Group evaluations as self-group distancing

Essien, Iniobong; Otten, Sabine; Degner, Juliane

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2020

Link to publication in University of Groningen/UMCG research database

RESEARCH ARTICLE

Group evaluations as self-group distancing: Ingroup typicality moderates evaluative intergroup bias in stigmatized groups

Iniobong Essien[1,2] | Sabine Otten[3] | Juliane Degner[2]

[1]FernUniversität in Hagen, Hagen, Germany

[2]Universität Hamburg, Hamburg, Germany

[3]University of Groningen, Groningen, The Netherlands

Correspondence

Iniobong Essien, Institute of Psychology, FernUniversität in Hagen, Universitätsstr 33, 58097 Hagen, Germany.
Email: iniobong.essien@fernuni-hagen.de

Funding information

Deutscher Akademischer Austauschdienst, Grant/Award Number: 91620926

Abstract

Outgroup favoritism among members of stigmatized groups can be seen as a form of self-group distancing. We examined how intergroup evaluations in stigmatized groups vary as a function of ingroup typicality. In Studies 1 and 2, Black participants ($N = 125,915$; $N = 766$) more strongly preferred light-skinned or White relative to dark-skinned or Black individuals the lighter their own skin tone. In Study 3, overweight participants ($N = 147,540$) more strongly preferred normal-weight relative to overweight individuals the lower their own body weight. In Study 4, participants with disabilities ($N = 35,058$) more strongly preferred non-disabled relative to disabled individuals the less visible they judged their own disability. Relationships between ingroup typicality and intergroup evaluations were at least partially mediated by ingroup identification (Studies 2 and 3). A meta-analysis across studies yielded an average effect size of $r = .12$. Furthermore, higher ingroup typicality was related to both ingroup and outgroup evaluations. We discuss ingroup typicality as an individual constraint to self-group distancing among stigmatized group members and its relation to intergroup evaluations.

KEYWORDS

ingroup favoritism, ingroup typicality, intergroup evaluations, self-group distancing, social identity, stigmatized groups

1 | INTRODUCTION

Picture Tyree and Jamal, who both view themselves as Black Americans. Whereas Jamal is very dark-skinned, Tyree's complexion is very light. This difference is also reflected in their daily life experiences. Most people agree that Jamal is Black; fewer people are that confident when judging Tyree. May these different perceptions and experiences influence how these two individuals generally think and feel about Black Americans relative to White Americans? The present research explores how categorizing oneself as a member of a stigmatized group, yet appearing more or less typical for or similar to that group, may shape how we feel

about our ingroup. More specifically, we argue that ingroup and outgroup evaluations can reflect a tendency to distance the self from a stigmatized identity, and that this tendency is constrained by the extent to which an individual appears more or less typical for the ingroup.

2 | SELF-GROUP DISTANCING AMONG MEMBERS OF STIGMATIZED GROUPS

In many societies, members of stigmatized groups—social groups that are ascribed comparatively less prestige than others (Brown-Iannuzzi, Lundberg, Kay, & Payne, 2015; Tajfel & Turner, 1979)—frequently

experience discrimination and threats to their social identity (Schmitt, Branscombe, Postmes, & Garcia, 2014). However, individuals differ in how they cope with their ingroup's disadvantage. Experiences of disadvantage prompt some individuals to pursue group-level strategies aimed at improving the status of the ingroup (collective action; Tajfel & Turner, 1979); yet others pursue individual-level strategies aimed at improving their personal situation (see de Lemus & Stroebe, 2015). Individual-level and group-level strategies to cope with group disadvantage are often incompatible with one another. *Self-group distancing* represents one such individual-level strategy that comes at the expense of group-level outcomes. Broadly, self-group distancing describes strategic behaviors displayed by members of stigmatized groups, who sacrifice group goals in order to pursue individual goals (Derks, van Laar, & Ellemers, 2016). For example, individuals might distance themselves by perceiving or emphasizing their dissimilarities with the ingroup (Weiss & Lang, 2012), endorsing negative stereotypes about the ingroup (Derks, Ellemers, van Laar, & Groot, 2011), describing themselves more in terms of (positive) outgroup characteristics (Derks, Ellemers, et al., 2011; Derks, van Laar, Ellemers, & Raghoe, 2015), psychologically distancing themselves from the ingroup (Derks, van Laar, Ellemers, & Groot, 2011), or evaluating the stigmatized ingroup negatively (Guimond, Dif, & Aupy, 2002). The current research focuses on relative preferences for an outgroup over the ingroup, which we argue also represents a way of distancing the self from a stigmatized social identity.

Scholars attribute self-group distancing among members of stigmatized groups to the biased and identity threatening social contexts they have to navigate in their daily lives (see Derks et al., 2016). According to this view, self-group distancing is a consequence of stigmatized group members' effort to assimilate to outgroup contexts, in which they are frequently exposed to threats to their social identity. These outgroup contexts are shaped by dominant groups, and stigmatized group members thus likely encounter negative ingroup stereotypes as well as positive outgroup stereotypes, decreasing their willingness to be categorized according to their (stigmatized) group membership (Branscombe, Ellemers, Spears, & Doosje, 1999). Consequently, distancing the self from negative ingroup stereotypes and/or associating the self with positive outgroup stereotypes seems to provide a way for stigmatized group members to cope with social identity threats, and to personally thrive in social contexts which are biased against them. Lastly, not all members of stigmatized groups react to identity threatening contexts in a similar fashion. Instead, self-group distancing seems to be more likely among individuals who are less identified with their stigmatized ingroup (Derks et al., 2016). Thus, the extent to which individuals identify with their stigmatized ingroup seems to provide an important antecedent to self-group distancing.

So far, most research has investigated self-group distancing in organizational and work settings, and specifically among women in leadership roles. For example, low gender identified senior policewomen described themselves in more masculine terms when they were reminded of experiences in which they were stereotyped (Derks, van Laar, et al., 2011). Similarly, low gender identified senior women leaders in another study who reported having experienced more gender discrimination characterized themselves with more masculine traits (Derks, Ellemers, et al., 2011). Lastly, self-group distancing has also been observed in other social identity domains. For example, older adults were more likely to distance themselves from their age group when they were exposed to negative age stereotypes (Weiss & Freund, 2012). Other research suggests that Gay men might distance themselves from negative ingroup stereotypes by displaying stereotypically male behaviors (Bishop, Kiss, Morrison, Rushe, & Specht, 2014; Clarkson, 2006; Eguchi, 2009).

Taken together, self-group distancing represents an individual-level strategy aimed at individual mobility, often at the expense of group interests; is thought to reflect stigmatized group members' responses to social identity threats; and is more likely observed among individuals who are less identified with their ingroup. Given these insights, we argue that ingroup and/or outgroup evaluations among members of stigmatized groups may reflect self-group distancing.

3 | GROUP EVALUATIONS AS A MANIFESTATION OF SELF-GROUP DISTANCING

Ingroup favoritism is a robust phenomenon, influencing feelings, beliefs, and behaviors (Greenwald & Pettigrew, 2014; Hewstone, Rubin, & Willis, 2002; Mullen, Brown, & Smith, 1992). Theorizing suggests that identifying with and favoring one's ingroup affects well-being (e.g., self-esteem; Tajfel & Turner, 1979), and might also provide a buffer against social rejection (Branscombe, Schmitt, & Harvey, 1999). The occurrence of ingroup favoritism, however, also depends on a group's position within status and power hierarchies in society. Specifically, members of stigmatized groups do not always display ingroup favoritism.

Whereas stigmatized group members often self-report similar degrees of ingroup liking to non-stigmatized group members (e.g., Mullen et al., 1992), they sometimes display evaluative outgroup favoritism on implicit measures (e.g., Nosek, Banaji, & Greenwald, 2002).[1] In our view, outgroup favoritism on evaluative measures reflects, at least to some degree, self-group distancing (see also Derks et al., 2016). Self-group distancing among members of stigmatized groups is often a response to stereotyping in biased contexts (e.g., Derks, Ellemers, et al., 2011; Derks, van Laar, et al., 2011; Weiss & Freund, 2012). Because the valence of stereotypes is causally related to the valence of group evaluations (e.g., Phills, Hahn, & Gawronski, 2020), stigmatized group members' exposure to stereotypes is likely to have downstream consequences for group evaluations. Consequently, outgroup favoritism may directly follow from the activation or endorsement of negative ingroup stereotypes and/or positive outgroup stereotypes. This

---

[1]This article uses the term "implicit" to refer to indirect measurement procedures and their outcomes. However, this "implicit-as-indirect" conceptualization (Corneille & Hütter, 2020, p. 1) does not equate implicit and explicit measurement outcomes with different mental representations or features of automaticity.

rationale is further supported by findings that outgroup favoritism is more pronounced in groups to the extent that they are stigmatized by others (i.e., negatively evaluated; Essien, Calanchini, & Degner, 2020; Rudman, Feinberg, & Fairchild, 2002). Taken together, we view outgroup favoritism among members of stigmatized groups as one manifestation of self-group distancing, because such group evaluations might at least in part reflect the activation or endorsement of negative ingroup and/or positive outgroup stereotypes.

Findings regarding stigmatized group members' intergroup evaluations on implicit measures vary greatly, and different studies have documented ingroup favoritism, outgroup favoritism, or no group preferences (Axt, Ebersole, & Nosek, 2014; Degner, Essien, & Reichardt, 2016; Essien et al., 2020; Rae, Newheiser, & Olson, 2015; Rudman et al., 2002). Thus far, these variations in intergroup evaluations have not been exhaustively explained. Here, we suggest that one moderator of group evaluations may be the degree of similarity individuals perceive between themselves, their ingroup, and a non-stigmatized outgroup. More specifically, we assume that group members who *appear* as less phenotypically prototypical for the stigmatized ingroup and thus more similar to a non-stigmatized outgroup may be more likely to display outgroup favoritism as a form of self-group distancing.

