UC Riverside

UC Riverside Previously Published Works

Title

Moderators of intergroup evaluation in disadvantaged groups: A comprehensive test of predictions from system justification theory.

Permalink

https://escholarship.org/uc/item/07p7q06m

Journal

Journal of personality and social psychology, 120(5)

ISSN

0022-3514

Authors

Essien, Iniobong
Calanchini, Jimmy
Degner, Juliane

Publication Date

2021-05-01

DOI

10.1037/pspi0000302

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Title page with all author information

Moderators of Intergroup Evaluation in Disadvantaged Groups: A Comprehensive Test of Predictions from System Justification Theory

Iniobong Essien[1,2], Jimmy Calanchini[3], & Juliane Degner[2]

[1] FernUniversität in Hagen

[2] Universität Hamburg

[3] University of California Riverside

Author Note

'Department of Psychology, Community Psychology, FernUniversität in Hagen, Germany.' 'Institute of Psychology, Universität Hamburg, Germany.' Department of Psychology, University of California, Riverside, CA, USA.

The first author wishes to thank Marleen Stelter for her insights and helpful feedback on previous drafts of this manuscript and Carsten Sander for help with additional data preparation and literature search.

Correspondence concerning this article should be addressed to Iniobong Essien, Department of Psychology, Community Psychology, FernUniversität in Hagen, 58084 Hagen, Germany. E-mail: iniobong.essien@fernuni-hagen.de

Masked Manuscript without Author Information

Abstract

We examined hypotheses proposed by System Justification Theory (SJT; Jost, Banaji, & Nosek, 2004) regarding intergroup evaluation in disadvantaged groups, using large samples of online participants (total N = 715,721), spanning eight intergroup domains and 14 nations. Using a meta-analytic approach, we tested these hypotheses at the individual level (as SJT is generally articulated), as well as at the social group level. Consistent with SJT, individual-level analyses revealed that disadvantaged groups demonstrated outgroup favoritism on IATs (i.e., implicit measures), but demonstrated ingroup favoritism or no intergroup preference on self-report (i.e., explicit) measures. Additionally, these average effects were characterized by high heterogeneity, and follow-up exploratory analyses revealed that intergroup evaluation in disadvantaged groups was moderated by the intergroup domain: Whereas some disadvantaged groups consistently displayed outgroup favoritism (e.g., age; weight), others consistently displayed ingroup favoritism (e.g., sexual orientation; religion), and yet others displayed diverging patterns on implicit and explicit measures (e.g., race; ethnicity). Consistent with SJT, intergroup evaluation on all measures was moderated by self-reported conservatism. Furthermore, the magnitude of these relationships depended on the level of analysis, with small effects emerging at the individual level and medium-sized effects emerging at the social group level. Social group-level analyses also indicated that intergroup evaluation in disadvantaged groups was moderated by stigma. Overall, these findings support and extend the predictions of SJT, but the relatively complex patterns of intergroup evaluation in disadvantaged groups identified here illustrate a need for further theory development and more theory-driven research in this domain.

Keywords: system justification theory; disadvantaged groups; stigma; implicit bias; explicit bias

Word count: 15,889 (excluding references)

Moderators of Intergroup Evaluation in Disadvantaged Groups: A Comprehensive Test of Predictions from System Justification Theory

> "Ask yourself what would happen to your own personality if you heard it said over and over again that you were lazy, a simple child of nature, expected to steal, and had inferior blood. Suppose this opinion were forced on you by the majority of your fellow-citizens. And suppose nothing that you could do would change this opinion-because you happen to have black skin." (Allport, 1954, p. 142)

In his classic work *The Nature of Prejudice*, Allport (1954) speculated about the consequences of being socialized in a society with a predominantly negative attitude towards one's own social group. He reasoned that societal devaluation of one's ingroup might—among other things—lead to feelings of insecurity, rejection of ingroup membership, identification with dominant outgroups, self-hate, or even aggression against the ingroup. Similar ideas were formulated by Clark and Clark (1950) who reasoned that Black children, early in their development, become aware of the inferior status position of their ingroup in society, leading to "feelings of inadequacy and inferiority" (Clark & Clark, 1950, p. 350). In their seminal doll studies, they observed that a substantial number of Black children displayed *outgroup favoritism*, preferring White dolls over Black dolls. More recently, System Justification Theory (SJT; Jost et al., 2004) posited that such feelings of inferiority are nursed by a general motive to accept the current state of affairs, which in turn gives rise to *false consciousness*: an internalization of stigma by members of disadvantaged groups.

The present research investigates intergroup evaluations across a wide range of disadvantaged groups[1]. Specifically, we provide a large-scale test of hypotheses proposed by Jost and colleagues (2004), who suggested that intergroup evaluation in disadvantaged

[1] The SJT key publication (Jost et al., 2004) uses the terms *disadvantaged* and *low-status* interchangeably. Here, we use *disadvantaged* because it can be broadly applied to groups that are stigmatized, numerical minorities, of lower socio-economic status, and lacking access to resources.

groups should be moderated by a number of factors, including aspects of the measurement procedure, system justification tendencies, political ideology, and stigma. The present research advances the literature by testing these hypotheses at the individual level—as SJT is primarily articulated—as well as at the level of the social group, thereby extending SJT with novel tests of generalizability and boundary conditions.

System Justification Theory

Decades of social psychology research have shown that people often think, feel, and act in self-interested ways (e.g., Miller, 1999) or in ways that serve the interests of the social groups they belong to (Greenwald & Pettigrew, 2014; Tajfel & Turner, 1979). Additionally, according to SJT (Jost & Banaji, 1994; Jost et al., 2004), people are also generally motivated to justify and defend the social systems in which they live. However, many societies are characterized by high levels of inequality, and some social groups are more advantaged than other social groups in terms of status, power, and access to resources. Consequently, for members of advantaged groups, system justification motives (e.g., to support the status quo) are congruent with motivations to see themselves and the ingroup positively (e.g., self-esteem; group pride); but for members of disadvantaged groups, system justification motives are at odds with motivations to see themselves or the ingroup positively (Jost, 2019; Jost, Gaucher, & Stern, 2015; Jost & van der Toorn, 2012). SJT was developed, in part, as an advancement of Social Identity Theory (Tajfel & Turner, 1979), to better account for findings of outgroup favoritism in disadvantaged groups (Jost, Burgess, & Mosso, 2001). According to SJT (Jost & Banaji, 1994), system justifying beliefs often contain stereotypes that offer ostensible explanations why social groups inhabit their positions in society. Stereotypes not only describe the characteristics of social groups, but also justify the higher status of dominant groups (e.g., “they hold high positions by virtue”) and the lower status of disadvantaged groups (e.g., “they are poorly off because they did not work hard enough”).

In SJT's initial formulation, Jost and Banaji (1994) emphasized that stereotypes about social groups help to rationalize the social order and legitimize the positions that social groups hold in society. Although stereotypes of disadvantaged groups are often negative, members of disadvantaged groups are assumed to accept and internalize beliefs that are detrimental to one's own or the ingroup's interests but legitimize the status quo (i.e., false consciousness; Jost & Banaji, 1994). In the course of socialization, stereotypical beliefs become deeply entrenched and highly accessible to both advantaged and disadvantaged group members (Jost et al., 2002; Mentovich & Jost, 2008).

SJT posits that even those who are clearly disadvantaged by their positions in society are motivated to justify the status quo because system justifying beliefs serve a palliative function and make people feel better about how things are (Jost & Hunyady, 2002). Acknowledging that the social system is unfair or unjust is emotionally taxing, especially for those who are disadvantaged by the system (Jost, Wakslak, & Tyler, 2008). Consequently, SJT assumes that disadvantaged group members internalize society's negative perceptions of their ingroup to regulate negative emotions about the unfairness of society.

Predictions Regarding Outgroup Favoritism

In articulating SJT as a framework for understanding outgroup favoritism in disadvantaged groups, Jost and colleagues (2004) proposed a number of hypotheses specifying the conditions under which members of disadvantaged groups should be especially likely to display outgroup favoritism. Of these hypotheses, we test two in the context of the present research. First, outgroup favoritism in members of disadvantaged groups is more likely observed on implicit than explicit measures. Second, outgroup favoritism is more likely when disadvantaged group members endorse system justifying beliefs. Additionally, a typical reading of SJT (Jost & Banaji, 1994) suggests a third, testable hypothesis: that disadvantaged group members' evaluations of their own group are

more negative to the extent that their group is viewed negatively (i.e., stigmatized) by society (e.g., Dasgupta, 2004; Lane, Mitchell, & Banaji, 2005; Livingston, 2002; Rudman, Feinberg, & Fairchild, 2002). In the following section, we elaborate on each of these hypotheses and review existing empirical evidence for each hypothesis.

Implicit versus explicit measures. Previous research has revealed that disadvantaged groups' intergroup evaluations are moderated by measurement method. Intergroup evaluations can be measured *explicitly*, often through direct self-report, as well as *implicitly*, based on the interpretation of speed or accuracy of responses rather than the contents of the response, *per se*.[2] Implicit measures were introduced, in part, because explicit measures have been shown to be susceptible to socially desirable responding (Gawronski & Hahn, 2019). Implicit measures are thought to circumvent self-presentation and social desirability through task procedures designed to minimize the extent to which people can deliberately feign responses (but see Czellar, 2006; Fiedler & Bluemke, 2005; Steffens, 2004). Nosek, Banaji, and Greenwald (2002) examined intergroup evaluation among members of disadvantaged groups who completed both implicit and explicit measures at the Project Implicit demonstration website. Black participants displayed a large effect of ingroup favoritism on the explicit measure, but a small effect of outgroup favoritism on the implicit measure. Older participants demonstrated an even more pronounced divergence by measurement method, with a small effect of outgroup favoritism on the explicit measure but a large effect of outgroup favoritism on the implicit measure.

In response to such demonstrations of implicit outgroup favoritism but explicit ingroup favoritism among disadvantaged groups, Jost and colleagues (2004) reasoned that outgroup favoritism in disadvantaged groups should be more likely to be observed on implicit than on explicit measures (Hypothesis 6'; Jost et al., 2004, p. 893). They argued

[2] Throughout this paper, we use the term "implicit" in reference to indirect measurement tools ("implicit measures") and their behavioral outcomes ("implicit evaluations" and "implicit bias"). Thus, our use of the term "implicit" does not make assumptions about underlying mental representations or process characteristics (Corneille & Hütter, 2020).

that members of disadvantaged groups feel intense social pressures to show ingroup pride, so they should be reluctant to openly endorse beliefs that disfavor the ingroup. Because disadvantaged group members would not want to be seen as identifying with the dominant outgroup, they display ingroup favoritism under conditions that readily allow for such deliberate responding, i.e., explicit measures. However, disadvantaged group members' internalized negativity may be more readily expressed under conditions that constrain deliberate responding, i.e., implicit measures. Taken together, SJT (Jost et al., 2004) predicts that implicit measures rather than explicit measures should more readily reveal outgroup favoritism in disadvantaged groups because they minimize socially-desirable responding.

As evidence of implicit outgroup favoritism in disadvantaged groups, a substantial body of work in the SJT tradition (Jost et al., 2004, 2015; Jost & van der Toorn, 2012; Mentovich & Jost, 2008) relies upon research in which disadvantaged group members complete an Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) measuring their evaluative preferences for their own group relative to an advantaged outgroup (Ashburn-Nardo, Knowles, & Monteith, 2003; Jost & Hunyady, 2002; Livingston, 2002; Nosek et al., 2002; Rudman et al., 2002; Uhlmann, Dasgupta, Elgueta, Greenwald, & Swanson, 2002). In line with the predictions of SJT, some studies provide evidence of outgroup favoritism among disadvantaged groups. For example, Black participants displayed a preference for White relative to Black targets (i.e., outgroup favoritism) on an evaluative IAT (Ashburn-Nardo et al., 2003; see also Nosek et al., 2002). Another study reported the opposite result: Black participants displayed a preference for Black relative to White targets (i.e., ingroup favoritism) on an evaluative IAT, albeit to a lower degree than on an explicit feeling thermometer measure (Livingston, 2002, Experiment 1b). Still other studies have found more nuanced results in other intergroup domains. Students from a prestigious university demonstrated implicit ingroup favoritism, but students from a less prestigious university demonstrated no evaluative preference, i.e., neither ingroup nor

outgroup favoritism (Jost et al., 2002, Study 1). A similar pattern was observed among Hispanic participants, who did not display any intergroup preference on two evaluative Hispanic-White IATs (Uhlmann et al., 2002). Taken together, some studies provide evidence that disadvantaged groups display outgroup favoritism on implicit measures (Jost et al., 2004), but other studies report either no preference or ingroup favoritism among disadvantaged group members on implicit measures.

System justifying tendencies and ideological beliefs. The second prediction articulated by Jost and colleagues is that disadvantaged group members are more likely to display outgroup favoritism to the extent that they exhibit system justifying tendencies (Hypothesis 8; Jost et al., 2004, p. 901). System justifying tendencies have been operationalized in a number of ways. For example, Jost and Thompson (2000) developed a scale to assess *economic system justification*, which measures beliefs about (in)equality and the (un)fairness of the economic system. System justifying tendencies have also been operationalized in terms of social dominance orientation (SDO; Pratto, Sidanius, Stallworth, & Malle, 1994).[3]

System justifying tendencies can also manifest as ideological beliefs (Jost, 2019). Specifically, SJT predicts that disadvantaged group members should be more likely to display outgroup favoritism to the extent that they hold conservative beliefs (Hypothesis 8'; Jost et al., 2004, p. 901). Jost and colleagues (2003) reasoned that conservatism is comprised of two potentially interrelated core aspects: the tendency to oppose change and maintain the status quo, and the preference for inequality. These two core aspects of conservatism—opposition to change and preference for inequality—are often correlated because, in unequal societies, opposing change usually implies sustaining inequality and keeping traditionally advantaged groups in power (Jost et al., 2003). As such, conservatism can be regarded as a system justifying belief because it provides an intellectual basis for rationalizing the current state of affairs (see Jost, 2019, for an overview of studies on the

[3] See Jost and Hunyady (2005) for a discussion of constructs related to system justifying tendencies.

relationship between system justification and ideological beliefs). In short, because conservatism is regarded as a system justifying ideology, disadvantaged group members should display outgroup favoritism to the extent that they hold conservative beliefs.

