. D., McDonald, M. M., Molina, L. E., & Sidanius, J. (2010). Prejudice at the nexus of race and gender: An outgroup male target hypothesis. *Journal of Personality and Social Psychology, 98,* 933–945. http://dx.doi.org/10.1037/a0017931

Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: The role of target gender. *Psychological Science, 20,* 155–158. http://dx.doi.org/10.1111/j.1467-9280.2009.02273.x

Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology, 53,* 431–444. http://dx.doi.org/10.1037/0022-3514.53.3.431

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 105,* 11087–11092. http://dx.doi.org/10.1073/pnas.0805664105

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion, 9,* 128–133. http://dx.doi.org/10.1037/a0014520

Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron, 48,* 175–187. http://dx.doi.org/10.1016/j.neuron.2005.09.025

Pillemer, J., Graham, E. R., & Burke, D. M. (2014). The face says it all: CEOs, gender, and predicting corporate performance. *The Leadership Quarterly, 25,* 855–864. http://dx.doi.org/10.1016/j.leaqua.2014.07.002

Purdie-Vaughns, V., & Eibach, R. P. (2008). Intersectional invisibility: The distinctive advantages and disadvantages of multiple subordinate-group identities. *Sex Roles, 59,* 377–391. http://dx.doi.org/10.1007/s11199-008-9424-4

Ratner, K. G., & Amodio, D. M. (2013). Seeing "us vs. them": Minimal group effects on the neural encoding of faces. *Journal of Experimental Social Psychology, 49,* 298–301. http://dx.doi.org/10.1016/j.jesp.2012.10.017

Ratner, K. G., Dotsch, R., Wigboldus, D. H. J., van Knippenberg, A., & Amodio, D. M. (2014). Visualizing minimal ingroup and outgroup faces: Implications for impressions, attitudes, and behavior. *Journal of Personality and Social Psychology, 106,* 897–911. http://dx.doi.org/10.1037/a0036498

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57,* 199–226. http://dx.doi.org/10.1146/annurev.psych.57.102904.190208

Rogers, K. H., & Biesanz, J. C. (2014). The accuracy and bias of interpersonal perceptions in intergroup interactions. *Social Psychological and Personality Science, 5,* 918–926. http://dx.doi.org/10.1177/1948550614537307

Rule, N. O., & Ambady, N. (2009). She's got the look: Inferences from female chief executive officers' faces predict their success. *Sex Roles, 61,* 644–652. http://dx.doi.org/10.1007/s11199-009-9658-9

Sachdev, I., & Bourhis, R. Y. (1991). Power and status differentials in minority and majority group relations. *European Journal of Social Psychology, 21,* 1–24. http://dx.doi.org/10.1002/ejsp.2420210102

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion, 9,* 260–264. http://dx.doi.org/10.1037/a0014681

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47,* 609–612. http://dx.doi.org/10.1016/j.jrp.2013.05.009

Secord, P. F. (1958). Facial features and inference processes in interpersonal perception. In R. Tagiuri & L. Petrullo (Eds.), *Person perception and interpersonal behavior* (pp. 300–315). California: Stanford University Press.

Secord, P. F., Dukes, W. F., & Bevan, W. (1954). Personalities in faces: I. An experiment in social perceiving. *Genetic Psychology Monographs, 49,* 231–279.

Sporer, S. L. (2001). Recognizing faces of other ethnic groups: An integration of theories. *Psychology, Public Policy, and Law, 7,* 36–97. http://dx.doi.org/10.1037/1076-8971.7.1.36

Stanley, D. A., Sokol-hessner, P., Fareri, D. S., Perino, M. T., Delgado, M. R., Banaji, M. R., & Phelps, E. A. (2012). Race and reputation: Perceived racial group trustworthiness infuences the neural correlates of trust decisions. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367,* 744–753. http://dx.doi.org/10.1098/rstb.2011.0300

Stolier, R. M., & Freeman, J. B. (2016). Neural pattern similarity reveals the inherent intersection of social categories. *Nature Neuroscience, 19,* 795–797. http://dx.doi.org/10.1038/nn.4296

Strohminger, N., Gray, K., Chituc, V., Heffner, J., Schein, C., & Heagins, T. B. (2016). The MR2: A multi-racial, mega-resolution database of facial stimuli. *Behavior Research Methods, 48,* 1197–1204. Retrieved from https://web.stanford.edu/group/mcslab/cgi-bin/wordpress/examine-the-research/. http://dx.doi.org/10.3758/s13428-015-0641-9

Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition, 127,* 105–118. http://dx.doi.org/10.1016/j.cognition.2012.12.001

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology, 1,* 149–178. http://dx.doi.org/10.1002/ejsp.2420010202

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago, IL: Nelson.

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66,* 519–545.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences, 12,* 455–460. http://dx.doi.org/10.1016/j.tics.2008.10.001

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168,* 242–249. http://dx.doi.org/10.1016/j.psychres.2008.05.006

Tropp, L. R., & Pettigrew, T. F. (2005). Relationships between intergroup contact and prejudice among minority and majority status groups. *Psychological Science, 16,* 951–957. http://dx.doi.org/10.1111/j.1467-9280.2005.01643.x

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge, MA: Basil Blackwell.

Van Bavel, J. J., Packer, D. J., & Cunningham, W. A. (2008). The neural substrates of in-group bias: A functional magnetic resonance imaging investigation. *Psychological, 19,* 1131–1139. http://dx.doi.org/10.1111/j.1467-9280.2008.02214.x

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.

Verkuyten, M. (2005). Ethnic group identification and group evaluation among minority and majority groups: Testing the multiculturalism hypothesis. *Journal of Personality and Social Psychology, 88,* 121–138. http://dx.doi.org/10.1037/0022-3514.88.1.121

Wang, X., & Tang, X. (2009). Face photo-sketch synthesis and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 31,* 1955–1967. http://dx.doi.org/10.1109/TPAMI.2008.222

Webster, M. A., & Macleod, D. I. A. (2011). Visual adaptation and face perception. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 366,* 1702–1725. http://dx.doi.org/10.1098/rstb.2010.0360

West, T. V. (2011). Interpersonal perception in cross-group interactions: Challenges and potential solutions. *European Review of Social Psychology, 22,* 364–401. http://dx.doi.org/10.1080/10463283.2011.641328

West, T. V., & Kenny, D. A. (2011). The truth and bias model of judgment. *Psychological Review, 118,* 357–378. http://dx.doi.org/10.1037/a0022936

Wilson, J. P., Hugenberg, K., & Rule, N. O. (2017). Racial bias in judgments of physical size and formidability: From size to threat. *Journal of Personality and Social Psychology, 113,* 59–80. http://dx.doi.org/10.1037/pspi0000092

Zebrowitz, L. A., Kikuchi, M., & Fellous, J.-M. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology, 98,* 175–189. http://dx.doi.org/10.1037/a0017990

Zebrowitz, L. A., & Montepare, J. M. (2005). Psychology. Appearance DOES matter. *Science, 308,* 1565–1566. http://dx.doi.org/10.1126/science.1114170

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass, 2,* 1497–1517. http://dx.doi.org/10.1111/j.1751-9004.2008.00109.x

Received September 11, 2017
Revision received August 2, 2018
Accepted August 6, 2018 ■