4 | PHENOTYPIC PROTOTYPICALITY AS (DIS)SIMILARITY WITH GROUP PROTOTYPES

Phenotypic prototypicality refers to the degree to which individuals' appearances are perceived as similar to a group prototype (Davies, Hutchinson, Osborne, & Eberhardt, 2016). For example, Blacks with darker skin tone or with more Afrocentric facial features (e.g., broader nose, fuller lips) are more readily perceived as prototypically Black. Such phenotypic racial prototypicality has been linked to many real-life outcomes (Maddox, 2004). For example, Black individuals with darker skin tone were evaluated more negatively than those with lighter skin tone on both, implicit and explicit measures (e.g., Hagiwara, Kashy, & Cesario, 2012). Furthermore, Black individuals who were perceived as more prototypical were more likely to be rejected by non-Black outgroup members (Hebl, Williams, Sundermann, Kell, & Davies, 2012), more strongly associated with attributes stereotypically linked with Blacks (e.g., Blair, Judd, Sadler, & Jenkins, 2002), and more likely to be perceived as threatening than those who were perceived as less prototypical (e.g., Dixon & Maddox, 2005; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Kahn & Davies, 2011; Ma & Correll, 2011). Such effects influence outcomes in other important areas of everyday life, such as the educational system. Compared with individuals with lighter skin tone, dark-skinned individuals were judged as less competent, and these judgments were associated with lower educational expectations (Meeus, Mayor, González, Brown, & Manzi, 2017). Together, these studies suggest that higher phenotypic prototypicality among members of stigmatized groups is related to more negative experiences in outgroup contexts.

These findings regarding prototypicality also suggest that phenotypic appearance, in addition to group membership per se, shapes interaction experiences. We propose that, eventually, these experiences may relate to stigmatized group members' own perceptions and attitudes about both ingroup and outgroups. In other words, we suggest that to the extent that ingroup typicality reflects phenotypic appearance, it might influence stigmatized group members' ability to (psychologically) distance themselves from or connect with their ingroup and outgroups. However, most research has investigated how stigmatized group members are perceived, judged, and treated by *others*, and only few studies have actually looked at how phenotypic prototypicality may relate to their *own* perceptions and group evaluations.

Perceived ingroup prototypicality may be associated with a number of psychological processes that construe the self as proximal to or distant from the ingroup and, in turn, may relate to differences in group evaluations. For example, less prototypical Black individuals displayed less ingroup identification (Harvey, LaBeach, Pridgen, & Gocial, 2005). Similarly, Black participants with lighter skin tone reported less closeness to Black people than those with darker skin tone (Brown, Ward, Lightbourn, & Jackson, 1999). Lastly, less prototypical Black and Latino individuals were less likely to be identified with their racial ingroups than more prototypical individuals (Wilkins, Kaiser, & Rieck, 2010). In sum, to the extent that members of stigmatized groups are perceived as more prototypical they seem to be more identified with their ingroup. Crucially, these relationships with identification were not only observed for self-reported prototypicality (Harvey et al., 2005), but also for other-rated prototypicality (Brown et al., 1999; Wilkins et al., 2010). Thus, relationships between ingroup prototypicality and identification may at least in part reflect actual phenotypic differences in appearance between members of stigmatized groups. In our view, this suggests that ingroup prototypicality may constrain stigmatized group members' ability to identify with the ingroup. Because ingroup identification is related to ingroup liking and satisfaction with the ingroup (Leach et al., 2008), low ingroup prototypicality may also be related to lower levels of ingroup favoritism.

Furthermore, perceived ingroup typicality may be associated with stigmatized group members' perception of group boundaries. Stigmatized group members who appear less phenotypically prototypical may perceive group boundaries as more permeable, thus perceiving a higher liberty to distance the self from the ingroup. Social Identity Theory (Tajfel & Turner, 1979) assumes that the perception of permeability of group boundaries influences whether group members who are unsatisfied with the lower status of their ingroup attempt to leave the group, a strategy termed *individual mobility*. According to social identity theory, individual mobility might manifest itself not only in the sense of physically leaving a group, but also in the sense of psychologically distancing the self from the group, that is by disidentifying from the former ingroup (Tajfel & Turner, 1979). Hence, we argue that ingroup typicality might pose an individual constraint to the permeability of group boundaries,

affecting stigmatized group members' tendency to distance themselves from the ingroup.

Lastly, weak group boundaries may decrease ingroup identification (cf. Reimer et al., 2016) and increase perspective-taking with outgroup members (e.g., Todd & Burgmer, 2013), thereby decreasing positive ingroup evaluations and enhancing positive outgroup evaluations (e.g., Rae et al., 2015). Taken together, we suggest that phenotypic appearance should be related to the extent to which members of stigmatized groups perceive themselves as typical for the ingroup. We further suggest that ingroup typicality and the resulting sense of (dis)similarity and (dis)connection with the ingroup may be related to ingroup and outgroup favoritism in stigmatized group members' intergroup evaluations on implicit and explicit measures.

5 | THE PRESENT RESEARCH

We present four studies in which we investigated effects of ingroup typicality on evaluative ingroup favoritism on implicit and explicit measures. We analyzed data from two sources: Project Implicit and the American National Election Studies (ANES). Project Implicit is a website where visitors can complete various Implicit Association Tests (IAT; Greenwald, McGhee, & Schwartz, 1998) as well as explicit measures. For many target domains, data have been collected for more than a decade, yielding massive, openly available datasets (https://osf.io/y9hiq/; see Xu et al., 2017). ANES provides survey data from representative probability samples of adult U.S. citizens, which focus on political behavior and also include measures of intergroup attitudes (see http://www.electionstudies.org). We reviewed these sources for datasets, in which (a) participants self-categorized as members of a stigmatized social identity and that (b) included at least one variable differentiating between different levels of phenotypic prototypicality as a proxy for ingroup typicality. Three Project Implicit datasets and two ANES waves (2012 and 2016) fulfilled these criteria, investigating intergroup evaluations regarding skin tone and racial attitudes (Study 1 and 2), weight status (Study 3), and disability (Study 4). Different variables were available as indirect indicators of ingroup typicality. In Studies 1 and 2, we assumed that Black participants with lighter skin tone have lower perceived/phenotypic typicality for Blacks (or higher similarity to Whites). We investigated effects of self-reported skin tone (Study 1) and other-observed skin tone (Study 2). In Study 3, we assumed that overweight participants have lower perceived/phenotypic typicality for the overweight category (or higher similarity to the normal-weight category) the lower their self-reported weight status and the lower they believed others judge their weight status. Finally, in Study 4, we assumed that disabled participants have lower perceived/phenotypic typicality for the category disabled and higher similarity to non-disabled people the more they were able to hide their disability, the less they felt affected by their disability in daily life, and the less severe they judged their disability. In all four

studies, we examined how these variables were related to group evaluations. In addition, we explored in Study 2 and 3 whether ingroup typicality was related to ingroup identification. Lastly, we conducted a series of meta-analyses across studies, which examined (a) the overall magnitude of the relationship between ingroup typicality and intergroup evaluations and (b) whether ingroup typicality was related to ingroup and/or outgroup evaluations. All analyses were done using R.[2] Analyses scripts are accessible via Open Science Framework (https://osf.io/z4xwx/).

6 | STUDY 1

In Study 1, we explored how Black participants' evaluations of light-skinned and dark-skinned people varied depending on their own skin tone perception.

6.1 | Method

6.1.1 | Participants

Participants were visitors of the Project Implicit demonstration website (https://implicit.harvard.edu) between 2004 and 2015 who self-categorized as Black or African American ($N = 125,915$). Given such large sample size, achieved power for small effects ($\rho = .1$) was $1 - \beta = 1.0$ (two-tailed).

6.1.2 | Measures

Self-reported skin tone
Black participants' self-reported skin tone was measured on a scale ranging from 1 (*very dark*) to (*very light*)—see Table 2 for all response options and descriptive statistics.

Skin tone IAT
In the skin tone IAT, attribute stimuli were positive and negative words that had to be categorized as *good* versus *bad*. The target stimuli were dark-skinned and light-skinned male and female faces—see https://osf.io/y9hiq/, for a complete overview of stimuli and procedures.

Self-reported preference
Participants indicated their relative preference for dark-skinned over light-skinned people on a scale from 1 (*strong preference for dark-skinned people*) to 7 (*strong preference for light-skinned people*), with the midpoint indicating no preference.

TABLE 1 Mean IAT D Scores, self-reported preference scores, feeling thermometer difference scores, and their standard deviations as a function of Black participants' self-reported skin tone in Study 1

	IAT			Preference			Thermometer		
	M	SD	n	M	SD	n	M	SD	n
All Black participants	0.09	0.45	125,915	3.79	1.08	115,738	−0.37	1.86	121,597
I consider my skin to be									
(1) Very dark	0.05	0.46	3,834	3.46	1.33	3,536	−1.13	2.58	3,749
(2) Dark	0.04	0.45	24,534	3.57	1.13	22,853	−0.83	1.94	24,085
(3) Somewhat dark	0.06	0.45	21,910	3.71	1.06	20,497	−0.59	1.84	21,530
(4) Medium	0.09	0.45	42,636	3.86	1.02	39,623	−0.27	1.71	41,838
(5) Somewhat light	0.14	0.45	14,785	4.04	1.01	13,953	0.09	1.68	14,564
(6) Light	0.19	0.46	11,191	4.11	1.07	10,523	0.26	1.72	11,003
(7) Very light	0.21	0.48	1,843	3.98	1.17	1,742	0.14	2.13	1,811
White participants	0.40	0.40	465,925	4.27	0.95	434,900	0.54	1.72	450,702

Note: IAT = IAT D Score; Preference = self-reported preference score; Thermometer = feeling thermometer difference score. Higher means on IATs and self-report measures indicate a stronger preference for light-skinned over dark-skinned individuals. Variable sample sizes are based on missing values in the different dependent variables. For comparison, we also report average scores for the non-stigmatized group (i.e., White participants).