Previous research has examined relations between conservatism and intergroup evaluations in dominant groups (e.g., Cunningham, Nezlek, & Banaji, 2004) but, to our knowledge, only a handful of studies have directly investigated these relationships in disadvantaged groups. Hoffarth and Jost (2017) re-analyzed data from homosexual and bisexual participants and discovered that higher conservatism was related to less favorable evaluations of Gay people relative to Straight people on an IAT. Similarly, Jost and colleagues (2004) found that gay and lesbian participants displayed more outgroup favoritism on both IATs and self-report measures to the extent that they self-identified as conservative. However, a different pattern of results for Black and older participants. Black participants demonstrated a positive relationship between conservatism and explicit outgroup favoritism, but no relationship between conservatism and implicit outgroup favoritism, and intergroup evaluations were unrelated to conservatism among older participants. These findings illustrate heterogeneous effects of conservatism on intergroup evaluation in disadvantaged groups. Consequently, open questions remain regarding whether conservatism is related to outgroup favoritism in some groups but not others, as well as whether these effects might differ for implicit versus explicit measures of intergroup evaluation.

Relative group status and stigma. An idea frequently attributed to SJT (Jost & Banaji, 1994) is that disadvantaged groups are more likely to display outgroup favoritism to the extent that their group is stigmatized by society (e.g., Dasgupta, 2004; Lane et al., 2005; Livingston, 2002; Rudman et al., 2002). This notion is also inherent in Jost and colleagues' (2004) proposal that those who "suffer the most from the system are also those who have the most to explain, justify, and rationalize" (p. 909). Based on these ideas, intergroup evaluation in disadvantaged groups should depend on a social group's

relative status and/or level of stigma.

Evidence for a relationship between stigma and intergroup evaluation in disadvantaged groups is mixed. For example, in one comparative study disadvantaged groups with higher perceived group status (Jewish and Asian participants) displayed ingroup favoritism on implicit measures and those with lower perceived group status (overweight and poor participants) displayed outgroup favoritism on implicit measures, but both groups' intergroup evaluations on explicit measures were unrelated to group status (Rudman et al., 2002). Similarly, Black participants who believed that Black Americans were disliked by White Americans demonstrated less implicit ingroup favoritism, but more explicit ingroup favoritism, relative to Black participants who believed that Black Americans were liked by White Americans (Livingston, 2002). In other words, to the extent that Black people believed that their social group was disliked by White people, they demonstrated ingroup favoritism on an explicit measure but outgroup favoritism on an implicit measure. In sum, extant findings support the relationship between intergroup evaluations and group status/stigma predicted by SJT (Jost et al., 2004), but only on implicit measures.

Open Questions Regarding Intergroup Evaluation in Disadvantaged Groups

As the reviewed findings illustrate, extant empirical evidence supports some of the predictions derived from SJT (Jost et al., 2004) regarding outgroup favoritism in disadvantaged groups, but evidence for other predictions is mixed. Additionally, the accumulated evidence highlights two, related gaps in the SJT literature. First, many of the studies reviewed thus far focused on one disadvantaged group at a time (e.g., Ashburn-Nardo et al., 2003; Hoffarth & Jost, 2017; Livingston, 2002; Uhlmann et al., 2002), so it remains an open question whether or to what extent these results are specific to the investigated social groups. This narrow focus provides a streamlined experimental design to test specific predictions derived from SJT, but offers only limited insight into general

processes among disadvantaged groups. Given that SJT (Jost et al., 2004) is articulated as a generalized theory, analyses that incorporate multiple social identities are better positioned to test the claims made by SJT, as well as examine their boundary conditions. The second, related gap in the SJT literature is that a number of studies have examined intergroup attitudes across multiple social groups, but generally focus on documenting main effects of ingroup versus outgroup favoritism (e.g., Axt, Ebersole, & Nosek, 2014; Devos & Banaji, 2005; Nosek et al., 2002; Rudman & Ashmore, 2007), to the exclusion of psychological processes directly related to SJT, such as system justification tendencies, ideological beliefs, or stigma. To our knowledge, only two studies examining intergroup evaluation in disadvantaged groups fulfill both of these criteria and measure psychological processes related to SJT across multiple social groups (Jost et al., 2004; Rudman et al., 2002). Thus, SJT as a generalized theory of intergroup processes among disadvantaged groups would be strengthened by more process-level evidence from more groups.

Taken together, SJT (Jost et al., 2004) has articulated or inspired a number of hypotheses regarding intergroup evaluation in disadvantaged groups. Namely, SJT predicts that outgroup favoritism is more likely: to manifest on implicit than on explicit measures; to the extent that disadvantaged groups exhibit system justifying tendencies or endorse conservative beliefs; and to the extent that a social group is stigmatized by society. However, extant evidence provides varying levels of support for these predictions. These gaps in the SJT literature have motivated the present research into the relationships among system justifying tendencies, conservative beliefs, and stigma in the context of the implicit and explicit intergroup evaluations of large samples of many disadvantaged groups.

The Present Research

The primary aim of our research is to directly test hypotheses derived from SJT (Jost et al., 2004) regarding intergroup evaluations in disadvantaged groups. To do so, the present research uses very large samples drawn from 14 countries reflecting eight distinct

social identities, which is a broader and more diverse sample than has been examined in any previous SJT research. These data were collected by Project Implicit (https://implicit.harvard.edu/implicit/), a demonstration website where visitors can take different online versions of the IAT. Furthermore, Project Implicit has established a wide variety of international collaborations, setting up websites in many countries that additionally conduct country-specific studies with translated and adapted measures. One major advantage of these data is that the methodology is highly similar across countries, providing high levels of internal validity for between-country comparisons. Project Implicit data were made available by Xu and colleagues (Xu et al., 2017, 2018) at the Open Science Framework (OSF).

The present research focuses on intergroup evaluation in disadvantaged groups, including both implicit and explicit measures as dependent variables. As implicit measures, we use the IAT with participants' ingroup and outgroup as target categories. As explicit measures, we use two: one-item preference measures, which ask participants to judge how much they prefer the ingroup relative to the outgroup; and feeling thermometers, which ask participants to (separately) report their felt warmth or coldness towards the ingroup and the outgroup.

Using the Project Implicit data, we first test the prediction that disadvantaged group members "will be more likely to exhibit outgroup favoritism on implicit measures than on explicit measures" (Jost et al., 2004, p. 893) by examining the magnitude and direction of disadvantaged groups' intergroup evaluation, comparing effect sizes on an implicit measure with two explicit measures. Second, we test the prediction that "[a]s political conservatism increases, members of low-status groups will exhibit increased outgroup favoritism" (Jost et al., 2004, p. 901) by examining whether implicit and explicit intergroup evaluation in disadvantaged groups is moderated by self-reported political ideology. Third, we test the assumption that those who "suffer the most from the system are also those who have the most to explain, justify, and rationalize" (Jost et al., 2004, p. 909) by examining whether

stigma is related to outgroup favoritism among disadvantaged groups. One way to think about stigma is that a social group is stigmatized to the extent that the rest of society views that social group negatively (cf. Link & Phelan, 2001; Pinel, 1999). In our analyses, we thus conceptualize stigma in terms of the attitude measures taken from all Project Implicit visitors who are not members of the disadvantaged (i.e., stigmatized) group, and specifically operationalize stigma as non-disadvantaged group members' average evaluations of the disadvantaged group.

Importantly, the analyses reported here extend previous research in two key ways. First, we take a meta-analytic approach to test the hypotheses derived from SJT across a wide variety of intergroup domains. Previous SJT research has generally reported between-group comparisons of intergroup evaluations measured from one relatively advantaged group and one relatively disadvantaged group (e.g., Jost & Hunyady, 2002). Extending upon these relatively more focused comparisons, the present research compares intergroup evaluations measured from a wide variety of disadvantaged groups and, thus, assesses the generalizability of the predictions of SJT across different intergroup domains.

The second way in which the present research extends upon previous work is that we test SJT's predictions—whenever possible—at two levels of analysis: the individual and the social group. The predictions derived from SJT are formulated at the individual level, as applying to members of disadvantaged groups, so we test these predictions at the individual level: e.g., examining whether members of disadvantaged groups who endorse conservative beliefs also prefer the outgroup relative to the ingroup. We also test the predictions of SJT at the level of the social group by using group-level aggregates of measures of intergroup evaluation: e.g., examining whether disadvantaged groups who endorse more conservative beliefs also prefer the outgroup relative to the ingroup. In these analyses, the sample (i.e., social group) rather than the participant is treated as the unit of observation. This social-group level approach has three benefits. The first benefit is that it provides the opportunity to test the moderating influence of social group *per se* on

intergroup evaluation. The second benefit to this approach is that, relative to analyses based on individual-level measures, group-level aggregates yield more precise estimates of intergroup evaluation (Rushton, Brainerd, & Pressley, 1983). The issue of measurement precision is especially relevant in the context of the IAT, which has been criticized for having low measurement reliability relative to explicit measures of the same construct (Gawronski, Morrison, Phills, & Galdi, 2017). The third benefit of using the social group as the unit of observation is that it can reveal qualitatively different psychological processes than can analyses that use the individual as the unit of observation.

Rushton et al. (1983) highlight the utility of aggregation from the perspective of amplifying signal (i.e., the construct of interest) by canceling out noise (i.e., measurement error). However, aggregation also cancels out the influence of other psychological constructs that might vary between individuals that are not measurement error, but also are not specific to group identity. Consequently, individual-level analyses can be interpreted to reflect the influence of a variety of processes. In contrast, group-level analyses cancel out the influence of processes that vary among group members, thereby amplifying the influence of the common trait(s) shared among group members, i.e., the defining feature(s) of group membership. Taken together, individual-level analyses can be interpreted to reflect individual differences, and group-level analyses to reflect group processes. Because SJT is articulated at the level of the individual, these group-level analyses represent a novel extension of the theoretical perspective. By including both individual- and group-level analyses, the present research examines the extent to which the hypotheses proposed by SJT persist at both the individual level and the social group level.

To assume without empirical support that phenomena at one level of analysis persist at other levels is to commit the ecological fallacy (Selvin, 1958). In the classic demonstration of the ecological fallacy, English literacy rates were higher in regions of America with higher proportions of foreign-born (i.e., non-native English speaking) residents (Robinson, 1950). Follow-up analyses of this seemingly paradoxical finding

revealed that foreign-born individuals were less likely than native Americans to be English-literate, but largely settled in regions where the population is more literate (e.g., where there are more employment and educational opportunities). From this perspective, it would be premature to assume that any of the predictions of SJT that are based on individual-level data necessarily persist at the social group level. Thus, the present research is both statistically and theoretically positioned to extend SJT and the intergroup relations literature more broadly.

Method

Study Selection

Figure 1 depicts an adapted PRISMA flow diagram (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009), visualizing the process of identifying datasets and assessing their eligibility. At the time of data analysis and compilation of this report on April 5, 2019, $N = 110$ datasets were publicly available at OSF. Datasets can be accessed via osf.io/kaqi5 and osf.io/y9hiq[4]

[4] The two OSF projects differ in that one hosts only datasets collected on the US-based website whereas the other project hosts datasets collected on country-specific, non-US sites.

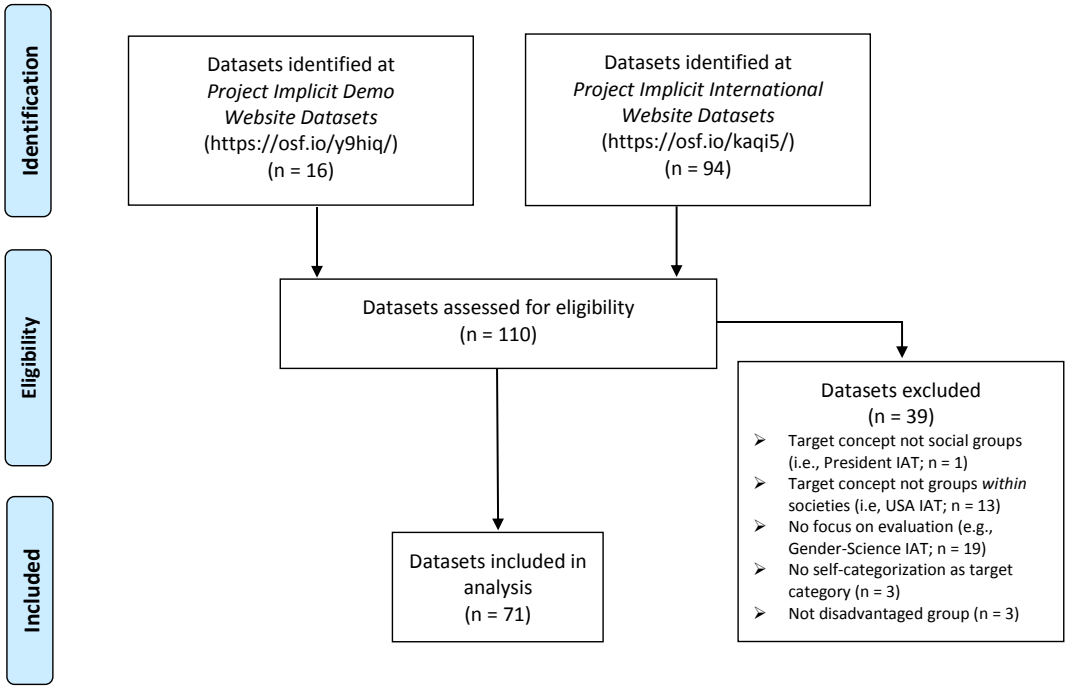

Figure 1. Adapted PRISMA flow diagram (Moher et al., 2009) showing the process of identifying, assessing for eligibility, and selecting datasets. In contrast to the standard PRISMA 2009 flow diagram, the screening phase is omitted here because all datasets were drawn from the same source: Project Implicit.