	1	2	3	4	5	M	SD
1. Typicality	–					3.67	1.34
2. IAT	0.09***	–				0.09	0.45
3. Preference	0.16***	0.11***	–			3.79	1.08
4. Thermometer	0.20***	0.10***	0.56***	–		−0.37	1.86
5. Ingroup	0.13***	0.08***	0.25***	0.41***	–	2.50	2.29
6. Outgroup	−0.03***	0.00	−0.20***	−0.41***	0.67***	2.87	2.29

TABLE 2 Zero-order correlations of Black participants' self-reported skin tone, IAT D Scores, self-reported preference scores, and individual feeling thermometers in Study 1

Note: Typicality = measure of ingroup typicality (i.e., self-reported skin tone); IAT = IAT D Score; Preference = self-reported preference score; Thermometer = feeling thermometer difference score; Ingroup = feeling thermometer regarding dark-skinned people; Outgroup = feeling thermometer regarding dark-skinned people. Higher means on IATs and self-report measures indicate a stronger preference for light-skinned over dark-skinned individuals.

*p < .05.

**p < .01.

***p < .001.

Feeling thermometers

Participants indicated their feelings regarding light-skinned and dark-skinned people using two scales from 1 (*very cold*) to 11 (*very warm*).

6.2 | Results

We used IAT D Scores and self-reported preference scores as reported in the dataset. In addition, we computed feeling thermometer difference scores by subtracting participants' evaluations of dark-skinned people from evaluations regarding light-skinned people. In all measures, more positive scores indicate a relative preference for light-skinned people over dark-skinned people. Table 1 presents descriptive statistics for measures of intergroup evaluations for the different levels of self-reported skin tone. As means of comparison, we also report average scores for the non-stigmatized group (e.g., White participants).

Overall, Black participants had a positive IAT D Score, which significantly differed from zero $t(125,914) = 69.69, p < .001, d_z = .20$, 95% CI [0.19, 0.20], indicating an overall small preference of light-skinned relative to dark-skinned people on the IAT. We also tested participants' mean self-reported preference scores against the scale midpoint, $t(115,737) = −66.77, p < .001, d_z = −.20$, 95% CI [−0.20, −0.19], and their thermometer difference scores against zero, $t(121,596) = −69.65, p < .001, d_z = −.20$, 95% CI [−0.21, −0.19], indicating a preference for dark-skinned relative to light-skinned people on both self-report measures.

Table 2 reports correlations between self-reported skin tone and group evaluations. Importantly, we observed a small positive correlations between participants' self-reported skin tone and their IAT D Scores, $r(120,731) = .09, p < .001$, 95% CI [0.09, 0.10], their self-reported preference scores, $r(112,725) = .16, p < .001$, 95% CI [0.16, 0.17], and their thermometer difference scores, $r(118,578) = .20, p < .001$, 95% CI [0.19, 0.20].

6.3 | Discussion

Black participants showed stronger preferences for light-skinned relative to dark-skinned individuals on the IAT and two self-report measures to the extent that they self-reported lighter skin tone. This provides initial support for the assumption that differences in ingroup typicality explain variance in intergroup evaluations. However, a reversed interpretation remains plausible. Participants may perceive their skin tone to be lighter *because* of their stronger relative preference for light-skinned relative to dark-skinned people and a relative disconnectedness with the ingroup. According to this reasoning, effects might be due to participants "adjusting" perceptions of their skin tone as a consequence of their attitudes, due to lower identification, or generally as a strategy to achieve cognitive consistency between their attitudes and their self-observations (cf. Gawronski, Brochu, Sritharan, & Strack, 2012). From this viewpoint, two people with the same skin tone may report very different levels of subjective skin tone, based on different levels of ingroup identification. Study 2 addresses this possible alternative interpretation.

7 | STUDY 2

Study 2—a pre-registered analysis of ANES 2012 and 2016 data—aimed at replicating and extending findings of the previous study. Again, we used skin tone as proxy for perceived ingroup typicality in Black participants. Unlike the previous study, the ANES datasets provided skin tone categorizations also as other-observations. Interviewers who conducted face-to-face interviews also reported participants' perceived skin tone. However, only self-report measures of intergroup evaluations were assessed (i.e., feeling thermometer scales).

From the results of Study 1, we expected Black participants to display higher levels of ingroup favoritism the darker their interviewer-assessed and self-reported skin tone. The pre-registration, materials, and analysis script are accessible via https://osf.io/kn7qv/.

In addition to the pre-registered analyses, we explored whether ingroup typicality was related to group identification among members of stigmatized groups. The results of these exploratory analyses may reveal whether lower ingroup typicality involves psychological distancing from a stigmatized identity (i.e., disidentification). Together, these analyses may point towards involved psychological processes underlying the correlation between skin tone measures and intergroup evaluations.

7.1 | Method

7.1.1 | Participants

Participants were 766 respondents (291 male, 472 female, 3 unknown; $MD_{age} = 43$, $SD = 15.99$) from the ANES 2012 ($n = 413$) and ANES 2016 ($n = 353$) Time Series, who self-identified as African American or Black. All ANES 2012 respondents participated in face-to-face interviews. For ANES 2016, 119 respondents participated in face-to-face interviews and 234 respondents participated online. Power analysis was based on the lower boundary of the confidence interval for the correlation between self-reported skin tone and self-reported preference scores observed in Study 1. Given $\alpha = .05$, and $1-\beta = 0.95$, a sample size of at least $N = 425$ was needed to detect an effect of $\rho = .158$ (one-tailed; see pre-registration at https://osf.io/95q4v/).

7.1.2 | Measures

Skin tone assessment

Skin tone was assessed using a skin color scale originally designed by Massey and Martin (2003), a 10-point graphical scale depicting a human hand in ten different shades (1 = very light; 10 = very dark). We recoded values in parallel to Study 1 such that higher values indicate lower ingroup typicality. Respondents' skin tone was recorded during two interviews; once at the end of the pre-election interview and once at the end of the post-election interview, respectively. Basing our analysis on intra-class correlations of $r_{ICC} = .85$, (95% CI [0.82, 0.87]), between interviewers' skin tone observations, we averaged the two skin tone observations. In ANES 2016, respondents additionally self-reported their skin tone using the same scale.[3]

Self-reported preference

In both studies, feeling thermometers were administered as part of the post-election data collection via computer-aided self-interviews. Evaluations of Blacks and Whites were assessed separately, using scales from 0 (*unfavorable/cold*) to 100 (*favorable/warm*). We computed a feeling thermometer difference score in parallel to Study 1 by subtracting evaluations regarding Blacks from evaluations regarding Whites. Positive values indicate more positive evaluations of Whites relative to Blacks.

Ingroup identification

Participants were asked how important being Black was to their identity on a scale ranging from 1 (*not at all important*) to 5 (*extremely important*).

7.2 | Results

Exploratory analyses of the main effect of ingroup favoritism revealed that Black participants had a negative feeling thermometer difference score ($M = -16.26$, $SD = 24.02$), significantly different from zero $t(707) = -18.01$, $p < .001$, $d_z = -.68$, 95% CI [−0.76, −0.59],

---

[3]Note that in the ANES 2016 study, face-to-face respondents self-reported their skin tone in addition to interviewer-assessed skin tone, whereas online skin tone was only assessed via self-report.

	1	2	3	4	5	M	SD
1. Skin Tone (Other)	–					5.52	2.24
2. Skin Tone (Self)	0.40***	–				5.48	1.96
3. Identification	0.12**	0.12*	–			4.29	1.07
4. Thermometer	0.11*	−0.06	0.19***	–		16.26	24.02
5. Ingroup	−0.01	−0.10	0.30***	0.42***	–	85.32	19.12
6. Outgroup	−0.12*	−0.02	0.05	−0.68***	0.38***	69.00	23.49

TABLE 3 Zero-order correlations of Black participants' interviewer-assessed skin tone, self-reported skin tone, feeling thermometer difference scores, and individual feeling thermometers in Study 2

Note: Skin Tone (Other) = interviewer-assessed skin tone; Skin Tone (Self) = self-reported skin tone; Identification = ingroup identification; Thermometer = feeling thermometer difference score; Ingroup = ingroup feeling thermometer; Outgroup = outgroup feeling thermometer. Higher means for skin tone measures indicate darker (interviewer-assessed and self-reported) skin tone; a higher feeling thermometer difference score indicates a stronger preference for Blacks relative to Whites; higher means for individual feeling thermometers indicate more favorable/warm evaluations of the respective target group.

*$p < .05$.

**$p < .01$.

***$p < .001$.

thus replicating the effect of self-reported ingroup favoritism observed in Study 1.

Table 3 reports correlations between skin tone ratings and group evaluations. As predicted, we observed a small positive correlation between face-to-face respondents' feeling thermometer difference scores and their interviewer-assessed skin tone, $r(474) = .11$, $p = .006$, 95% CI [0.04, 1.00]. Black participants showed higher levels of ingroup favoritism on feeling thermometer difference scores the darker their observed skin tone. Surprisingly however, participants' self-reported skin tone was not positively correlated with feeling thermometer difference scores, $r(329) = −.06$, $p = .854$, 95% CI [−0.15, 1.00].

In order to explore why the predicted effect did not replicate using respondents' self-reported skin tone ratings, we looked at the agreement between respondents' self-reported skin tone ratings and average interviewer-assessed skin tone ratings. Agreement was surprisingly low, $r_{ICC} = .57$, 95% CI [0.36, 0.71].

7.2.1 | Exploratory analyses

Next, we explored relationships between skin tone ratings and ingroup identification. We observed small positive correlations between participants' interviewer-observed skin tone and their levels of ingroup identification, $r(487) = .12$, $p = .010$, 95% CI [0.03, 0.20]. This indicates that participants with darker observed skin tone displayed higher levels of ingroup identification than did participants with lighter skin tone.

We then conducted a mediation analysis using the "medmod" package[4] with 1,000 bootstrap resamples. This analysis used participants' feeling thermometer difference scores as dependent variable, interviewer-observed skin tone as a predictor, and

ingroup identification as a mediator. Indeed, the indirect effect was significant, $b = −.25$, $SE = .11$, $p = .023$, 95% CI [−0.50, −0.07]. This indicates that the effect of skin tone on ingroup favoritism was mediated by participants' level of ingroup identification.

Regarding participants' self-reported skin tone, we observed a small positive correlation with their levels of ingroup identification, $r(329) = .12$, $p = .034$, 95% CI [0.01, 0.22]. This indicates that participants with darker skin tone reported higher levels of identification than those with lighter skin tone.

7.3 | Discussion

Study 2 indicates that Black participants displayed higher levels of outgroup favoritism on a self-report measure the lighter their skin tone. The replication of effects from Study 1 with the observer skin tone ratings in Study 2 strengthens our interpretation that lower ingroup typicality leads to less ingroup favoritism. It is less likely that skin tone perceptions were systematically biased by participants' intergroup attitudes—skin tone was not self-reported, but recorded by the interviewers. That said, it is also possible that skin tone observations were influenced by participants' responses during the interview, because skin tone observations were made at the end of the interview. On the other hand, it is also important to note that feeling thermometers and other sensitive information were assessed through computer-aided self-interviews without the interviewers' participation. Thus, it remains an open question whether or to what extent skin tone ratings might have been influenced by participants' behavior.

Contrary to expectations, the correlation of self-reported skin tone with ingroup favoritism observed in Study 1 did not replicate in Study 2. Explanations for this null finding might center on differences in measurement procedures between observed and

---

[4]The medmod package uses the "lavaan" package for computations.

self-reported skin tone. For example, observer skin tone ratings were assessed in face-to-face interviews whereas self-reported skin tone ratings were assessed during a self-administered online survey. Furthermore, observer skin tone ratings were assessed twice whereas self-reported skin tone was only assessed once. Consequently, measurement error of self-reported skin tone measures may have been higher compared to observed skin tone measures, which would have differentially attenuated correlations between skin tone measures and ingroup favoritism (e.g., Hunter & Schmidt, 2007).

Lastly, exploratory analyses indicate that the correlations between interviewer-assessed skin tone on ingroup favoritism were partially mediated by participants' level of ingroup identification. Findings of Study 2 thus provide preliminary evidence that lower ingroup typicality might be associated with psychological distancing from a stigmatized identity. Together, results of Study 1 and 2 suggest that how Black Americans evaluate the ingroup and outgroup is related to the degree to which they appear typical of their ingroup in terms of skin tone (i.e., ingroup typicality). In Studies 3 and 4 we explored whether these effects (a) extend to other social categories and (b) are observed when using different operationalizations of ingroup typicality. A replication of these effects would suggest the operation of comparable basic processes mediating the relationship between ingroup typicality and intergroup evaluations.

8 | STUDY 3

Study 3 investigated the link between ingroup typicality and intergroup bias regarding a different social categorization: weight status. Weight is an important factor in interpersonal perception, with overweight individuals being frequently negatively stigmatized (e.g., Crandall, 1994). However, although negative evaluations of overweight individuals are widely shared within many Western societies, own body weight has been shown to be related to weight-related automatic prejudice, with overweight individuals displaying less anti-fat bias and even ingroup favoritism the higher their actual body weight (e.g., Degner & Wentura, 2009; Schwartz, Vartanian, Nosek, & Brownell, 2006). Hence, Study 3 examined whether weight status categorization is related to intergroup evaluations on both the IAT and self-report measures. Furthermore, to further our understanding of the relationship between ingroup typicality and self-group distancing, we again explored relationships between ingroup typicality, group identification, and ingroup favoritism.

8.1 | Method

8.1.1 | Participants

Participants were visitors of the Project Implicit demonstration website between 2004 and 2015, who self-categorized as overweight

($N = 147,540$). Given such large sample size, achieved power for small effects ($\rho = .1$) was $1–\beta = 1.0$ (two-tailed).

8.1.2 | Measures

Weight IAT
The weight IAT followed the same procedure as the skin tone IAT in Study 1 with the exception that target stimuli were images of normal weight and overweight individuals (e.g., faces or body shapes)—see https://osf.io/y9hiq/, for a complete overview of experimental stimuli and procedures.

Self-reported preference
Participants indicated their relative preference for overweight over normal weight individuals on scales from 1 (*strong preference for overweight individuals*) to 7 (*strong preference for normal weight individuals*), with the midpoint indicating no preference. Participants indicated their feelings regarding overweight and normal weight individuals on scales from 1 (*very cold*) to 11 (*very warm*).

Perceived weight status
Participants' reported their perceived weight status and reported how they thought others would judge their weight on a scale ranging from 1 (*very underweight*) to 7 (*very overweight*). For comparability with the previous studies, weight status was recoded such that higher values indicate lower weight. We used both variables as separate proxies for participants' ingroup typicality—see Table 4 for all response options and descriptive statistics.[5]

Ingroup identification
Participants were asked how much they identified with overweight people using a scale from 1 (*Not at all*) to 5 (*Strongly*). Exact wording was "How much do you identify with people who are fat?".[6]

8.2 | Results

Overall, overweight participants had a positive IAT *D* Score, which significantly differed from zero $t(147,539) = 328.59$, $p < .001$, $d_z = .86$, 95% CI [0.85, 0.86], indicating outgroup favoritism onindicating outgroup favoritism on the IAT (see Table 5). We tested participants' mean self-reported preference score against the scale midpoint, $t(143,115) = 254.99$, $p < .001$, $d_z = .67$, 95% CI [0.67, 0.68], and their feeling thermometer difference score against zero,

---

[5]Note that we excluded overweight participants who reported that others would judge them as slightly, moderately, or very underweight ($n = 2,433$).

[6]In our view, the phrasing "fat" does not appear to be a neutral way of addressing people who are heavyweight. However, we do not have clear hypotheses about how this wording might have influenced responses toward this item.

TABLE 4 Mean IAT D Scores, self-reported preference scores, feeling thermometer difference scores, and their standard deviations as a function of participants' self-reported weight status and their report of how others would judge their weight status in Study 3

	IAT			Preference			Thermometer		
	M	SD	n	M	SD	n	M	SD	n
Overweight participants	0.37	0.43	147,540	4.73	1.08	143,116	0.47	2.32	146,555
Currently, I am									
(1) Very overweight	0.24	0.45	19,276	4.40	1.19	18,656	−0.16	2.56	19,139
(2) Moderately overweight	0.33	0.44	38,576	4.56	1.08	37,417	0.16	2.27	38,342
(3) Slightly overweight	0.41	0.42	89,688	4.87	1.03	87,043	0.74	2.24	89,074
Other people would say that I am									
(1) Very overweight	0.21	0.45	10,498	4.30	1.20	10,148	−0.30	2.57	10,415
(2) Moderately overweight	0.29	0.44	22,766	4.49	1.08	21,997	0.08	2.24	22,645
(3) Slightly overweight	0.38	0.43	44,546	4.72	1.02	43,146	0.53	2.20	44,254
(4) Normal weight	0.44	0.41	36,219	4.99	1.03	35,157	0.99	2.29	36,022
Normal weight participants	0.46	0.40	170,182	5.15	1.04	164,764	1.36	2.37	168,887

Note: IAT = IAT *D* Score; Preference = self-reported preference score; Thermometer = feeling thermometer difference score. Higher means on IATs and self-report measures indicate a stronger preference for normal weight over overweight individuals. Variable sample sizes are based on missing values in the different dependent variables. We excluded overweight participants who reported that others would judge them as slightly, moderately, or very underweight (*n* = 2,433). For comparison, we also report average scores for the non-stigmatized group (i.e., normal weight participants).

TABLE 5 Zero-order correlations of overweight participants' self-reported weight status, reports of how others would judge their weight status, IAT *D* Scores, self-reported preference scores, feeling thermometer difference scores, and individual feeling thermometers in Study 3

	1	2	3	4	5	6	7	M	SD
1. Weight (Self)	–							2.48	0.71
2. Weight (Others)	0.75***	–						2.93	0.94
3. Identification	0.38***	0.38***	–					3.21	1.06
4. IAT	0.14***	0.16***	0.14***	–				0.37	0.43
5. Preference	0.17***	0.21***	0.32***	0.22***	–			4.73	1.08
6. Thermometer	0.15***	0.18***	0.31***	0.18***	0.59***	–		0.47	2.32
7. Ingroup	0.10***	0.12***	0.30***	0.16***	0.41***	0.62***	–	5.06	2.08
8. Outgroup	−0.07***	−0.09***	−0.04***	−0.05***	−0.27***	−0.53***	0.34***	4.59	1.93

Note: Weight (Self) = self-reported weight status; Weight (Others) = reports of how others would judge participants' weight status; IAT = IAT *D* Score; Preference = self-reported preference score; Thermometer = feeling thermometer difference score; Ingroup = ingroup feeling thermometer; Outgroup = outgroup feeling thermometer. Higher means on IATs and self-report measures indicate a stronger preference for overweight relative to normal weight individuals.
*p < .05.
**p < .01.
***p < .001.

$t(146,554) = 77.50$, $p < .001$, $d_z = .20$, 95% CI [0.20, 0.21], indicating outgroup favoritism for both self-report measures.

Table 5 reports correlations between weight status, group evaluations, and ingroup identification. Crucially, we observed small correlations between IAT *D* Scores and participants' self-reported weight status, $r(147,538) = .14$, $p < .001$, 95% CI [0.14, 0.15], and with their reports of how others would judge their weight status, $r(114,027) = .16$, $p < .001$, 95% CI [0.16, 0.17]. We also observed small and small-to-medium correlations between self-reported preference scores and participants' self-reported weight status, $r(143,114) = .17$, $p < .001$, 95% CI [0.16, 0.17], and with their reports of how others would judge their weight status, $r(110,446) = .21$, $p < .001$, 95% CI [0.20, 0.21]. Lastly, we observed small correlations between feeling thermometer difference scores and participants' self-reported weight status, $r(146,553) = .15$, $p < .001$, 95% CI [0.14, 0.15], and their reports of how others would judge their weight status, $r(113,334) = .18$, $p < .001$, 95% CI [0.17, 0.18].