Inclusion and exclusion criteria. We included only studies with social groups or categories as target concepts. Studies measuring evaluations of individuals, such as evaluations of the President of the US, were outside the scope of the present study and were thus excluded ($n = 1$ dataset). Because we were interested in how disadvantaged group within societies evaluate their ingroup relative to an outgroup, we excluded studies focusing on nations as social categories (i.e., USA IATs; $n = 13$ datasets). Furthermore, the present research focused on intergroup *evaluations*, that is, relationships between positive versus negative attributes and the ingroup versus the outgroup. Consequently, we included only evaluative IATs and items and scales assessing liking and preference. Studies that did not focus on evaluations *per se* (e.g., semantic attributes, or stereotypes, such as American vs. Foreign in the Asian IAT, or Science vs. Arts in the Gender-Science IATs; $n = 19$ datasets) were outside the scope of this study and were thus excluded. From each dataset we included only data from participants who self-identified as members of the disadvantaged social group that the study focused on (e.g., a person identifying as being older who took the Age IAT). Data of all other participants, who did not self-identify as members of the disadvantaged target category, were excluded from the primary analyses (but were used to compute average stigma scores, see below). Studies were also excluded if one of the target categories did not unambiguously refer to a disadvantaged group ($n = 3$).[5] Additionally, studies were excluded if participants were not directly asked whether they self-identified as members of either of the target categories ($n = 3$) .[6] The final number of datasets included in the present study was $n = 71$, yielding a total of $k = 73$

[5] We excluded two studies, conducted in the US, using the target categories "Arab Muslims" versus "Other People" and "Judaism" versus "Other Religions", because Arab Muslim targets or Jewish targets were not unambiguously disadvantaged relative to *all* other outgroup targets. Furthermore, we excluded data from Jewish participants, who completed a Religion IAT with the target categories "Judaism" versus "Islam", because Judaism was not unambiguously disadvantaged relative to Islam in the US.

[6] We excluded one study that focused on evaluations of dark-skinned versus light-skinned people in Australia, because national/ethnic group membership but not self-reported skin tone was measured. Furthermore, we excluded a study that focused on evaluations of Black people versus White people in The Netherlands, because it assessed ethnic group memberships (e.g., Surinamese) but not self-categorization as Black. Lastly, we excluded one dataset, which used multi-category versions of the IAT to measure evaluations of religious groups, but did not measure self-reported religious group membership.

independent effect sizes.

Datasets reviewed. Over the years, different studies have been added to the Project Implicit demonstration websites. The time period of data collection spans from 2002 to the present, with some studies running for the full time period (e.g., Race IAT in the US) and others starting later (e.g., the Race IAT in Spain) or others discontinued at some point in time (e.g., the Religion IAT in the US). For each country-specific study, we used the dataset that included data for all years of data collection. The present research includes eight intergroup domains with distinct IAT versions and datasets from 14 countries[7]; see Table 1 for an overview of included samples and dependent measures and Table 2 for descriptive statistics of continuous moderators. The total sample size was $N =$ 715,721.

[7] For data from the US Project Implicit website, we included only participants who indicated US citizenship or who indicated that their current location was in the United States.

Table 1
Overview of Project Implicit samples, dependent variables, and descriptive statistics.

			IAT		Preference		Thermometer	
IAT Version	Group	Country	dz	N	dz	N	dz	N
Age	Older (55+) age	AUS	1.08	1033	0.06	1158	-0.05	1236
Age	Older (55+) age	BEL	1.52	165	0.46	160	0.27	185
Age	Older (55+) age	BRA	1.12	138	0.31	146	0.22	162
Age	Older (55+) age	CAN (EN)	1.07	1387	0.07	1474	-0.06	1536
Age	Older (55+) age	CAN (FR)	1.50	97	0.35	101	0.05	105
Age	Older (55+) age	CHN	1.68	18	0.50	21	0.30	21
Age	Older (55+) age	DEU	1.57	687	0.40	661	0.31	727
Age	Older (55+) age	ESP	1.42	658	0.49	651	-0.02	684
Age	Older (55+) age	FRA	1.67	836	0.44	836	0.25	905
Age	Older (55+) age	GBR	1.17	2485	0.14	2737	0.00	2861
Age	Older (55+) age	KOR	1.25	24	0.67	26	-0.04	28
Age	Older (55+) age	NLD	1.30	652	0.43	704	0.26	767
Age	Older (55+) age	SWE	1.32	773	0.22	663	-0.02	798
Age	Older (55+) age	USA	1.06	35208	0.04	35960	-0.14	37776
Arab	Arab and/or Muslim	FRA	0.33	1455	-0.24	1465	-0.30	125
Disability	Participants with disabilities	USA	0.89	48086	0.23	50589	-0.11	51832

Table 1 continued

			IAT		Preference		Thermometer	
IAT Version	Group	Country	dz	N	dz	N	dz	N
Race	Black or African (American)	AUS	-0.01	325	-0.61	382	-0.55	382
Race	Black or African (American)	BEL	0.08	86	-0.19	96	-0.50	102
Race	Black or African (American)	BRA	0.16	480	-0.22	677	-0.40	691
Race	Black or African (American)	CAN (EN)	0.12	2293	-0.53	2445	-0.56	2532
Race	Black or African (American)	CAN (FR)	0.16	152	-0.46	143	-0.51	158
Race	Black or African (American)	CHN	-0.14	3	1.17	4	-1.79	6
Race	Black or African (American)	DEU	0.34	194	-0.04	202	-0.11	225
Race	Black or African (American)	ESP	0.56	88	0.17	102	-0.21	110
Race	Black or African (American)	FRA	0.23	1022	-0.35	1078	-0.40	1132
Race	Black or African (American)	GBR	0.03	4197	-0.58	4864	-0.55	4988
Race	Black or African (American)	SWE	-0.08	207	-0.33	217	-0.61	229
Race	Black or African (American)	USA	-0.09	272872	-0.76	265536	-0.74	282975
Religion	Jewish	USA	-0.90	780	-2.06	761	-1.46	763
Religion	Muslim	USA	-0.63	405	-1.77	390	-0.97	392
Religion	Muslim	USA	-0.99	426	-1.85	412	-1.06	416
Sexuality	Gay/Lesbian	AUS	-0.31	1880	-0.47	2366	-0.34	2457
Sexuality	Gay/Lesbian	BEL	-0.31	367	-0.71	375	-0.42	435

Table 1 continued

			IAT		Preference		Thermometer	
IAT Version	Group	Country	dz	N	dz	N	dz	N
Sexuality	Gay/Lesbian	BRA	-0.61	1264	-0.70	1732	-0.73	1790
Sexuality	Gay/Lesbian	CAN (EN)	-0.31	2101	-0.47	2328	-0.42	2448
Sexuality	Gay/Lesbian	CAN (FR)	-0.39	305	-0.25	306	-0.24	339
Sexuality	Gay/Lesbian	CHN	-0.53	1855	-0.61	2352	-0.30	2454
Sexuality	Gay/Lesbian	DEU	-0.40	2896	-0.54	2954	-0.40	3276
Sexuality	Gay/Lesbian	ESP	-0.38	1204	-0.52	1371	-0.37	1491
Sexuality	Gay/Lesbian	FRA	-0.49	3068	-0.47	3252	-0.42	3592
Sexuality	Gay/Lesbian	GBR	-0.26	5904	-0.46	6963	-0.32	7198
Sexuality	Gay/Lesbian	KOR	-0.31	1375	-1.01	1818	-0.73	1909
Sexuality	Gay/Lesbian	NLD	-0.27	1020	-0.66	1183	-0.46	1272
Sexuality	Gay/Lesbian	RUS	-0.39	300	-0.62	294	-0.58	353
Sexuality	Gay/Lesbian	SWE	-0.52	962	-0.55	987	-0.56	1143
Sexuality	Gay/Lesbian	USA	-0.30	99669	-0.62	104699	-0.53	108612
Skin Tone	Dark-skinned	BEL	0.25	237	0.20	278	-0.09	319
Skin Tone	Dark-skinned	BRA	0.31	1055	0.27	1577	0.04	1647
Skin Tone	Dark-skinned	CAN (EN)	0.38	1332	0.01	1462	-0.17	1544
Skin Tone	Dark-skinned	CAN (FR)	0.38	91	-0.04	92	-0.19	102

Table 1 continued

			IAT		Preference		Thermometer	
IAT Version	Group	Country	dz	N	dz	N	dz	N
Skin Tone	Dark-skinned	CHN	0.98	1105	0.81	1343	0.01	1414
Skin Tone	Dark-skinned	DEU	0.55	957	0.24	1034	-0.12	1127
Skin Tone	Dark-skinned	ESP	0.62	915	0.33	1103	-0.05	1160
Skin Tone	Dark-skinned	FRA	0.49	1180	-0.01	1310	-0.06	1390
Skin Tone	Dark-skinned	GBR	0.34	1747	-0.03	1974	-0.23	2111
Skin Tone	Dark-skinned	KOR	0.82	1212	0.79	1559	-0.46	1615
Skin Tone	Dark-skinned	NLD	0.18	861	0.08	1055	-0.21	1138
Skin Tone	Dark-skinned	SWE	0.20	834	0.05	919	-0.19	996
Skin Tone	Dark-skinned	USA	0.27	58479	-0.21	59063	-0.31	61888
Weight	Overweight	AUS	0.93	4128	0.72	5280	0.33	5401
Weight	Overweight	BEL	0.83	657	0.70	741	0.26	783
Weight	Overweight	BRA	0.59	1605	0.80	2144	0.55	2164
Weight	Overweight	CAN (EN)	0.89	6233	0.72	6979	0.29	7129
Weight	Overweight	CAN (FR)	1.22	605	1.10	651	0.71	670
Weight	Overweight	CHN	0.12	2774	0.79	3450	-0.42	3541
Weight	Overweight	DEU	1.16	5334	0.65	5666	0.35	5964
Weight	Overweight	ESP	0.96	2778	0.90	3213	0.22	3304

Table 1 continued

			IAT		Preference		Thermometer	
IAT Version	Group	Country	dz	N	dz	N	dz	N
Weight	Overweight	FRA	1.14	4031	0.99	4391	0.59	4568
Weight	Overweight	GBR	0.83	9540	0.65	10859	0.30	11116
Weight	Overweight	KOR	0.60	2363	1.37	3168	-0.22	3264
Weight	Overweight	NLD	0.75	1402	0.84	1727	0.41	1815
Weight	Overweight	SWE	1.19	4692	0.89	5123	0.51	5409
Weight	Overweight	USA	0.85	98172	0.66	100396	0.24	103237

Note. Top row of column labels refer to dependent measures: IAT = IAT; Preference = one-item preference measure; Thermometer = feeling thermometer difference score. Second row of column labels refer to IAT Version, Group = disadvantaged group examined in primary analyses, dz = Cohen's d_z, with positive scores indicating outgroup favoritism and negative scores indicating ingroup favoritism from the perspective of the disadvantaged group; N = sample size. Abbreviations in third column denote country codes: AUS = Australia, BEL = Belgium, BRA = Brazil, CAN (EN) = Canada (English), CAN (FR) = Canada (French), CHN = China, DEU = Germany, ESP = Spain, FRA = France, GBR = United Kingdom, KOR = Korea, NLD = The Netherlands, RUS = Russia, SWE = Sweden, USA = United States.

Individual-level Measures

Intergroup evaluation. As implicit measures, we used the evaluative variant of the IAT for each sample. Attribute categories of IATs were always "Good" versus "Bad"; target categories differed according to the specific intergroup domain. We included two different explicit measures. One was a one-item preference measure, which asked participants to judge how much they preferred the ingroup relative to the outgroup, with responses made using either a 5-point scale (in the earlier years of Project Implicit) or a 7-point scale (in more recent years). For the other explicit measure, we used feeling thermometers, which asked participants to respond on scales from 0 (*extremely cold*) to 10 (*extremely warm*) how they felt towards the ingroup and the outgroup.

Conservatism. Conservatism was assessed using one-item self-placement measures. Using 6- or 7-point scales, participants were asked to place themselves along a continuum ranging from liberal to conservative.[8]

Group-level Measures

Intergroup evaluation. We calculated the sample averages of the IAT *D* Score, the one-item preference score, and the feeling thermometer difference score, separately for each sample of disadvantaged groups.

Conservatism. We calculated the sample average of the one-item conservatism measure separately for each sample of disadvantaged groups.

Stigma. We calculated average evaluation scores of each disadvantaged group based on the responses of Project Implicit visitors in each study who self-reported being members of social groups other than the disadvantaged group in the same study. We used these evaluation scores as proxies for the extent to which each disadvantaged group was

[8] Measures were adapted for country-specific websites with some studies using other but similar labels (e.g., left-wing vs. right-wing, conservative vs. progressive).

stigmatized by the rest of society. For each disadvantaged group, we thus calculated three measures of stigma, based on non-disadvantaged group members' IAT *D* Scores, one-item preference scores, and feeling thermometer difference scores.

Table 2
Descriptive statistics for continuous moderators.

				Stigma		
IAT Version	Group	Country	Conservatism	IAT	Preference	Thermometer
Age	Older (55+) age	AUS	-0.25	1.14	0.47	0.17
Age	Older (55+) age	BEL	-0.15	1.39	0.60	0.38
Age	Older (55+) age	BRA	-0.08	0.99	0.51	0.24
Age	Older (55+) age	CAN (EN)	-0.43	1.04	0.41	0.11
Age	Older (55+) age	CAN (FR)	-0.66	1.24	0.53	0.29
Age	Older (55+) age	CHN	-0.82	1.20	0.70	-0.05
Age	Older (55+) age	DEU	-0.62	1.38	0.57	0.28
Age	Older (55+) age	ESP	-0.56	1.20	0.52	0.07
Age	Older (55+) age	FRA	-0.39	1.39	0.61	0.31
Age	Older (55+) age	GBR	-0.05	1.14	0.38	0.12
Age	Older (55+) age	KOR	0.52	1.30	1.15	-0.27
Age	Older (55+) age	NLD	-0.88	1.28	0.61	0.26
Age	Older (55+) age	SWE	-0.01	1.19	0.46	0.08
Age	Older (55+) age	USA	-0.33	1.09	0.42	0.08
Arab	Arab and/or Muslim	FRA	-0.66	1.08	-0.24	0.53
Disability	Participants with disabilities	USA	-0.37	1.16	0.54	0.18

Table 2 continued

				Stigma		
IAT Version	Group	Country	Conservatism	IAT	Preference	Thermometer
Race	Black or African (American)	AUS	-0.44	0.86	0.49	0.25
Race	Black or African (American)	BEL	-0.78	0.82	0.60	0.41
Race	Black or African (American)	BRA	-0.80	0.86	0.45	0.23
Race	Black or African (American)	CAN (EN)	-0.42	0.89	0.42	0.26
Race	Black or African (American)	CAN (FR)	-0.32	1.08	0.57	0.40
Race	Black or African (American)	CHN	0.88	1.07	0.92	0.27
Race	Black or African (American)	DEU	-0.27	0.88	0.50	0.27
Race	Black or African (American)	ESP	-0.44	1.02	0.66	0.35
Race	Black or African (American)	FRA	-0.43	0.95	0.46	0.29
Race	Black or African (American)	GBR	-0.40	0.87	0.53	0.36
Race	Black or African (American)	SWE	-0.59	0.71	0.54	0.31
Race	Black or African (American)	USA	-0.34	0.84	0.43	0.29
Religion	Jewish	USA	-0.97	0.80	0.57	0.19
Religion	Muslim	USA	-0.53	0.99	0.82	0.42
Religion	Muslim	USA	-0.64	0.50	0.63	0.50
Sexuality	Gay/Lesbian	AUS	-0.83	0.62	0.50	0.46
Sexuality	Gay/Lesbian	BEL	-0.38	0.66	0.72	0.69

Table 2 continued

				Stigma		
IAT Version	Group	Country	Conservatism	IAT	Preference	Thermometer
Sexuality	Gay/Lesbian	BRA	-0.72	0.47	0.66	0.61
Sexuality	Gay/Lesbian	CAN (EN)	-1.01	0.58	0.48	0.45
Sexuality	Gay/Lesbian	CAN (FR)	-1.00	0.40	0.52	0.54
Sexuality	Gay/Lesbian	CHN	-0.72	0.27	0.88	0.70
Sexuality	Gay/Lesbian	DEU	-0.81	0.46	0.65	0.64
Sexuality	Gay/Lesbian	ESP	-1.11	0.59	0.61	0.51
Sexuality	Gay/Lesbian	FRA	-0.68	0.40	0.54	0.51
Sexuality	Gay/Lesbian	GBR	-0.84	0.65	0.53	0.51
Sexuality	Gay/Lesbian	KOR	-0.55	0.55	0.64	0.25
Sexuality	Gay/Lesbian	NLD	-1.19	0.76	0.65	0.67
Sexuality	Gay/Lesbian	RUS	-0.25	0.85	1.03	0.79
Sexuality	Gay/Lesbian	SWE	-0.40	0.36	0.49	0.48
Sexuality	Gay/Lesbian	USA	-1.21	0.62	0.48	0.47
Skin Tone	Dark-skinned	BEL	-0.47	0.76	0.54	0.31
Skin Tone	Dark-skinned	BRA	-0.52	0.88	0.58	0.38
Skin Tone	Dark-skinned	CAN (EN)	-0.47	0.85	0.43	0.26
Skin Tone	Dark-skinned	CAN (FR)	-0.51	1.09	0.46	0.37

Table 2 continued

				Stigma		
IAT Version	Group	Country	Conservatism	IAT	Preference	Thermometer
Skin Tone	Dark-skinned	CHN	-0.40	1.12	0.91	0.16
Skin Tone	Dark-skinned	DEU	-0.59	0.89	0.51	0.24
Skin Tone	Dark-skinned	ESP	-0.53	1.12	0.91	0.16
Skin Tone	Dark-skinned	FRA	-0.54	0.96	0.38	0.28
Skin Tone	Dark-skinned	GBR	-0.53	0.85	0.42	0.26
Skin Tone	Dark-skinned	KOR	-0.40	0.96	0.90	-0.39
Skin Tone	Dark-skinned	NLD	-0.67	0.68	0.57	0.30
Skin Tone	Dark-skinned	SWE	-0.33	0.67	0.54	0.35
Skin Tone	Dark-skinned	USA	-0.45	0.75	0.31	0.22
Weight	Overweight	AUS	-0.34	0.96	1.07	0.58
Weight	Overweight	BEL	-0.24	1.03	1.12	0.56
Weight	Overweight	BRA	-0.26	0.84	1.08	0.77
Weight	Overweight	CAN (EN)	-0.46	0.90	0.99	0.56
Weight	Overweight	CAN (FR)	-0.54	1.40	1.26	0.88
Weight	Overweight	CHN	-0.43	0.21	0.77	-0.01
Weight	Overweight	DEU	-0.56	1.28	0.98	0.65
Weight	Overweight	ESP	-0.59	1.05	1.05	0.38

Table 2 continued

				Stigma		
IAT Version	Group	Country	Conservatism	IAT	Preference	Thermometer
Weight	Overweight	FRA	-0.42	1.21	1.18	0.75
Weight	Overweight	GBR	-0.42	0.83	0.99	0.56
Weight	Overweight	KOR	-0.35	0.75	1.48	-0.04
Weight	Overweight	NLD	-0.66	0.96	1.24	0.58
Weight	Overweight	SWE	-0.09	1.24	1.09	0.67
Weight	Overweight	USA	-0.29	1.09	1.08	0.48

Table 2 continued

				Stigma		
IAT Version	Group	Country	Conservatism	IAT	Preference	Thermometer

Note. Group = disadvantaged group examined in primary analyses. Conservatism = Cohen's d_z for aggregated mean of disadvantaged group members' self-reported conservatism, with positive scores indicating conservative self-placement and negative scores indicating liberal self-placement. Stigma = Cohen's d_z for the three stigma measures: IAT = IAT *D* Score stigma; Preference = one-item preference score stigma; Thermometer = feeling thermometer difference score stigma. Stigma measures reflect the aggregated mean of non-disadvantaged group members' responses, with positive scores indicating negative evaluations of the disadvantaged group and negative scores indicating positive evaluations of the disadvantaged group. Abbreviations in third column denote country codes: AUS = Australia, BEL = Belgium, BRA = Brazil, CAN (EN) = Canada (English), CAN (FR) = Canada (French), CHN = China, DEU = Germany, ESP = Spain, FRA = France, GBR = United Kingdom, KOR = Korea, NLD = The Netherlands, RUS = Russia, SWE = Sweden, USA = United States.

Intergroup Domains

Old vs. Young. There were $k = 14$ studies focusing on age-related group evaluations. We included in the primary analyses only participants with a self-reported age of 55 years and older. The threshold of 55 for "older adults" has been used in previous research (e.g., Kite, Stockdale, Whitley, & Johnson, 2005; Neugarten, 1974). Target categories in the Age IAT, one-item preference measures, and feeling thermometers were "Old People" versus "Young People".

Arab vs. French. There was $k = 1$ study focusing on evaluations of Arab people relative to French people. We included in the primary analyses only participants who self-categorized as Arab or Muslim. Target categories in the French Arab IAT were "Maghreb People" versus "French People."[9]

Disabled vs. Abled. There was $k = 1$ study focusing on evaluations of disability relative to non-disability. We included in the primary analyses only participants who indicated that they had a disability. Target categories in the Disability IAT, one-item preference measures, and feeling thermometers were "Disabled People" versus "Abled People".

Black vs. White. There were $k = 12$ studies focusing on evaluations of Black people relative to White people. We included in the primary analyses only participants who self-categorized as Black or African, or joint identities, such as African American, Black British, or Black Caribbean. Target categories in the Race IAT, one-item preference measures, and feeling thermometers were "Black People" versus "White People", "Black" versus "White", or "African American" versus "European American", depending on the language- or country-specific demonstration website.

---

[9] Maghreb refers to a region in North and Northwestern Africa. The term is frequently used in French referring to North-African countries with Arabic as an official language, such as Morocco, Algeria, and Tunisia (Oxford Dictionaries, 2019).

Religious groups. There were $k = 3$ studies, all conducted in the USA, focusing on evaluations of religious groups relative to other religious groups. The studies used three variants of the Religion IAT and respective one-item preference measures and feeling thermometers. We included the sample of Jewish participants who completed the "Judaism" versus "Christianity" measures, and the two samples of Muslim participants who completed either the "Judaism" versus "Islam" or the "Islam" versus "Christianity" measures, respectively.

Gay vs. Straight. There were $k = 15$ studies focusing on evaluations of Gay people relative to Straight people. We included in the primary analyses only participants who self-categorized as homosexual. Target categories in the Sexuality IAT, one-item preference measures, and feeling thermometers were "Gay People" versus "Straight People" or "Homosexual" versus "Heterosexual", depending on the language- or country-specific demonstration website.

Dark-Skinned vs. Light-Skinned. There were $k = 13$ studies focusing on evaluations of dark-skinned people relative to light-skinned people. We included in the primary analyses only participants who self-categorized as somewhat dark-skinned, dark-skinned, or very dark-skinned. Target categories in the Skin Tone IAT, one-item preference measures, and feeling thermometers were "Dark Skinned People" versus "Light Skinned People".

Overweight vs. Normal Weight. There were $k = 14$ studies focusing on evaluations of overweight people relative to normal weight people. We included in the primary analyses only participants who self-categorized as being overweight. Target categories in the Weight IAT, one-item preference measures, and feeling thermometers were "Fat People" versus "Thin People" or "Fat" versus "Thin", depending on the language- or country-specific demonstration website.

Data analysis

We calculated average IAT *D* Scores (Greenwald, Nosek, & Banaji, 2003), one-item preference scores, and feeling thermometer difference scores for each sample within each dataset. Effect size estimates for all measures were coded such that positive scores indicated a preference for advantaged groups relative to disadvantaged groups and negative scores indicated a preference for disadvantaged groups relative to advantaged groups. We calculated feeling thermometer difference scores by subtracting ingroup feeling thermometers from outgroup feeling thermometers. We calculated effect size estimates Cohen's d_z for IAT effects by dividing each sample IAT *D* Score by its' standard deviation. For one-item preference measures and feeling thermometer difference scores, we took two steps to calculate effect size estimates. First, we performed one-sample *t*-tests, testing one-item preference scores against the scale midpoint and feeling thermometer difference scores against zero. We then calculated for each (sub-)sample Cohen's d_z, using the following formula (see Lakens, 2013; Rosenthal, 1991):

$$d_z = \frac{t}{\sqrt{n}}$$

where t is the test statistic obtained from one-item preference scores (versus the scale midpoint) or feeling thermometer difference scores (versus zero) and n is the respective sample size.

One-item preference measures were assessed using 5-point scales in earlier years and using 7-point scales in more recent years. Each sample could thus contribute up to two one-item preference scores. For samples that produced two effect sizes, we calculated an aggregated mean effect size, weighting effect sizes by their respective sample sizes. All dependent variables—IAT *D* Scores, one-item preference scores, and feeling thermometer difference scores—were coded such that positive scores indicated outgroup favoritism and negative scores indicated ingroup favoritism from the perspective of members of the

disadvantaged groups.[10]

We further calculated effect size estimates for conservatism using the same procedures detailed above. For samples that produced two effect sizes (i.e., corresponding to 6- and 7-point scales), we calculated an aggregated mean effect size, weighting effect sizes by their respective sample sizes. Effect size estimates for conservatism were coded such that positive scores indicated more conservative self-placement and negative scores indicated more liberal self-placement. Finally, we used the same approach to calculate effect size estimates for the three measures of stigma: IAT D Scores, one-item preference scores, and feeling thermometer difference scores, using the data of all participants of each study who did not self-categorize as belonging to the disadvantaged target group. Effect size estimates for stigma were coded such that positive scores indicated a preference for advantaged groups relative to disadvantaged groups and negative scores indicated a preference for disadvantaged groups relative to advantaged groups.

Analyses were conducted using a meta-analytic framework. We employed a random-effects model to allow for the assumption that different studies have different underlying true effects without assuming that there is only one true effect underlying the observed study results (e.g., Borenstein, Hedges, Higgins, & Rothstein, 2010; Cheung, 2015). As such, the studies included in our meta-analyses are assumed to be a random sample from a population of studies which, in principle, allows for the meta-analytic results to be generalized beyond the included studies (Cheung, 2015). For fitting a random-effects model, we weighted effects by their inverse variance to estimate an average population effect size. We calculated the inverse variance w of d_z for IAT D Scores, one-item preference scores, and feeling thermometer difference scores following the formula provided

[10] Note that the number of datasets does not equal the number of effect sizes within the present study. Instead, effect size estimates were calculated at the sample level and some datasets contributed multiple independent effect sizes from multiple independent samples. These were cases in which a dataset was comprised of studies that assessed evaluations toward different target categories. For example, the Religion IAT dataset included different independent studies, assessing evaluations towards Christianity vs. Judaism, Christianity vs. Islam, and Judaism vs. Islam. This dataset contributed three independent samples, because Jewish and Muslim participants participated in all three studies.

by Lipsey (2001, p. 72):

$$w = \frac{1}{SE^2}$$

where

$$SE^2 = \sqrt{\frac{n+n}{n*n} + \frac{d_z}{2(n+n)}}$$

.