8.2.1 | Exploratory analyses

Next, we explored relationships between the two indicators of ingroup typicality and ingroup identification. Participants' self-reported weight status correlated positively with ingroup identification, $r(102,575) = .38$, $p < .001$, 95% CI [0.38, 0.39], indicating that participants identified more strongly with overweight people the higher their self-reported weight status. Furthermore, participants' ratings of how others would judge their weight status correlated positively with their level of ingroup identification, $r(87,373) = .38$, $p < .001$, 95% CI [0.37, 0.38], indicating that participants identified more strongly with overweight people the higher they rated that others would judge their weight status.

We then explored in a series of mediation analyses whether the correlation between typicality and outgroup favoritism was mediated by participants' level of ingroup identification. Both measures of typicality were highly correlated, $r(114,027) = .75$, $p < .001$, 95% CI [0.75, 0.76], so we calculated an ingroup typicality index by averaging the two items.

First, we conducted a mediation analysis with participants' IAT D Scores as dependent variable, ingroup typicality as a predictor, and ingroup identification as a mediator. Indeed, the indirect effect was significant, $b = -.02$, $SE = .00$, $p < .001$, 95% CI [−0.02, −0.02]. This indicates that the correlation between ingroup typicality and outgroup favoritism was partially mediated by participants' level of ingroup identification.

Again, we conducted a mediation analysis, this time using participants' self-reported preference scores as dependent variable, ingroup typicality as a predictor, and ingroup identification as a mediator. The indirect effect was significant, $b = -.16$, $SE = .00$, $p < .001$, 95% CI [−0.17, −0.16]. This indicates that the correlation between ingroup typicality and outgroup favoritism on self-report measures was partially mediated by participants' level of ingroup identification.

Lastly, we conducted a mediation analysis, this time using participants' feeling thermometer difference scores as dependent variable, ingroup typicality as a predictor, and ingroup identification as a mediator. The indirect effect was significant, $b = .34$, $SE = .01$, $p < .001$, 95% CI [0.33, 0.35]. This indicates that the correlation between ingroup typicality and outgroup favoritism on feeling thermometers was partially mediated by participants' level of ingroup identification.

8.3 | Discussion

Results of Study 3 indicate that overweight participants displayed a stronger preference for normal weight relative to overweight individuals the lower their self-reported weight status and the lower their reports of how they believed others would judge their weight status—thus the less typical they appeared for the overweight category. Importantly, this effect was observed for intergroup evaluations on both the IAT and self-report measures. We also observed that both indicators of ingroup typicality were correlated with overweight participants' level of ingroup identification. Lastly, we observed that higher levels of outgroup favoritism were in part due to

the fact that overweight participants who reported being less typical for their group were also less likely to identify with that group.

9 | STUDY 4

Study 4 focuses on yet another domain of intergroup perception—disability. While people might self-categorize as either disabled or abled, people who self-categorize as disabled might still perceive themselves as more or less conforming to a prototypical image of a disabled person, which might in turn influence ingroup and outgroup evaluations.

9.1 | Method

9.1.1 | Participants

Participants were visitors of the Project Implicit demonstration website, who indicated that they had a disability ($N = 35,058$). Given such a large sample size, the achieved power for small effects ($\rho = .1$) was $1–\beta = 1.0$ (two-tailed).

9.1.2 | Measures

Disability IAT
The disability IAT followed the same procedures as IATs in Study 1 and 3 except that target stimuli were symbols indicating disability (e.g., crutches, guide dog, wheelchair) and ability (e.g., persons who walk, run, or ski)—see https://osf.io/y9hiq/, for a complete overview of experimental stimuli and procedures.

Self-reported preference
Participants indicated their relative preference for disabled over abled people on scales from 1 (*strong preference for disabled people*) to 7 (*strong preference for abled people*), with the midpoint indicating no preference.

Perceived ingroup typicality
Participants completed several measures that can be interpreted as proxies of perceived ingroup typicality or similarity to non-disabled people: the ability to hide their disability on a scale from 1 (*impossible to hide*) to 4 (*very able to hide*), how much their disability affected things they do in life on a scale from 1 (*nothing I do*) to 6 (*everything I do*), and the perceived severity of their disability on a scale from 1 (*very slight*) to 5 (*very severe*)—see Table 6 for all response options and descriptive statistics.

9.2 | Results

Overall, participants with disabilities had a positive IAT D Score, which significantly differed from zero, $t(35,057) = 162.38$, $p < .001$,

TABLE 6 Mean IAT *D* Scores, self-reported preference scores, feeling thermometer difference scores, and their standard deviations as a function of three measures of ingroup typicality in Study 4

	IAT			Preference			Thermometer		
	M	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*
Disabled participants	0.37	0.43	147,540	4.73	1.08	143,116	0.47	2.32	146,555
My disability is									
(1) Impossible to hide	0.34	0.49	3,687	4.20	1.21	3,479	−0.24	2.48	3,637
(2) Slightly able to hide	0.40	0.48	5,609	4.19	1.02	5,299	−0.28	1.97	5,529
(3) Moderately able to hide	0.43	0.48	8,783	4.20	0.99	8,353	−0.25	1.88	8,665
(4) Very able to hide	0.43	0.46	13,310	4.30	0.95	12,726	−0.03	1.86	13,174
My disability affects									
(1) Everything I do	0.35	0.49	3,307	4.07	1.25	3,156	−0.52	2.72	3,264
(2) Very many things I do	0.35	0.49	3,918	4.14	1.03	3,716	−0.39	2.00	3,864
(3) Many things I do	0.41	0.47	9,972	4.22	0.96	9,491	−0.21	1.82	9,865
(4) Few things I do	0.44	0.46	8,660	4.29	0.92	8,240	−0.01	1.80	8,548
(5) Very few things I do	0.46	0.46	4,725	4.36	0.99	4,493	0.08	1.80	4,661
(6) Nothing I do	0.43	0.50	735	4.37	1.29	693	0.08	2.44	722
My disability is									
(1) Very severe	0.35	0.52	1,078	4.01	1.49	1,022	−0.41	3.48	1,060
(2) Severe	0.38	0.48	5,859	4.12	1.05	5,565	−0.42	2.06	5,773
(3) Moderate	0.41	0.48	15,110	4.22	0.96	14,380	−0.19	1.85	14,937
(4) Slight	0.45	0.46	6,598	4.35	0.96	6,280	0.04	1.79	6,522
(5) Very slight	0.45	0.45	2,658	4.40	0.99	2,535	0.18	1.94	2,624
Non-disabled participants	0.49	0.43	234,676	4.42	0.97	227,100	0.45	1.98	232,983

Note: IAT = IAT *D* Score; Preference = self-reported preference score; Thermometer = feeling thermometer difference score. Higher means on IATs and self-report measures indicate a stronger preference for non-disabled over disabled individuals. Variable sample sizes are based on missing values in the different dependent variables. As means of comparison, we also report average scores for the non-stigmatized group (i.e., non-disabled participants).

$d_z = .87$, 95% CI [0.85, 0.88], indicating outgroup favoritism on the IAT. Testing participants' mean self-reported preference score against the scale midpoint, $t(33,412) = 23.43$, $p < .001$, $d_z = .13$, 95% CI [0.12, 0.14], and their thermometer difference score against zero, $t(34,580) = −13.24$, $p < .001$, $d_z = −.07$, 95% CI [−0.08, −0.06], indicated small effects of self-reported outgroup and ingroup favoritism, respectively.

Table 7 reports correlations between measures of ingroup typicality and group evaluations. Crucially, we observed correlations between IAT *D* Scores and participants' self-reported ability to hide their disability, $r(31,387) = .06$, $p < .001$, 95% CI [0.05, 0.07], their judgments of how much the disability affects their lives, $r(31,315) = .07$, $p < .001$, 95% CI [0.06, 0.08], and the perceived severity of their disability, $r(31,301) = .05$, $p < .001$, 95% CI [0.04, 0.06], indicating very small correlations between ingroup typicality and outgroup favoritism on the IAT (see Table 6). Next, we correlated participants' self-reported preference and feeling thermometer difference scores with the different proxy variables. We observed similar correlations with the ability to hide the disability,

$r(29,855) = .04$, $p < .001$, 95% CI [0.03, 0.05] and $r(31,003) = .05$, $p < .001$, 95% CI [0.04, 0.06], with the degree to which the disability affects their lives, $r(29,787) = .09$, $p < .001$, 95% CI [0.08, 0.10] and $r(30,922) = .10$, $p < .001$, 95% CI [0.08, 0.11], and with perceived severity, $r(29,780) = .09$, $p < .001$, 95% CI [0.08, 0.11] and $r(30,914) = .09$, $p < .001$, 95% CI [0.08, 0.10], respectively. This indicates that the correlations between measures of ingroup typicality and intergroup evaluations on self-report measures were very small.

9.3 | Discussion

Study 4 provided evidence that disabled participants displayed higher levels of outgroup favoritism the more they were able to hide their disability, the less they felt affected by their disability, and the less severe they judged their disability. Although these effects were substantially smaller than in the previous studies, it is worth highlighting that the pattern of results was consistent across both IAT and self-report measures.