First, we estimated the heterogeneity of effects within each measure type, and then conducted follow-up moderator analyses with a series of mixed-effects meta-regressions. In these meta-regressions, political ideology and stigma were included as continuous moderators, and social group was dummy-coded as a categorical moderator. All analyses were done using R.[11] Analyses scripts are accessible at https://osf.io/cxp9z/.[12]

Results

Testing the Predictions of SJT at the Individual Level

The predictions derived from SJT (Jost et al., 2004) are formulated at the individual level, as applying to members of disadvantaged groups. Consequently, we first report a set of analyses based on individual-level data.

Implicit versus explicit measures. We fitted three separate random-effects models, using the three measures of intergroup evaluation as dependent variables.

IAT* D *Scores.

[11] R (Version 3.6.2; R Core Team, 2017) and the R-packages *apaTables* (Version 2.0.5; Stanley, 2018), *bookdown* (Version 0.17; Xie, 2016), *cowplot* (Version 1.0.0; Wilke, 2017), *data.table* (Version 1.12.8; Dowle & Srinivasan, 2017), *here* (Version 0.1; Müller, 2017), *knitr* (Version 1.28; Xie, 2015), *metafor* (Version 2.1.0; Viechtbauer, 2010), *papaja* (Version 0.1.0.9,942; Aust & Barth, 2018), *png* (Version 0.1.7; Urbanek, 2013), *tidyverse* (Version 1.3.0; Wickham, 2017), and *xtable* (Version 1.8.4; Dahl, 2016)

[12] Additional analyses can be found in the Supplement, including: individual-level correlational analyses using other measures of system justifying beliefs among a sub-sample of studies; country-level analyses using cultural value dimensions; parallel analyses assessing the relationship between ideology and intergroup evaluations in advantaged groups and group members; and additional figures.

We observed a significant mean effect of $d_z = 0.43$, $z = 5.34$, $p < .001$, 95% *CI* [0.27; 0.59] on IAT *D* Scores. This medium-sized positive effect indicates that, on average, members of disadvantaged groups displayed outgroup favoritism on the IAT. The estimated amount of total heterogeneity was $\tau^2 = 0.46$, $Q(72) = 80{,}468.60$, $p < .001$, accounting for a large proportion of the total variability. The percentage of the heterogeneity not attributable to sampling error was $I^2 = 99.94\%$, indicating that a high percentage of the estimated heterogeneity was due to genuine between-sample variability (Higgins, Thompson, Deeks, & Altman, 2003).

One-item preference scores.

We observed a mean effect of $d_z = 0.02$, $z = 0.22$, $p = .827$, 95% *CI* [-0.14; 0.18] on one-item preference scores. This null effect indicates that, on average, members of disadvantaged groups displayed neither ingroup nor outgroup favoritism on the one-item preference measures. The estimated amount of total heterogeneity was $\tau^2 = 0.47$, $Q(72) = 186{,}325.30$, $p < .001$, accounting for a large proportion of the total variability, and $I^2 = 99.95\%$, indicating high total heterogeneity due to genuine between-sample variability.

Feeling thermometer difference scores.

We observed a mean effect of $d_z = -0.20$, $z = -3.83$, $p < .001$, 95% *CI* [-0.30; -0.10] on feeling thermometer difference scores. This small negative effect indicates that, on average, members of disadvantaged groups displayed ingroup favoritism on feeling thermometer difference scores. The estimated amount of total heterogeneity was $\tau^2 = 0.18$, $Q(72) = 45{,}597.29$, $p < .001$, accounting for a substantial proportion of the total variability, and $I^2 = 99.83\%$, indicating high total heterogeneity due to genuine between-sample variability.

Conservatism. To test the relationship between disadvantaged group members' conservative beliefs and intergroup evaluation at the individual level, we calculated the correlation between conservatism and the three measures of intergroup evaluation within each sample and fitted three separate random-effects models, weighting each correlation

coefficient by its corresponding sample size.

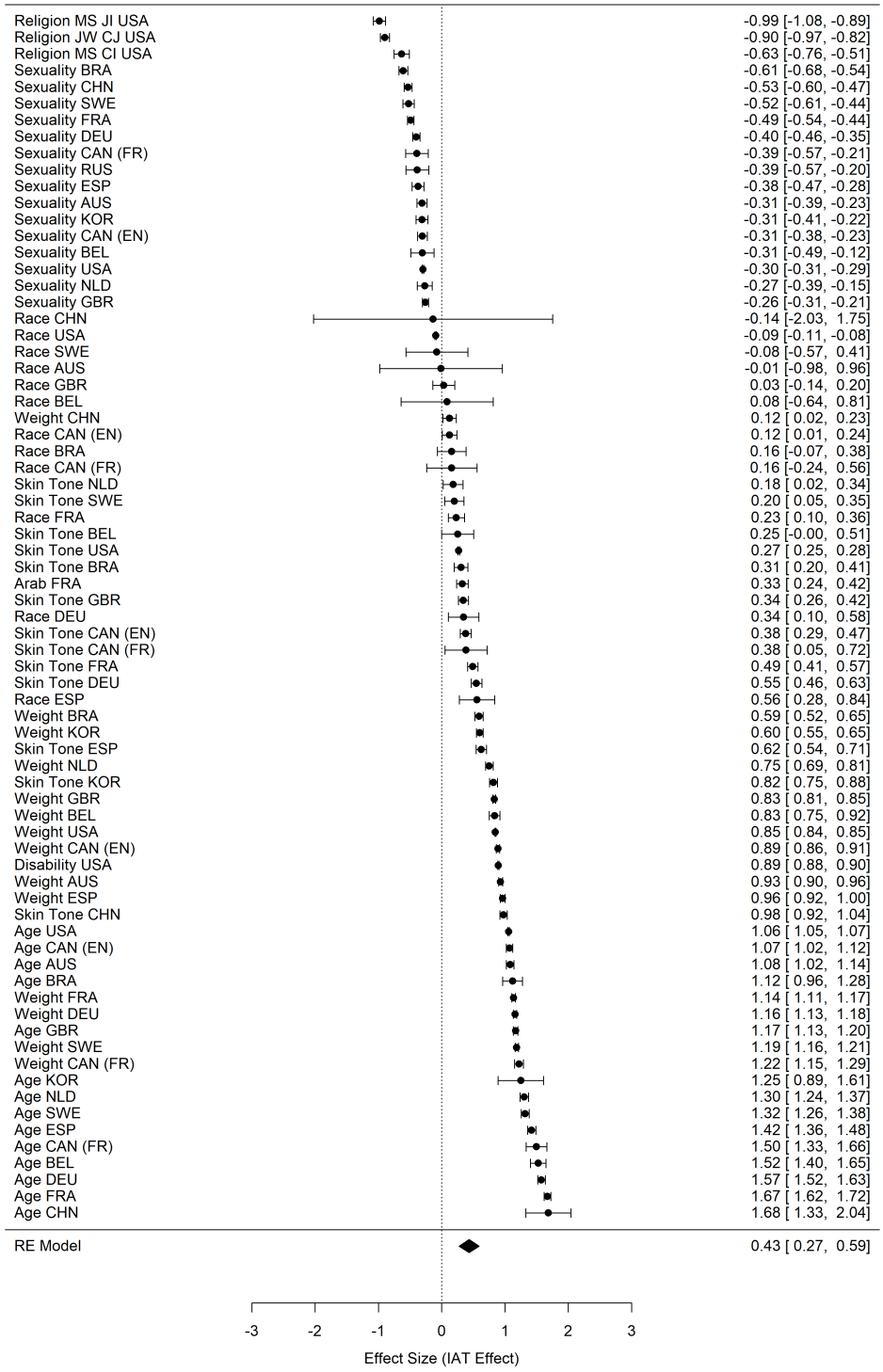

Figure 2. Caterpillar plot of random-effects meta-analysis of IAT effects (IAT *D* Scores) with study effects ordered by effect size. Positive scores indicate outgroup favoritism and negative scores indicate ingroup favoritism from the perspective of the disadvantaged groups. Error bars depict 95% confidence intervals and values in squared brackets indicate lower and upper bounds of confidence intervals. Dataset labels denote the intergroup domain for each study and the respective country. Abbreviations for religious groups: MS = Muslim participants, JW = Jewish participants, JI = Judaism vs. Islam, CJ = Christianity vs. Judaism, CI = Christianity vs. Islam. Country codes: AUS = Australia, BEL = Belgium, BRA = Brazil, CAN (EN) = Canada (English), CAN (FR) = Canada (French), CHN = China, DEU = Germany, ESP = Spain, FRA = France, GBR = United Kingdom, KOR = Korea, NLD = The Netherlands, RUS = Russia, SWE = Sweden, USA = United States.

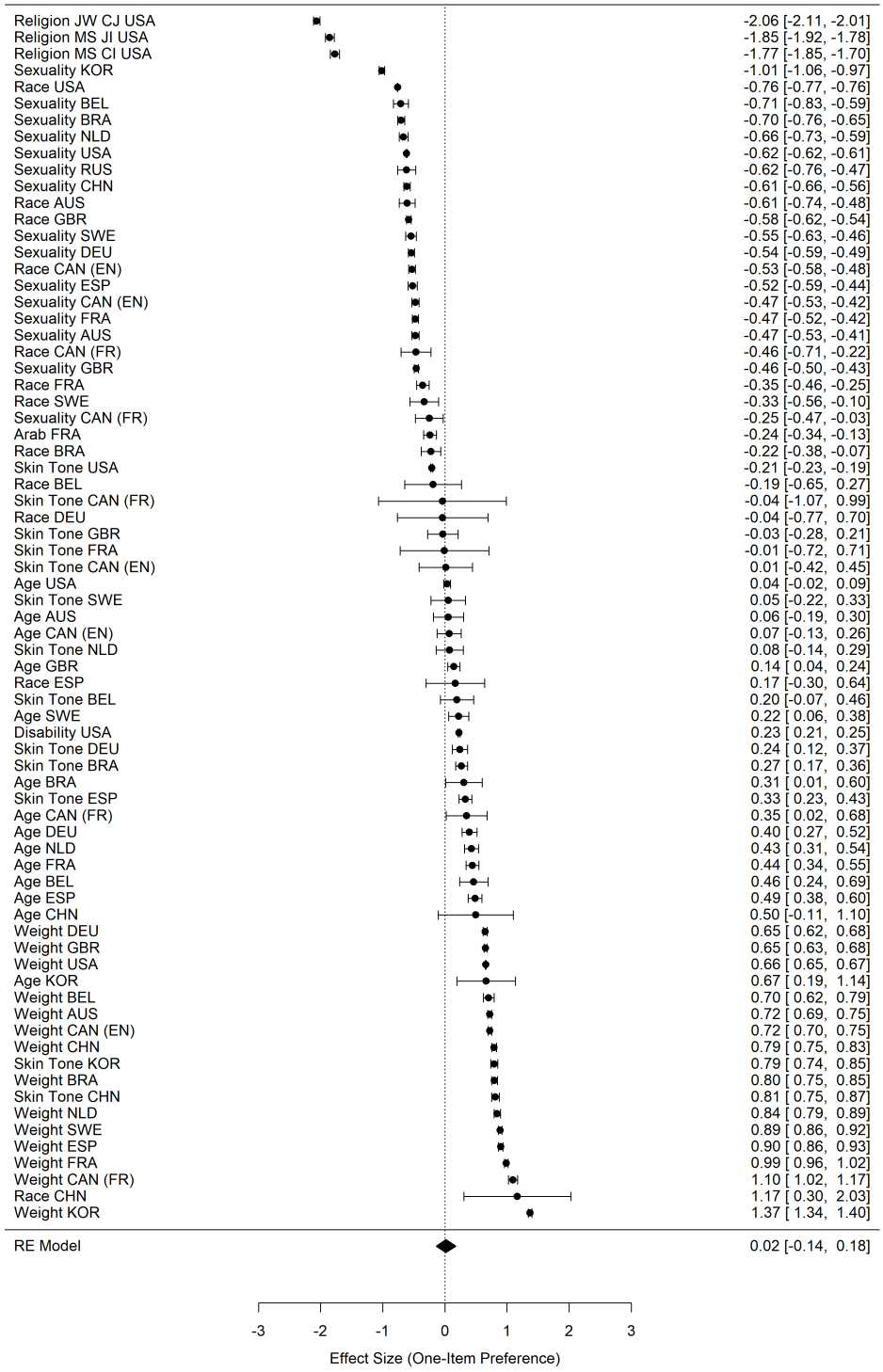

Figure 3. Caterpillar plot of random-effects meta-analysis of one-item preference scores with study effects ordered by effect size. Positive scores indicate outgroup favoritism and negative scores indicate ingroup favoritism from the perspective of the disadvantaged groups. Error bars depict 95% confidence intervals and values in squared brackets indicate lower and upper bounds of confidence intervals. Dataset labels denote the intergroup domain for each study and the respective country. Abbreviations for religious groups: MS = Muslim participants, JW = Jewish participants, JI = Judaism vs. Islam, CJ = Christianity vs. Judaism, CI = Christianity vs. Islam. Country codes: AUS = Australia, BEL = Belgium, BRA = Brazil, CAN (EN) = Canada (English), CAN (FR) = Canada (French), CHN = China, DEU = Germany, ESP = Spain, FRA = France, GBR = United Kingdom, KOR = Korea, NLD = The Netherlands, RUS = Russia, SWE = Sweden, USA = United States.