TABLE 7 Zero-order correlations of disabled participants' self-reported measures of ingroup typicality (i.e., the ability to hide their disability; how much their disability affected things they do in life; and the perceived severity of their disability), IAT *D* Scores, self-reported preference scores, feeling thermometer difference scores, and individual feeling thermometers in Study 4

|	1	2	3	4	5	6	7	*M*	*SD*
1. Typicality 1	–							3.01	1.04
2. Typicality 2	0.23***	–						3.31	1.24
3. Typicality 3	0.30***	0.58***	–					3.12	0.93
4. IAT	0.06***	0.07***	0.05***	–				0.41	0.47
5. Preference	0.04***	0.09***	0.09***	0.15***	–			4.13	1.03
6. Thermometer	0.05***	0.10***	0.09***	0.12***	0.46***	–		−0.14	1.98
7. Ingroup	0.04***	0.03***	0.06***	0.08***	0.27***	0.48***	–	3.62	2.21
8. Outgroup	0.00	−0.06***	−0.02***	−0.04***	−0.15***	−0.43***	0.58***	3.76	2.15

Note: Typicality 1 = ability to hide disability; Typicality 2 = extent to which affected by disability; Typicality 3 = perceived severity of disability; IAT = IAT *D* Score; Preference = self-reported preference score; Thermometer = feeling thermometer difference score; Ingroup = ingroup feeling thermometer; Outgroup = outgroup feeling thermometer. Higher means on IATs and self-report measures indicate a stronger preference for disabled individuals relative to non-disabled individuals.

*$p < .05$.

**$p < .01$.

***$p < .001$.

FIGURE 1 Forest plot of random-effects meta-analysis of results from Studies 1–4. Error bars depict 95% confidence intervals
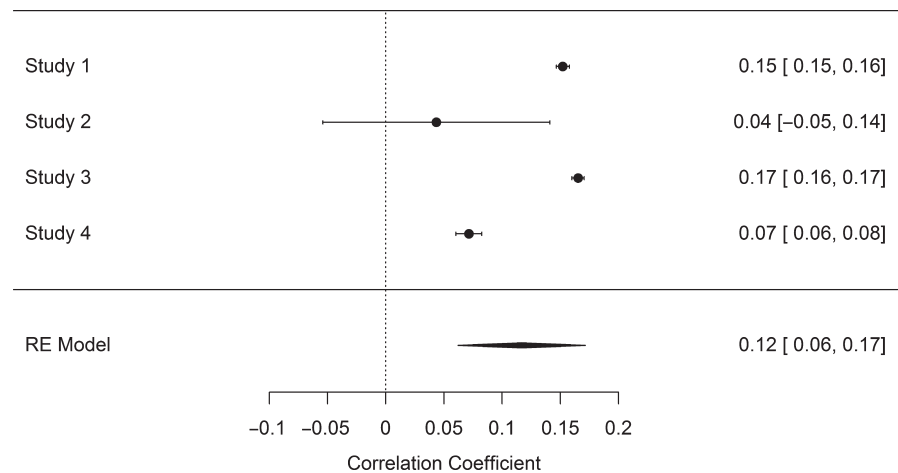
10 | META-ANALYSIS

Studies 1–4 indicate that ingroup typicality is associated with ingroup and outgroup favoritism on IATs and self-report measures of intergroup evaluations. In order to compute a meta-analytic average effect size across studies, we first obtained one mean effect by averaging across *all* effects of each study that were weighted by their respective sample sizes. Next, we conducted a random effects meta-analysis of correlation coefficients across studies, using the metafor package (Viechtbauer, 2010), with effect sizes being weighted by their inverse sampling variance. This analysis yielded an average effect size of $r = .12$, $z = 4.18$, $p < .001$, 95% CI [0.06, 0.17], indicating an overall small effect of ingroup typicality on intergroup evaluations (see Figure 1).[7]

[7]We also observed a substantial amount of heterogeneity, $\tau^2 = .00$, $Q(3) = 228.16$, $p < .001$, accounting for a large proportion of the total variability, $I^2 = 99.36\%$.

In addition to combining effects across relative preference measures (i.e., IATs, self-reported preference scores, and feeling thermometer difference scores), the present data also provide the opportunity to conduct meta-analyses for correlations with single group evaluations (i.e., individual feeling thermometers). Correlations with relative preference measures versus single group evaluations are both informative. Correlations with relative preference measures allow inferences whether ingroup typicality is related to the extent to which people prefer the ingroup relative to the outgroup. In addition, correlations with single group evaluations allow inferences whether ingroup typicality is related to ingroup and/or outgroup evaluations. First, we calculated average effect sizes as described above. Next, we conducted two separate random effects meta-analyses of correlation coefficients for ingroup and outgroup evaluations across studies, with effect sizes being weighted by their inverse sampling variance.

The meta-analysis using ingroup feeling thermometers yielded an average effect size of $r = .07$, $z = 2.11$, $p = .035$, 95% CI [0.00, 0.14]. This indicates that the effect of ingroup typicality on ingroup feeling thermometers was significant but very small. The meta-analysis using outgroup feeling thermometers yielded an average effect size of $r = -.05$, $z = -2.97$, $p = .003$, 95% CI [−0.08, −0.02].[8] This indicates that the effect of ingroup typicality on outgroup feeling thermometers was significant but very small. In sum, meta-analyses of correlations with single group evaluations suggest that higher levels of ingroup typicality are related to more positive ingroup evaluations *and* more negative outgroup evaluations. Consequently, both ingroup and outgroup evaluations contributed to correlations between ingroup typicality and feeling thermometer difference scores.[9]

11 | GENERAL DISCUSSION

The present research investigated whether members of stigmatized groups who appear less typical for their ingroup (e.g., light-skinned Black individuals; individuals with a disability that is less visible) are more likely to distance themselves from their group by evaluating the stigmatized ingroup less positively relative to a non-stigmatized outgroup. Across three social categories, intergroup evaluations on IATs and self-report measures varied with stigmatized group members' relative typicality for the ingroup. In Study 1, Black participants displayed a stronger preference for light-skinned relative to dark-skinned individuals the lighter their self-reported skin tone. This effect was partially replicated in Study 2, with Black participants displaying a stronger preference for Whites relative to Blacks the lighter their other-observed skin tone, but independent of their self-reported skin tone. In Study 3, overweight participants displayed a stronger preference for normal-weight relative to overweight individuals the lower their self-reported weight and the lower they believed others would judge their weight. In Study 4, participants with disabilities displayed a stronger preference for non-disabled relative to disabled individuals the more they reported being able to hide their disability, the less severe they judged their disability, and the less their disability affected their lives. In addition, exploratory analyses revealed that correlations between ingroup typicality and intergroup evaluations were at least partially mediated by stigmatized group members' level of ingroup identification (Studies 2 and 3). Together, these results highlight the role of subjective representations of ingroup typicality as one potential explanation for why members of stigmatized groups may (psychologically) distance themselves from their ingroup, in turn shaping group evaluations.

It is important to highlight that ingroup typicality was related to both ingroup and outgroup evaluations. In other words, to the extent that stigmatized group members displayed lower levels of ingroup typicality they also displayed more negative ingroup evaluations *and* more positive outgroup evaluations. The observed effects of ingroup typicality on intergroup evaluations are consistent across social categories and across measures, but they are small. This might at least in part reflect the use of single-item measures with relatively few response categories (e.g., Loo, 2002; Lozano, García-Cueto, & Muñiz, 2008), and future research might use more reliable multi-item scales. Another possibility, of course, is that the investigated effect itself *is* small. Albeit small, the effect is far from negligible, because even statistically small effects may have large consequences on a societal level—if they apply to many people or if they apply repeatedly to the same individuals (Greenwald, Banaji, & Nosek, 2015). Furthermore, some scholars argue that small effects may have less explanatory power for single events, but are likely to be "consequential in the not-very-long run" (Funder & Ozer, 2019, p. 156). For example, while a person's phenotypic ingroup typicality may not affect *all* their daily social interactions, it may affect the relative frequency of positive or negative experiences with ingroup and outgroup members, thus having a cumulative effect over time. In sum, we observed small effects of ingroup typicality, which may still be consequential at a societal level.

The observed relationships between ingroup typicality and intergroup evaluations fit into the literature on self-group distancing for a number of reasons. In our view, outgroup favoritism can be understood as a way of distancing the self from a stigmatized social identity. This notion of outgroup favoritism as self-group distancing is based on our finding that outgroup favoritism was more prevalent among stigmatized group members who reported lower levels of ingroup typicality. Similarly, previous research on self-group distancing has demonstrated that individuals distance themselves from a stigmatized social identity by perceiving or emphasizing dissimilarities with the ingroup or similarities with a non-stigmatized outgroup (e.g., Derks, Ellemers, et al., 2011; Derks et al., 2015; Weiss & Lang, 2012). Second, the notion of outgroup favoritism as self-group distancing is further corroborated by our finding that outgroup favoritism was more likely among stigmatized group members who reported lower levels of ingroup identification. This negative relationship between outgroup favoritism and ingroup identification is also consistent with the self-group distancing literature, which suggests that self-group distancing is more likely among stigmatized group members who are less identified with their ingroup (Derks, Ellemers, et al., 2011; Derks et al., 2016; Derks, van Laar, et al., 2011). Third, the notion of outgroup favoritism as self-group distancing is consistent with research suggesting that group members who distance themselves from a stigmatized social identity may endorse negative ingroup stereotypes and display negative ingroup evaluations (e.g., Derks, Ellemers, et al., 2011; Guimond et al., 2002), which are arguably related to outgroup favoritism (see Tables 2, 3, 5

---

[8]We observed substantial amounts of between-study heterogeneity for both ingroup evaluations, $\tau^2 = .00$, $Q(3) = 212.87$, $p < .001$, $I^2 = 99.57\%$, and outgroup evaluations, $\tau^2 = .00$, $Q(3) = 176.07$, $p < .001$, $I^2 = 97.72\%$.

[9]In addition to examining whether ingroup typicality is related to ingroup and/or outgroup evaluations, an important question is whether the magnitude of effect sizes differs for ingroup and outgroup evaluations. Unfortunately, due to the relatively small number of studies, the present data do not provide sufficient power to detect differences between the two subgroups of studies (Harrer, Cuijpers, Furukawa, & Ebert, 2019; Higgins & Thompson, 2004).

and 7 for correlations between group evaluations and ingroup favoritism). Taken together, we argue that outgroup favoritism is a form of self-group distancing because of its relations with ingroup typicality, ingroup identification, and group evaluations.