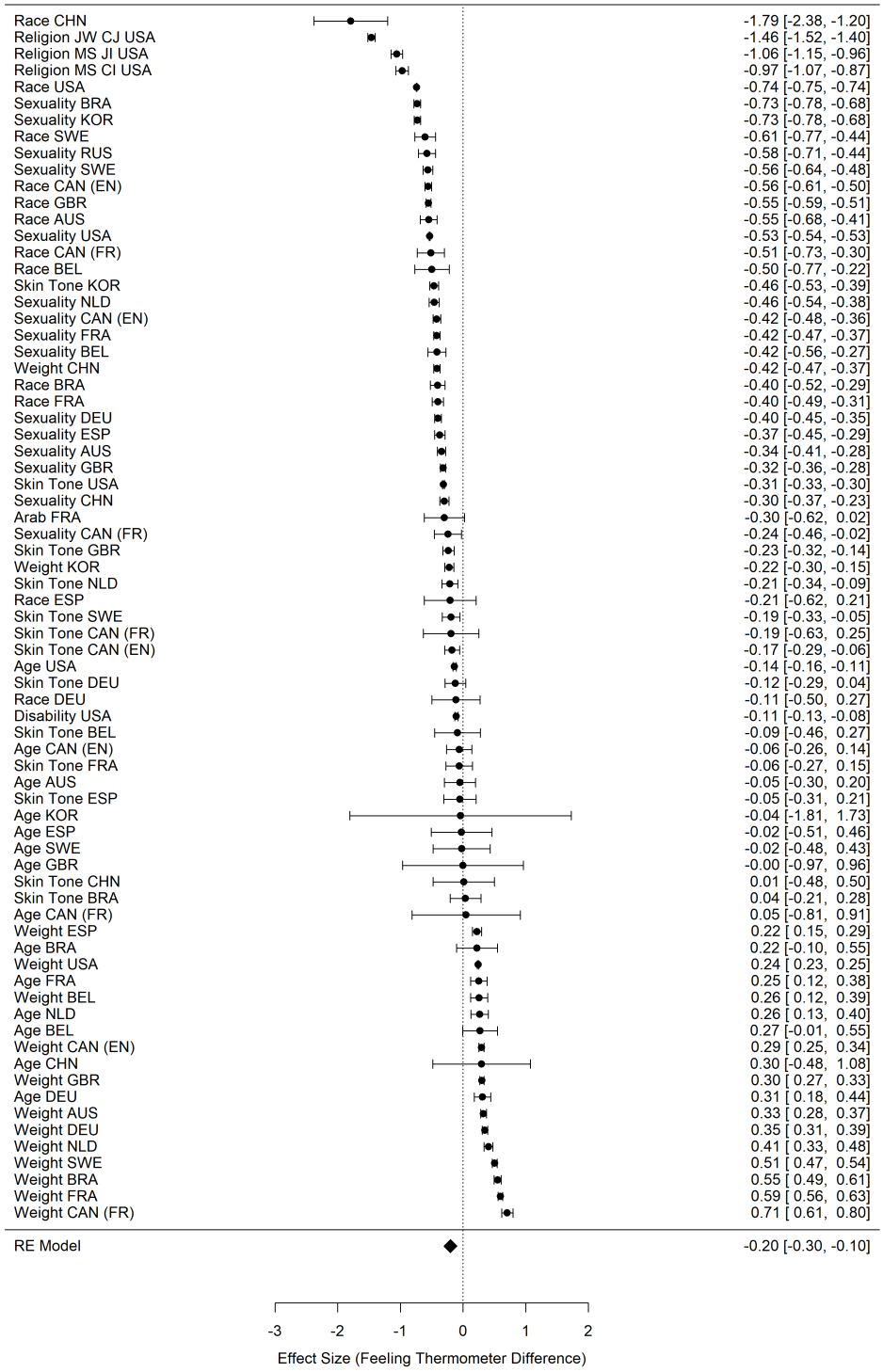

Figure 4. Caterpillar plot of random-effects meta-analysis of feeling thermometer (difference) scores with study effects ordered by effect size. Positive scores indicate outgroup favoritism and negative scores indicate ingroup favoritism from the perspective of the disadvantaged groups. Error bars depict 95% confidence intervals and values in squared brackets indicate lower and upper bounds of confidence intervals. Dataset labels denote the intergroup domain for each study and the respective country. Abbreviations for religious groups: MS = Muslim participants, JW = Jewish participants, JI = Judaism vs. Islam, CJ = Christianity vs. Judaism, CI = Christianity vs. Islam. Country codes: AUS = Australia, BEL = Belgium, BRA = Brazil, CAN (EN) = Canada (English), CAN (FR) = Canada (French), CHN = China, DEU = Germany, ESP = Spain, FRA = France, GBR = United Kingdom, KOR = Korea, NLD = The Netherlands, RUS = Russia, SWE = Sweden, USA = United States.

IAT* D *Scores.

We observed an average effect of $r = .08$, $z = 10.80$, $p < .001$, 95% *CI* [0.06; 0.09], indicating that the correlation between conservatism and IAT *D* Scores at the individual level was very small. The estimated amount of total heterogeneity was $\tau^2 = 0.0023$, $Q(71) = 2{,}532.83$, $p < .001$, accounting for a large proportion of the total variability, $I^2 = 94.73\%$.

One-item preference scores.

We observed an average effect of $r = .07$, $z = 6.79$, $p < .001$, 95% *CI* [0.05; 0.09], indicating that the correlation between conservatism and one-item preference scores at the individual level was very small. The estimated amount of total heterogeneity was $\tau^2 = 0.01$, $Q(71) = 2{,}486.61$, $p < .001$, accounting for a large proportion of the total variability, $I^2 = 97.80\%$.

Feeling thermometer difference scores.

We observed an average effect of $r = .08$, $z = 7.12$, $p < .001$, 95% *CI* [0.06; 0.10], indicating that the correlation between conservatism and feeling thermometer difference scores at the individual-level was very small. The estimated amount of total heterogeneity was $\tau^2 = 0.01$, $Q(71) = 2{,}915.18$, $p < .001$, accounting for a large proportion of the total variability, $I^2 = 98.37\%$.

Taken together, individual-level analyses indicate that members of disadvantaged groups displayed outgroup favoritism on implicit measures, but no intergroup preference or ingroup favoritism on explicit measures. Furthermore, individuals who self-report being more conservative also displayed more favorable evaluations of the advantaged group relative to the disadvantaged group on all three measures. Additionally, and importantly, these average effects of intergroup evaluation were characterized by high levels of heterogeneity. Therefore, we conducted follow-up group-level analyses to examine whether this heterogeneity could be explained by a number of theoretically-derived moderators.

Testing and Extending the Predictions of SJT at the Social Group Level

The present research compares intergroup evaluations measured from a wide variety of disadvantaged groups. SJT does not make specific predictions about how relative levels of intergroup evaluation should vary across disadvantaged social groups *per se*. However, visual inspection of the caterpillar plots in Figure 2, 3, and 4 suggests that effects are clustered by social groups. For example, in all three figures, studies examining evaluations of religious groups and sexuality are clustered in the upper end of the distribution, indicating ingroup favoritism, whereas studies examining age- and weight-related evaluations are largely clustered in the bottom of the distribution, indicating outgroup favoritism. These data provide the opportunity to examine social group as a moderator in exploratory analyses, in order to assess the generalizability of the predictions of SJT across different intergroup domains. Additionally, we examined whether the relationships between conservatism and outgroup favoritism observed at the individual level also persists at the social group level. More specifically, we examined whether different disadvantaged groups display different levels of conservatism (i.e., are positioned differently on a conservatism-liberalism dimension) and whether group levels of conservatism are related to their average preferences for the ingroup relative to the outgroup. Finally, we examined the relationship between stigma and intergroup evaluation among disadvantaged groups. Whereas most previous research in this domain has relied on individuals' self-reported perceptions of stigma against their ingroup (e.g., Pinel, 1999), in the present research we operationalized stigma in terms of the intergroup biases of everybody else in the sample, which aligns more closely with classic conceptualizations of stigma (e.g., Link & Phelan, 2001) and treats stigma as an objectively measurable cultural phenomenon.

Social group. To explore the degree to which intergroup evaluation in disadvantaged groups was moderated by the intergroup domain, we treated the different IAT versions as proxies for the intergroup domain in a series of mixed-effects meta

regressions.

IAT* D *Scores.

We first fitted a mixed-effects model, treating IAT version as a categorical moderator, dummy-coding each level of the moderator (i.e., each IAT version) and using disadvantaged group members' IAT *D* Scores as the dependent variable. This analysis yielded a significant moderation effect, $Q_M(10) = 877.85$, $p < .001$. We observed negative IAT effects for the Religion IAT, d_z = -0.90, $SE = 0.22$, $p < .001$, 95% *CI* [-1.32; -0.47], and Sexuality IAT, d_z = -0.63, $SE = 0.22$, p .005, 95% *CI* [-1.07; -0.20], indicating ingroup favoritism. In contrast, we observed positive IAT effects for the Skin Tone IAT, d_z = -0.99, $SE = 0.22$, $p < .001$, 95% *CI* [-1.42; -0.55], the Weight IAT, d_z = -0.39, $SE = 0.06$, $p < .001$, 95% *CI* [-0.50; -0.27], the Age IAT, $d_z = 1.33$, $SE = 0.06$, $p < .001$, 95% *CI* [1.21; 1.45], and the Disability IAT, $d_z = 0.89$, $SE = 0.21$, $p < .001$, 95% *CI* [0.47; 1.31], indicating outgroup favoritism. The remaining null IAT effects indicated neither ingroup nor outgroup favoritism on the Arab IAT, $d_z = 0.33$, $SE = 0.22$, $p = .137$, 95% *CI* [-0.10; 0.76], and the Race IAT, $d_z = 0.15$, $SE = 0.08$, $p = .061$, 95% *CI* [-0.01; 0.30].

One-item preference scores.

We next fitted the same mixed-effects model, now using disadvantaged group members' one-item preference scores as the dependent variable. This analysis also yielded a significant moderation effect, $Q_M(10) = 600.46$, $p < .001$. We observed negative effects for one-item preference scores for the Black vs. White comparisons, d_z = -0.39, $SE = 0.07$, $p <$ .001, 95% *CI* [-0.53; -0.24], religious ingroup vs. outgroup comparisons, d_z = -2.06, $SE =$ 0.22, $p < .001$, 95% *CI* [-2.49; -1.63], and Gay vs. Straight comparisons, d_z = -1.77, $SE =$ 0.22, $p < .001$, 95% *CI* [-2.20; -1.34], indicating ingroup favoritism. In contrast, we observed positive effects for dark-skinned vs. light-skinned comparisons, d_z = -1.85, $SE =$ 0.22, $p < .001$, 95% *CI* [-2.28; -1.42], overweight vs. normal weight comparisons, d_z = -0.58, $SE = 0.06$, $p < .001$, 95% *CI* [-0.69; -0.47], and old vs. young comparisons, $d_z = 0.31$, *SE*

= 0.07, $p < .001$, 95% *CI* [0.18; 0.43], indicating outgroup favoritism. The remaining null effects indicated neither ingroup nor outgroup favoritism for disabled vs. abled comparisons, $d_z = 0.23$, $SE = 0.22$, $p = .291$, 95% *CI* [-0.20; 0.66] and Arab vs. White comparisons, $d_z = -0.24$, $SE = 0.22$, $p = .285$, 95% *CI* [-0.68; 0.20].

Feeling thermometer difference scores.

We fitted a third mixed-effects model in a similar fashion, now using disadvantaged group members' feeling thermometer difference scores as the dependent variable. This analysis also yielded a significant moderation effect, $Q_M(10) = 288.28$, $p < .001$. We observed negative effects for Black vs. White evaluations, $d_z = -0.54$, $SE = 0.07$, $p < .001$, 95% *CI* [-0.67; -0.41], religious ingroup vs. outgroup evaluations, $d_z = -1.46$, $SE = 0.20$, $p < .001$, 95% *CI* [-1.85; -1.07], Gay vs. Straight evaluations, $d_z = -0.97$, $SE = 0.21$, $p < .001$, 95% *CI* [-1.37; -0.57], and dark-skinned vs. light-skinned evaluations, $d_z = -1.06$, $SE = 0.20$, $p = < .001$, 95% *CI* [-1.46; -0.66], indicating ingroup favoritism. In contrast, we observed positive effects for overweight vs. normal weight evaluations, $d_z = -0.46$, $SE = 0.05$, $p < .001$, 95% *CI* [-0.56; -0.35], indicating outgroup favoritism. The remaining null effects indicated neither ingroup nor outgroup favoritism for old vs. young evaluations, $d_z = 0.11$, $SE = 0.07$, $p = .116$, 95% *CI* [-0.03; 0.25], Arab vs. White evaluations, $d_z = -0.30$, $SE = 0.26$, $p = .246$, 95% *CI* [-0.80; 0.21], and disabled vs. abled evaluations, $d_z = -0.11$, $SE = 0.20$, $p = .589$, 95% *CI* [-0.50; 0.28].

Conservatism. To test the relationship between disadvantaged groups' conservatism and intergroup evaluation at the level of the social group, we fitted three separate mixed-effects meta-regression models with self-reported conservatism aggregated at the sample level as continuous moderators (see Table 2, column 4, for sample level aggregates of conservatism), and using disadvantaged groups' sample-aggregated IAT *D* Scores, one-item preference scores, and feeling thermometer difference scores as dependent variables. Figure 5 depicts the relationship between sample averages of conservatism (y-axes) and sample averages of intergroup evaluations (x-axes) among disadvantaged

groups, with each panel corresponding to a different measure of intergroup evaluation.