That said, there is one noteworthy conceptual difference between our interpretation of the present findings and our reading of the self-group distancing literature. Our reading of this literature is that it conceptualizes ingroup identification as a *moderator* of self-group distancing. According to this view, self-group distancing should be more likely among low identified stigmatized group members and less likely among high identified stigmatized group members (e.g., Derks, Ellemers, et al., 2011; Derks, van Laar, et al., 2011). Because the self-group distancing literature views ingroup identification as a moderator, such research seems less focused on explaining why stigmatized group members display varying levels of ingroup identification in the first place. The present findings extend the self-group distancing literature by suggesting that (lower) identification with the ingroup may itself be one possible consequence of (lower) ingroup typicality. Thus, to the extent that ingroup typicality is based on phenotypic appearance (e.g., variations in skin tone or facial features; weight status; the visibility of a disability), it may influence stigmatized group members' tendency to identify with their ingroup (e.g., Brown et al., 1999; Wilkins et al., 2010), with potential downstream consequences for group evaluations. According to this alternative view, ingroup identification is conceptualized as a *mediator* of the relationship between ingroup typicality and group evaluations. This view is consistent with our finding that ingroup identification partially mediated the effects of ingroup typicality on intergroup evaluations (Studies 2 and 3).

Think back to our opening example of Tyree and Jamal, who might not be equally categorized by others as Black. We argue that these differences in (perceived) ingroup typicality might feed back into stigmatized group members' tendency to self-categorize as an ingroup member and to identify with the ingroup. Hence, group members might experience more or less flexibility to identify themselves with or distance themselves from the ingroup, eventually affecting their evaluations of the ingroup and outgroups. In sum, we argue that phenotypic appearance places a boundary on stigmatized group members' ingroup typicality, affecting the tendency to psychologically connect with the ingroup, and eventually affecting group evaluations. Future longitudinal research would seem best positioned to establish these hypothesized causal relationships between ingroup typicality, ingroup identification, and group evaluations.

In addition to ingroup identification, our explanation of the relations between ingroup typicality and group evaluations centers on the perceived permeability of group boundaries and individual mobility (Tajfel & Turner, 1979). In our view, ingroup typicality might introduce an *individual constraint* with implications for the perceived permeability of group boundaries for members of stigmatized groups. For example, varying perceived ingroup typicality might influence the likelihood for a given member of a stigmatized group to be categorized, and thus to self-categorize, as a group member. In other words, to the extent that ingroup typicality reflects phenotypic

appearance, it might affect a person's likelihood of being perceived and to perceive themselves as a group member. These speculations are consistent with social identity theory (Tajfel & Turner, 1979), which highlights that individual mobility necessarily implies that stigmatized group members dis-identify with the ingroup. In fact, previous research has documented that individual mobility is related to ingroup identification, and that those who anticipate upward mobility also tend to have more negative attitudes regarding the ingroup (e.g., Chipeaux, Kulich, Iacoviello, & Lorenzi-Cioldi, 2017).

Lastly, low perceived ingroup typicality might also limit stigmatized group members' ability to form social bonds with fellow ingroup members. Group members who are perceived, or who perceive themselves, as less typical, may have more difficulty forming connections with other ingroup members. For example, previous research suggests that differences in skin tone also play an important role within Black communities (e.g., Harvey, Tennial, & Hudson Banks, 2017), affecting group members' feelings of acceptance. One study found that Black university students with darker skin tone also felt more accepted by their Black peers compared to lighter skinned students (Harvey et al., 2005). Hence, in addition to constraining the permeability of group boundaries, ingroup typicality might also constrain stigmatized group members' ability to form attachment with fellow ingroup members or with the ingroup as a whole (i.e., ingroup identification).

However, we recognize that we cannot draw firm conclusions regarding the causal nature of the relationship between ingroup typicality, ingroup identification, and group evaluations. Thus, our data cannot confirm the implied causal model, where ingroup typicality provides an individual constraint for ingroup identification and the permeability of group boundaries. Furthermore, prominent theorizing suggests that constructs with similarities to ingroup typicality are in fact sub-components of identification. In particular, Leach et al. (2008) suggest that the extent to which individuals see themselves as similar to the ingroup (i.e., self-stereotyping) is one facet of ingroup identification. Thus, this model would at least suggest bi-directional relationships between self-stereotyping and other forms of identification. Moreover, it is certainly possible that causality might work the other way around. For example, stigmatized group members who are less identified might perceive themselves as less typical for the ingroup; or stigmatized group members who evaluate the ingroup more negatively might be less identified *and* perceive themselves as less typical. These outlined alternative relations point to the need for future research. Ideally, future research needs to investigate causal effects, for example through experimentally varying perceptions of ingroup typicality (e.g., by changing characteristics of ingroup and/or outgroup prototypes) and directly measuring perceptions of group boundary permeability, thus furthering our understanding of the psychological processes underlying the effects of ingroup typicality on intergroup evaluations.

Previous research has spent a lot of effort investigating how members of advantaged groups perceive, judge, and treat stigmatized group members; at the same time, research has largely

overlooked how perceptions and attitudes within stigmatized groups might differ due to within-group differences, and how these may affect intergroup relations. The observed consistent pattern of effects for three different and unrelated social categories and across different measures indicates that research on ingroup typicality may be a promising approach to study why stigmatized group members sometimes do or do not display ingroup favoritism.

CONFLICT OF INTEREST

The authors confirm they have no conflict of interest to declare. The authors also confirm that this article adheres to the ethical guidelines specified in the APA Code of Conduct as well as the authors' national ethics guidelines

AUTHORS' CONTRIBUTION

I.E., J.D., and S.O. developed the study concept. I.E. analyzed and interpreted the data under the supervision of J.D. and S.O. I.E. provided the first draft of the manuscript. S.O. and J.D. provided critical revisions of the manuscript. All authors approved the final version of the manuscript for submission. We confirm that the research was conducted ethically, results are reported honestly, the submitted work is original and not (self-)plagiarized, and authorship reflects the authors' contributions.

TRANSPARENCY STATEMENT

Analyses scripts are accessible via Open Science Framework (https://osf.io/z4xwx/)

ORCID

Iniobong Essien [ID] https://orcid.org/0000-0003-0277-7908
Sabine Otten [ID] https://orcid.org/0000-0001-5369-7241
Juliane Degner [ID] https://orcid.org/0000-0002-6634-8099

REFERENCES

Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. Retrieved from https://github.com/crsh/papaja

Axt, J. R., Ebersole, C. R., & Nosek, B. A. (2014). The rules of implicit evaluation by race, religion, and age. *Psychological Science*, 25(9), 1804–1815. https://doi.org/10.1177/0956797614543801

Bishop, C. J., Kiss, M., Morrison, T. G., Rushe, D. M., & Specht, J. (2014). The association between gay men's stereotypic beliefs about drag queens and their endorsement of hypermasculinity. *Journal of Homosexuality*, 61(4), 554–567. https://doi.org/10.1080/00918 369.2014.865464

Blair, I. V., Judd, C. M., Sadler, M. S., & Jenkins, C. (2002). The role of Afrocentric features in person perception: Judging by features and categories. *Journal of Personality and Social Psychology*, 83(1), 5–25. https://doi.org/10.1037//0022-3514.83.1.5

Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity* (pp. 35–58). Oxford, UK: Blackwell Science.

Branscombe, N. R., Schmitt, M. T., & Harvey, R. D. (1999). Perceiving pervasive discrimination among African Americans: Implications for group identification and well-being. *Journal of Personality and Social Psychology*, 77(1), 135–149. https://doi.org/10.1037/0022-3514.77.1.135

Brown, K. T., Ward, G. K., Lightbourn, T., & Jackson, J. S. (1999). Skin tone bias and racial identity among African Americans: A theoretical and research framework. In R. L. Jones (Ed.), *Advances in African American psychology* (pp. 191–215). Hampton, VA: Cobb Publishers.

Brown-Iannuzzi, J. L., Lundberg, K. B., Kay, A. C., & Payne, B. K. (2015). Subjective status shapes political preferences. *Psychological Science*, 26(1), 15–26. https://doi.org/10.1177/0956797614553947

Chipeaux, M., Kulich, C., Iacoviello, V., & Lorenzi-Cioldi, F. (2017). I want, therefore I am: Anticipated upward mobility reduces ingroup concern. *Frontiers in Psychology*, 8, 1451. https://doi.org/10.3389/fpsyg.2017.01451

Clarkson, J. (2006). Everyday joe versus pissy, bitchy, queens: Gay masculinity on straightacting.com. *The Journal of Men's Studies*, 14(2), 191–207. https://doi.org/10.3149/jms.1402.191

Conigrave, J. (2019). *Corx: Create and format correlation matrices*. Retrieved from https://CRAN.R-project.org/package=corx

Corneille, O., & Hütter, M. (2020). Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. *Personality and Social Psychology Review*, 24(3), 212–232. https://doi.org/10.1177/1088868320911325

Crandall, C. S. (1994). Prejudice against fat people: Ideology and self-interest. *Journal of Personality and Social Psychology*, 66(5), 882–894. https://doi.org/10.1037/0022-3514.66.5.882

Davies, P. G., Hutchinson, S., Osborne, D., & Eberhardt, J. L. (2016). Victims' race and sex leads to eyewitness misidentification of perpetrator's phenotypic stereotypicality. *Social Psychological and Personality Science*, 7(6), 491–499. https://doi.org/10.1177/19485 50616644655

de Lemus, S., & Stroebe, K. (2015). Achieving social change: A matter of all for one? *Journal of Social Issues*, 71(3), 441–452. https://doi.org/10.1111/josi.12122

Degner, J., Essien, I., & Reichardt, R. (2016). Effects of diversity versus segregation on automatic approach and avoidance behavior towards own and other ethnic groups. *European Journal of Social Psychology*, 46(6), 783–791. https://doi.org/10.1002/ejsp.2234

Degner, J., & Wentura, D. (2009). Not everybody likes the thin and despises the fat: One's weight matters in the automatic activation of weight-related social evaluations. *Social Cognition*, 27(2), 202–221. https://doi.org/10.1521/soco.2009.27.2.202