IAT* D *Scores.

First, we fitted a mixed-effects model with sample aggregates of self-reported conservatism as a continuous moderator, using disadvantaged groups' IAT *D* Scores as dependent variable. We observed a significant moderating effect, $Q_M(1) = 16.53$, $p < .001$, which accounted for 20.18% of the heterogeneity. This indicates that samples with higher averages of self-reported conservatism were more likely to demonstrate implicit outgroup favoritism (see Figure 5, Panel A).

One-item preference scores.

Again, we fitted a mixed-effects model with sample aggregates of self-reported conservatism as a continuous moderator, now using disadvantaged groups' one-item preference scores as dependent variable. We observed a significant moderating effect, $Q_M(1) = 14.83$, $p < .001$, which accounted for 15.98% of the heterogeneity. This indicates that samples with higher averages of self-reported conservatism were more likely to demonstrate outgroup favoritism on one-item preference scores (see Figure 5, Panel B).

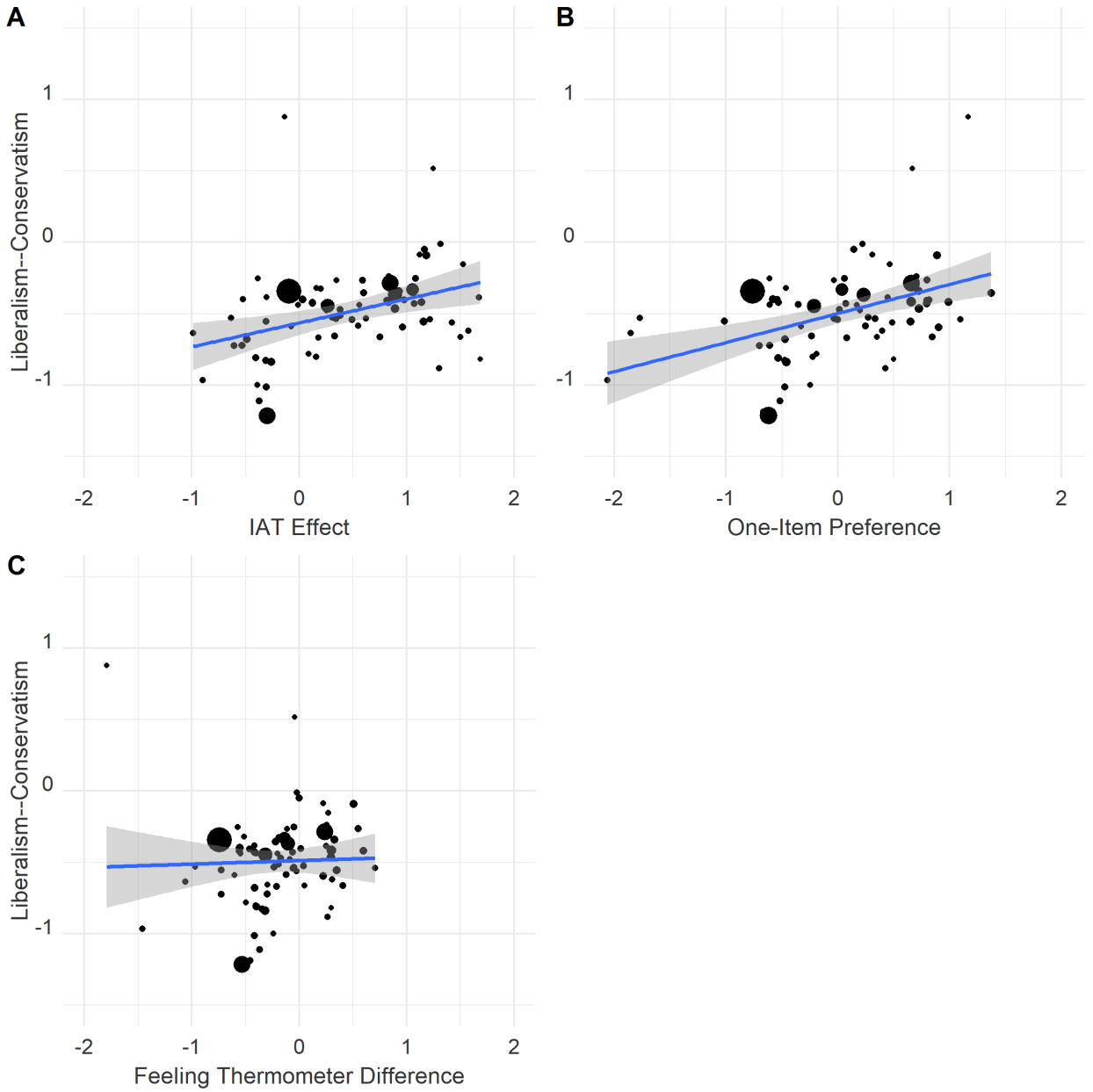

Figure 5. Scatterplot depicting the relationship between conservatism and intergroup evaluations among disadvantaged groups. Y-axes reflect disadvantaged groups' sample-level mean Cohen's d_z for conservatism, measured on a one-item 6- and/or 7-point scale. X-axes reflect mean Cohen's d_z for IAT *D* Scores (Panel A), one-item preference scores (Panel B), and feeling thermometer difference scores (Panel C). Each circle corresponds to a different social group, with circle size reflecting sample size. Positive values on the y-axes indicate more conservative attitudes. Positive values on the x-axes indicate more favorable evaluations of the advantaged group relative to the disadvantaged group, which reflects outgroup favoritism for these samples of disadvantaged groups.

Feeling thermometer difference scores.

Lastly, we fitted the same mixed-effects model with feeling thermometer difference scores as the dependent variable and sample aggregates of self-reported conservatism as a

continuous moderator. Here, we did not observe a significant moderating effect, $Q_M(1) = 0.78$, $p = .376$, with conservatism accounting for only 1.18% of the heterogeneity in feeling thermometer difference scores. This indicates that sample averages of self-reported conservatism were unrelated to sample averages of feeling thermometer difference scores (see Figure 5, Panel C).

Stigma. We operationalized stigma in terms of how a disadvantaged group is evaluated by people who are not members of that social group (i.e., "the rest of society"). Specifically, stigma estimators were manifest in the present analyses using three measures: IAT *D* Scores, one-item preference scores, and feeling thermometer difference scores collected from all participants who indicated not belonging to the disadvantaged group in each study. We used these stigma measures as continuous moderators in a series of mixed-effects meta-regression models, with disadvantaged groups' IAT *D* Scores, one-item preference scores, and feeling thermometer difference scores as dependent variables.

Figure 6 depicts the relationship between the three measures of stigma and the three measures of intergroup evaluation among disadvantaged groups, with positive values on both the x- and y-axes indicating more favorable evaluations of the advantaged group relative to the disadvantaged group. Consequently, positive values on the y-axes reflect outgroup favoritism among disadvantaged groups, and positive values on the x-axes reflect higher levels of stigma against disadvantaged groups.

IAT* D *Scores.

We first fitted three separate mixed-effects meta-regression models with the three stigma measures as continuous moderators, using disadvantaged groups' IAT *D* Scores as dependent variable. We observed a significant moderation effect of IAT *D* Score stigma on disadvantaged groups' IAT *D* Scores, $Q_M(1) = 170.55$, $p < .001$, accounting for 71.31% of heterogeneity; a non-significant moderation effect of one-item preference score stigma on disadvantaged groups' IAT *D* Scores, $Q_M(1) = 3.57$, $p = .059$, accounting for 3.53% of

heterogeneity; and a significant moderation effect of feeling thermometer difference score stigma on disadvantaged groups' IAT *D* Scores, $Q_M(1) = 7.28$, $p = .007$, accounting for 8.08% of heterogeneity.[13] Taken together, these results indicate that disadvantaged groups displayed more implicit outgroup favoritism to the extent that their social group was stigmatized by others in terms of the implicit measure, an effect which accounted for substantial amounts of the heterogeneity. The same descriptive but non-significant trend was observed in terms of explicit, one-item preference score stigma. However, the opposite trend was observed in terms of explicit, feeling thermometer difference score stigma (see Figure 6, Panels A, B, C).

One-item preference scores.

We fitted the same three mixed-effects meta-regression models with the three stigma measures as continuous moderators, this time using disadvantaged groups' one-item preference scores as dependent variable. We observed a significant moderation effects of IAT *D* Score stigma on disadvantaged groups' one-item preference scores, $Q_M(1) = 23.28$, $p < .001$, accounting for 23.81% of heterogeneity; a significant moderation effect of one-item preference score stigma on disadvantaged groups' one-item preference scores, $Q_M(1) = 25.05$, $p < .001$, accounting for 25.44% of heterogeneity; and a non-significant moderation effect of feeling thermometer difference score stigma on disadvantaged groups' one-item preference scores, $Q_M(1) = 1.38$, $p = .239$, accounting for 0.36% of heterogeneity. Taken together, these results indicate that disadvantaged groups displayed more explicit outgroup favoritism to the extent that their social group was stigmatized by others in terms of the implicit measure and the explicit one-item preference measure. However, feeling thermometer difference score stigma was unrelated to disadvantaged groups' one-item preference scores (see Figure 6, Panel D, E, F).

[13] Visual inspection of the Figure 6 (Panel C) suggests that this moderation effect was in the opposite direction than expected. Disadvantaged groups' IAT *D* Scores were higher (indicating more outgroup favoritism) the *less* they were stigmatized by others on feeling thermometers.

Feeling thermometer difference scores.

Finally, we fitted the same three mixed-effects meta-regression models with the three stigma measures as continuous moderators, this time using disadvantaged groups' feeling thermometer difference scores as dependent variable. We observed a significant moderation effects of IAT *D* Score stigma on disadvantaged groups' feeling thermometer difference scores, $Q_M(1) = 27.62$, $p < .001$, accounting for 30.00% of heterogeneity; a significant moderation effect of one-item preference score stigma on disadvantaged groups' feeling thermometer difference scores, $Q_M(1) = 12.68$, $p < .001$, accounting for 17.29% of heterogeneity; and a non-significant moderation effect of feeling thermometer difference score stigma on disadvantaged groups' feeling thermometer difference scores, $Q_M(1) = 3.36$, $p = .067$, accounting for 4.62% of heterogeneity. These results indicate that disadvantaged groups displayed less explicit ingroup favoritism on feeling thermometer difference scores to the extent that their social group was stigmatized by others on the implicit measure and (descriptively) both explicit measures (see Figure 6, Panel G, H, I).

Conservatism versus stigma. According to SJT, both conservatism and stigma should moderate intergroup bias among disadvantaged groups, but the theory makes no predictions about the relative influences of each of these constructs. The present research provides an opportunity to examine whether conservatism and stigma account for heterogeneity above and beyond the heterogeneity accounted for by the other. We examined this in a meta-analytic framework, by fitting hierarchical multivariate meta-regression models with intergroup evaluation as the dependent variable and sample averages of conservatism and the three stigma measures as moderators (see Harrer, Cuijpers, Furukawa, & Ebert, 2019).[14] For each measure of intergroup evaluation, we first fitted a reduced model with disadvantaged groups' sample aggregates of conservatism as a

[14] Correlational analyses of the three measures of stigma indicate that they are not highly correlated, thus making multicollinearity unlikely: IAT *D* Score stigma and one-item preference score stigma, $r(71) = .15$, $p = 0.20$, 95% *CI* [-0.08; 0.37]; IAT *D* Score stigma and feeling thermometer difference score stigma, $r(71) = -.21$, $p = 0.08$, 95% *CI* [-0.42; 0.02]; one-item preference score stigma and feeling thermometer difference score stigma, $r(71) = .22$, $p = 0.06$, 95% *CI* [-0.01; 0.43].

continuous moderator. Next, we fitted the full model, adding the three measures of stigma. We then compared the model fit of both models by using a likelihood ratio test (LRT) and by comparing *Akaike's Information Criterion* (AIC) values for both models. Lastly, in order to control for Type I error, we tested the robustness of the full model by performing a permutation test with 1,000 iterations (see Higgins & Thompson, 2004; Viechtbauer, Lopez-Lopez, Sanchez-Meca, & Marin-Martinez, 2015).