Derks, B., Ellemers, N., van Laar, C., & de Groot, K. (2011). Do sexist organizational cultures create the queen bee? *The British Journal of Social Psychology*, 50(3), 519–535. https://doi.org/10.1348/014466610X 525280

Derks, B., van Laar, C., & Ellemers, N. (2016). The queen bee phenomenon: Why women leaders distance themselves from junior women. *The Leadership Quarterly*, 27(3), 456–469. https://doi.org/10.1016/j.leaqua.2015.12.007

Derks, B., van Laar, C., Ellemers, N., & de Groot, K. (2011). Gender-bias primes elicit queen-bee responses among senior policewomen. *Psychological Science*, 22(10), 1243–1249. https://doi.org/10.1177/0956797611417258

Derks, B., van Laar, C., Ellemers, N., & Raghoe, G. (2015). Extending the queen bee effect: How Hindustani workers cope with disadvantage by distancing the self from the group. *Journal of Social Issues*, 71(3), 476–496. https://doi.org/10.1111/josi.12124

Dixon, T. L., & Maddox, K. B. (2005). Skin tone, crime news, and social reality judgments: Priming the stereotype of the dark and dangerous

black criminal. *Journal of Applied Social Psychology*, *35*(8), 1555–1570. https://doi.org/10.1111/j.1559-1816.2005.tb02184.x

Dowle, M., & Srinivasan, A. (2017). Data.table: Extension of 'data.frame'. Retrieved from https://CRAN.R-project.org/package=data.table

Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy. *Psychological Science*, *17*(5), 383–386. https://doi.org/10.1111/j.1467-9280.2006.01716.x

Eguchi, S. (2009). Negotiating hegemonic masculinity: The rhetorical strategy of "straight-acting" among gay men. *Journal of Intercultural Communication Research*, *38*(3), 193–209. https://doi.org/10.1080/17475759.2009.508892

Essien, I., Calanchini, J., & Degner, J. (2020). Moderators of intergroup evaluation in disadvantaged groups: A comprehensive test of predictions from system justification theory. *Journal of Personality and Social Psychology*. Advance online publication. http://doi.org/10.1037/pspi0000302

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). Irr: Various coefficients of interrater reliability and agreement. Retrieved from https://CRAN.R-project.org/package=irr

Gawronski, B., Brochu, P. M., Sritharan, R., & Strack, F. (2012). Cognitive consistency in prejudice-related belief systems: Integrating old-fashioned, modern, aversive, and implicit forms of prejudice. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency: A fundamental principle in social cognition* (pp. 369–389). New York, NY: Guilford Press.

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the implicit association test can have societally large effects. *Journal of Personality and Social Psychology*, *108*(4), 553–561. https://doi.org/10.1037/pspa0000016

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Greenwald, A. G., & Pettigrew, T. F. (2014). With malice toward none and charity for some: Ingroup favoritism enables discrimination. *American Psychologist*, *69*(7), 669–684. https://doi.org/10.1037/a0036056

Guimond, S., Dif, S., & Aupy, A. (2002). Social identity, relative group status and intergroup attitudes: When favorable outcomes change intergroup relations… For the worse. *European Journal of Social Psychology*, *32*(6), 739–760. https://doi.org/10.1002/ejsp.118

Hagiwara, N., Kashy, D. A., & Cesario, J. (2012). The independent effects of skin tone and facial features on whites' affective reactions to blacks. *Journal of Experimental Social Psychology*, *48*(4), 892–898. https://doi.org/10.1016/j.jesp.2012.02.001

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. (2019). Doing meta-analysis in R: A hands-on Guide. https://doi.org/10.5281/zenodo.2551803

Harvey, R. D., LaBeach, N., Pridgen, E., & Gocial, T. M. (2005). The intragroup stigmatization of skin tone among black Americans. *Journal of Black Psychology*, *31*(3), 237–253. https://doi.org/10.1177/0095798405278192

Harvey, R. D., Tennial, R. E., & Hudson Banks, K. (2017). The development and validation of a colorism scale. *Journal of Black Psychology*, *43*(7), 740–764. https://doi.org/10.1177/0095798417690054

Hebl, M. R., Williams, M. J., Sundermann, J. M., Kell, H. J., & Davies, P. G. (2012). Selectively friending: Racial stereotypicality and social rejection. *Journal of Experimental Social Psychology*, *48*(6), 1329–1335. https://doi.org/10.1016/j.jesp.2012.05.019

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, *53*, 575–604. https://doi.org/10.1146/annurev.psych.53.100901.135109

Higgins, J. P. T., & Thompson, S. G. (2004). Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine*, *23*(11), 1663–1682. https://doi.org/10.1002/sim.1752

Hunter, J. E., & Schmidt, F. L. (2007). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Kahn, K. B., & Davies, P. G. (2011). Differentially dangerous? Phenotypic racial stereotypicality increases implicit bias among ingroup and outgroup members. *Group Processes & Intergroup Relations*, *14*(4), 569–580. https://doi.org/10.1177/1368430210374609

Kelley, K. (2018). *MBESS: The mbess r package*. Retrieved from https://CRAN.R-project.org/package=MBESS

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., … Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, *95*(1), 144–165. https://doi.org/10.1037/0022-3514.95.1.144

Loo, R. (2002). A caveat on using single–item versus multiple–item scales. *Journal of Managerial Psychology*, *17*(1), 68–75. https://doi.org/10.1108/02683940210415933

Lozano, L. M., García-Cueto, E., & Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, *4*(2), 73–79. https://doi.org/10.1027/1614-2241.4.2.73

Ma, D. S., & Correll, J. (2011). Target prototypicality moderates racial bias in the decision to shoot. *Journal of Experimental Social Psychology*, *47*(2), 391–396. https://doi.org/10.1016/j.jesp.2010.11.002

Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review*, *8*(4), 383–401. https://doi.org/10.1207/s15327957pspr0804_4

Massey, D. S., & Martin, J. A. (2003). *The NIS skin color scale*. Princeton, NJ: Princeton University Press.

Meeus, J., Mayor, J. P., González, R., Brown, R., & Manzi, J. (2017). Racial phenotypicality bias in educational expectations for both male and female teenagers from different socioeconomic backgrounds. *European Journal of Social Psychology*, *47*(3), 289–303. https://doi.org/10.1002/ejsp.2247

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, *22*(2), 103–122. https://doi.org/10.1002/ejsp.2420220202

Müller, K. (2017). *Here: A simpler way to find your files*. Retrieved from https://CRAN.R-project.org/package=here

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, *6*(1), 101–115. https://doi.org/10.1037/1089-2699.6.1.101

Phills, C. E., Hahn, A., & Gawronski, B. (2020). The bidirectional causal relation between implicit stereotypes and implicit prejudice. *Personality and Social Psychology Bulletin*, *46*, 1318–1330. https://doi.org/10.1177/0146167219899234

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rae, J. R., Newheiser, A.-K., & Olson, K. R. (2015). Exposure to racial out-groups and implicit race bias in the united states. *Social Psychological and Personality Science*, *6*(5), 535–543. https://doi.org/10.1177/1948550614567357

Reimer, N. K., Becker, J. C., Benz, A., Christ, O., Dhont, K., Klocke, U., … Hewstone, M. (2016). Intergroup contact and social change: Implications of negative and positive contact for collective action in advantaged and disadvantaged groups. *Personality and Social*

Psychology Bulletin, 43(1), 121–136. https://doi.org/10.1177/0146167216676478

Rudman, L. A., Feinberg, J., & Fairchild, K. (2002). Minority members' implicit attitudes: Automatic ingroup bias as a function of group status. *Social Cognition, 20*(4), 294–320. https://doi.org/10.1521/soco.20.4.294.19908

Schmitt, M. T., Branscombe, N. R., Postmes, T., & Garcia, A. (2014). The consequences of perceived discrimination for psychological well-being: A meta-analytic review. *Psychological Bulletin, 140*(4), 921–948. https://doi.org/10.1037/a0035754

Schwartz, M. B., Vartanian, L. R., Nosek, B. A., & Brownell, K. D. (2006). The influence of one's own body weight on implicit and explicit anti-fat bias. *Obesity, 14*(3), 440–447. https://doi.org/10.1038/oby.2006.58

Selker, R., Love, J., & Dropmann, D. (2018). Jmv: The 'jamovi' analyses. Retrieved from https://CRAN.R-project.org/package=jmv

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In S. Worchel & W. G. Austin (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Todd, A. R., & Burgmer, P. (2013). Perspective taking and automatic intergroup evaluation change: Testing an associative self-anchoring account. *Journal of Personality and Social Psychology, 104*(5), 786–802. https://doi.org/10.1037/a0031999

Viechtbauer, W. (2010). Conducting meta-analyses in r with the meta-for package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Weiss, D., & Freund, A. M. (2012). Still young at heart: Negative age-related information motivates distancing from same-aged people.

Psychology and Aging, 27(1), 173–180. https://doi.org/10.1037/a0024819

Weiss, D., & Lang, F. R. (2012). They are old but I feel younger: Age-group dissociation as a self-protective strategy in old age. *Psychology and Aging, 27*(1), 153–163. https://doi.org/10.1037/a0024887

Wickham, H. (2017). *Tidyverse: Easily install and load the "tidyverse".* Retrieved from https://CRAN.R-project.org/package=tidyverse

Wilkins, C. L., Kaiser, C. R., & Rieck, H. (2010). Detecting racial identification: The role of phenotypic prototypicality. *Journal of Experimental Social Psychology, 46*(6), 1029–1034. https://doi.org/10.1016/j.jesp.2010.05.017

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Boca Raton, FL: Chapman; Hall/CRC. Retrieved from https://yihui.name/knitr/.

Xu, F. K., Nosek, B. A., Greenwald, A. G., Ratliff, K. A., Bar-Anan, Y., Umansky, E., Smith, C. (2017). Project implicit demo website datasets. https://doi.org/10.17605/OSF.IO/Y9HIQ