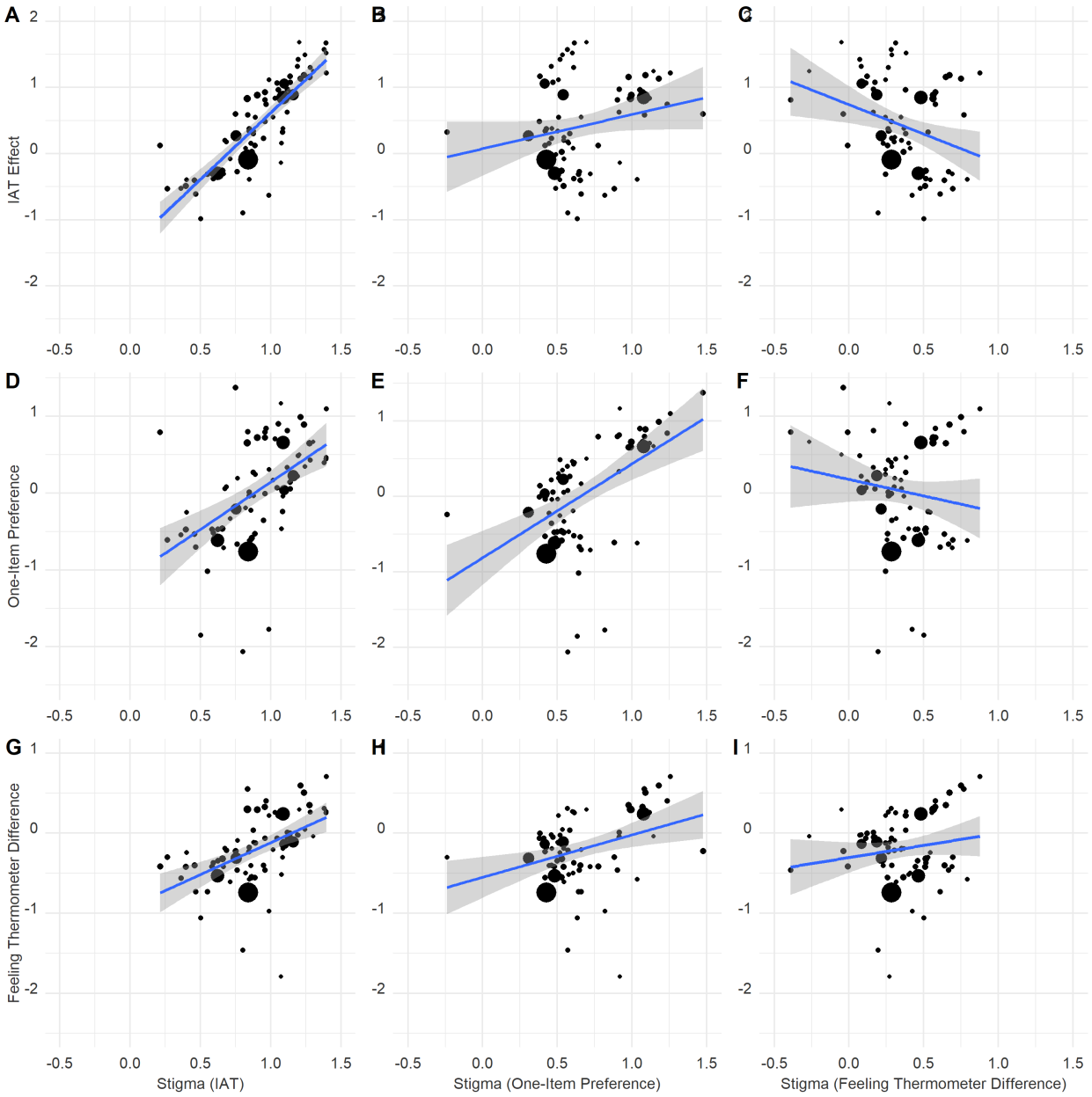

Figure 6. Scatterplots depicting the relationship between stigma and intergroup evaluation among disadvantaged groups. Y-axes reflect mean Cohen's d_z for disadvantaged groups' IAT *D* Scores (Panels A-C), one-item preference scores (Panel D-F), and feeling thermometer difference scores (Panel G-I). X-axes reflect stigma, operationalized as mean Cohen's d_z for non-disadvantaged group members' IAT *D* Scores (left column), one-item preference scores (middle column), and feeling thermometer difference scores (right column). Positive values on both the x- and y-axes indicate more favorable evaluations of the advantaged group relative to the disadvantaged group. Consequently, positive values on the y-axes reflect outgroup favoritism among disadvantaged groups, and positive values on the x-axes reflect stigma against disadvantaged groups.

IAT D *Scores.*

We first fitted a mixed-effects meta-regression model with sample averages of conservatism as a continuous moderator, using disadvantaged groups' sample averages of

IAT *D* Scores as dependent variable. We observed a significant overall moderating effect, $F_{1,71} = 16.36$, $p < .001$, which accounted for 21.29% of the heterogeneity. Next, we fitted the full mixed-effects meta-regression model, adding the three measures of stigma. Again, we observed a significant overall moderating effect, $F_{4,68} = 52.30$, $p < .001$. Crucially, the moderating effect of conservatism became non-significant in the full model, $\beta = 0.17$, $p = .325$, 95% *CI* [-0.17; 0.51], whereas IAT *D* Score stigma, $\beta = 1.85$, $p < .001$, 95% *CI* [1.52; 2.17], one-item preference score stigma, $\beta = 0.33$, $p = .040$, 95% *CI* [0.02; 0.64], and feeling thermometer difference score stigma, $\beta = -0.49$, $p = .014$, 95% *CI* [-0.87; -0.10], remained significant predictors of disadvantaged groups' IAT *D* Scores.[15] The full model accounted for 76.68% of the heterogeneity. The likelihood ratio test indicated that the full model indeed had a better fit than the reduced model, $\chi^2 = 85.24$, $p < .001$, and AIC values were lower (indicating better fit) for the full model, $AIC_c = 64.69$, than for the reduced model, $AIC_c = 143.01$. Lastly, we performed a permutation test, providing evidence for the robustness of the full model, $F_{4,68} = 52.30$, $p = .001$. This indicates that the effect of disadvantaged groups' level of conservatism on implicit outgroup favoritism was fully accounted for by stigma. Moreover, stigma accounted for heterogeneity above and beyond the heterogeneity accounted for by conservatism. In other words, at the social group level, stigma explained more variance in implicit intergroup evaluation than conservatism, and conservatism had no independent effect on IAT *D* Scores.

One-item preference scores.

Again, we fitted a mixed-effects meta-regression model with sample averages of conservatism as a continuous moderator, now using disadvantaged groups' sample averages of one-item preference scores as dependent variable. We observed a significant overall moderating effect, $F_{1,71} = 15.22$, $p < .001$, which accounted for 17.06% of the heterogeneity. Next, we fitted the full mixed-effects meta-regression model, adding the three measures of

[15] Note however that other participants' feeling thermometer difference scores were negatively correlated with disadvantaged groups' IAT *D* Scores.

stigma. Again, we observed a significant overall moderating effect, $F_{4,68} = 16.16$, $p < .001$. Crucially, the moderating effect of conservatism became non-significant in the full model, $\beta = 0.27$, $p = .238$, 95% *CI* [-0.18; 0.73], whereas IAT *D* Score stigma, $\beta = 0.87$, $p < .001$, 95% *CI* [0.40; 1.34] and one-item preference score stigma, $\beta = 1.13$, $p < .001$, 95% *CI* [0.67; 1.59], remained significant predictors of disadvantaged groups' one-item preference scores. The effect of feeling thermometer difference score stigma on disadvantaged groups' one-item preference scores was non-significant, $\beta = -0.42$, $p = .139$, 95% *CI* [-0.98; 0.14]. The full model accounted for 48.38% of the heterogeneity. The likelihood ratio test indicated that the full model indeed had a better fit than the reduced model, $\chi^2 = 33.64$, $p < .001$, and AIC values were lower for the full model, $AIC_c = 118.11$, than for the reduced model, $AIC_c = 144.82$. Lastly, we performed a permutation test, providing evidence for the robustness of the full model, $F_{4,68} = 16.16$, $p = .001$. This indicates that the effect of disadvantaged groups' level of conservatism on explicit intergroup evaluation was fully accounted for by stigma. Moreover, stigma accounted for heterogeneity above and beyond the heterogeneity accounted for by conservatism. In other words, at the social group level, stigma explained more variance in explicit intergroup evaluation than conservatism, and conservatism had no independent effect on one-item preference scores.

Feeling thermometer difference scores.

Lastly, we fitted a mixed-effects meta-regression model with sample averages of conservatism as a continuous moderator, now using disadvantaged groups' sample averages of feeling thermometer difference scores as dependent variable. The overall moderating effect was non-significant, $F_{1,71} = 0.81$, $p = .371$, and accounted for 2.60% of the heterogeneity. Next, we fitted the full mixed-effects meta-regression model, adding the three measures of stigma. Here, we observed a significant overall moderating effect, $F_{4,68} = 13.25$, $p < .001$. Crucially, the moderating effect of conservatism remained non-significant in the full model, $\beta = -0.14$, $p = .368$, 95% *CI* [-0.46; 0.17], whereas IAT *D* Score stigma, $\beta = 0.87$, $p < .001$, 95% *CI* [0.55; 1.18] and one-item preference score stigma, $\beta = 0.44$, $p =$

.007, 95% *CI* [0.12; 0.75], were significant predictors of disadvantaged groups' feeling thermometer difference scores, and the effect of feeling thermometer difference score stigma was non-significant, $\beta = 0.38$, $p = .054$, 95% *CI* [-0.01; 0.77]. The full model accounted for 48.36% of the heterogeneity. The likelihood ratio test indicated that the full model indeed had a better fit than the reduced model, $\chi^2 = 40.14$, $p < .001$, and AIC values were lower for the full model, $\text{AIC}_\text{c} = 61.13$, than for the reduced model, $\text{AIC}_\text{c} = 94.34$. Lastly, we performed a permutation test, providing evidence for the robustness of the full model, $F_{4,68} = 13.25$, $p = .001$. This indicates that the effect of disadvantaged groups' level of conservatism on explicit intergroup evaluation was fully accounted for by stigma. Moreover, stigma accounted for heterogeneity above and beyond the heterogeneity accounted for by conservatism. In other words, at the social group level, stigma explained more variance in explicit intergroup evaluation than conservatism, and conservatism had no independent effect on feeling thermometer difference scores.

Taken together, group-level meta-analyses extend SJT by testing its predictions at a new unit of analyses. Exploratory analyses using social group as a moderator revealed that intergroup evaluations in disadvantaged groups were moderated by the intergroup domain: Whereas some disadvantaged groups consistently displayed outgroup favoritism, others consistently displayed ingroup favoritism, and yet others displayed diverging patterns on implicit and explicit measures. Additionally, group-level conservatism consistently moderated disadvantaged groups' implicit intergroup evaluations, but inconsistently moderated their explicit intergroup evaluations. Similarly, stigma operationalized in terms of others' implicit intergroup evaluations consistently moderated disadvantaged groups' intergroup evaluations, whereas stigma operationalized in terms of others' explicit intergroup evaluations inconsistently moderated disadvantaged groups' intergroup evaluations. Finally, when both conservatism and stigma were entered into the same models, only stigma was consistently related to intergroup evaluations of disadvantaged groups.

General Discussion

The present research used large datasets from 73 samples of online participants collected in 14 countries to investigate moderators of intergroup evaluation in a wide variety of disadvantaged groups. We tested SJT's predictions, whenever possible, at two levels of analysis: the individual level and the social group level. At the individual level, members of disadvantaged groups on average displayed a medium-sized effect of outgroup favoritism on the IAT, but either a small effect of ingroup favoritism or no intergroup preference on two explicit measures. These findings are consistent with SJT's predicted dissociation between implicit and explicit measures (Jost et al., 2004). In follow-up, exploratory analyses that treated social group as a moderator, intergroup evaluation among disadvantaged groups was moderated by the intergroup domain: Whereas some disadvantaged groups consistently displayed outgroup favoritism, others consistently displayed ingroup favoritism, and yet others displayed diverging patterns on implicit and explicit measures. Additionally, and supporting the predictions of SJT, implicit and explicit intergroup evaluations were moderated by self-reported conservatism. Importantly, the magnitude of effect sizes depended on the level of analysis, indicating small effects at the individual level and medium-sized effects at the social group level. Lastly, at the social group level, disadvantaged groups displayed higher levels of outgroup favoritism the more negatively their own social group was evaluated relative to an advantaged outgroup in their societal context, which supports the hypothesis that stigma is related to outgroup favoritism in disadvantaged groups (Jost & Banaji, 1994; Jost et al., 2004).

SJT at the Level of the Individual Versus the Social Group

The present research underscores the utility of conducting analyses at both the individual and social group level. By testing the predictions of SJT at two levels of analysis, our findings provide insight into qualitatively distinct psychological processes: the

individual-level analyses reflect individual differences, whereas the group-level analyses reflect group processes.

Conservatism. We observed marked differences in the magnitude of the relationships between conservatism and intergroup evaluations between units of analysis: Whereas moderator analyses conducted at the social group level (i.e., between samples) revealed that conservatism accounted for considerable proportions of the variance of intergroup evaluations between disadvantaged groups, analyses conducted at the individual level (i.e., within-samples) revealed that conservatism accounted for little variance in intergroup evaluations within disadvantaged groups. These large differences between between-sample and within-sample analyses are striking—but have been observed in other research on implicit bias as well (sf. Payne, Vuletich, & Lundberg, 2017). One explanation for these large differences might be that within-sample correlations were attenuated by the relative unreliability of the measures (e.g., Hunter & Schmidt, 2007), whereas aggregation at the social group level in our between-sample analyses likely reduced measurement error (e.g., Rivers, Rees, Calanchini, & Sherman, 2017; Rushton et al., 1983). However, the differences appear too pronounced to gloss over as statistical artifact. We therefore offer and discuss speculations about the underlying processes that may explain these differences between levels of analysis that we hope may inspire future research in this domain.

At the individual level, conservatism of individual members of disadvantaged groups is the unit of analysis. We assume that disadvantaged group members' level of conservatism reflects individual differences in the preference for inequality, acceptance of the status quo, or system justifying beliefs more generally (Jost et al., 2004; Jost et al., 2003), as well as other psychological correlates of conservatism (for a review, see Hodson & Dhont, 2015). In turn, individual differences in conservatism are related to the extent to which some group members display more outgroup favoritism than other group members. The small average effect size of these correlations suggests that the differences between individual members' level of conservatism are relatively weakly related to their individual

tendency to display ingroup or outgroup favoritism.

At the social group level, conservatism of the disadvantaged group as a whole is the unit of analysis. Disadvantaged groups inhabit different positions on the conservatism-liberalism spectrum, with some groups on average leaning less liberal (e.g., overweight participants) than others (e.g., Gay and Lesbian participants).[16] We propose that disadvantaged groups' position along the conservatism-liberalism spectrum reflect group-level processes, such as group histories, social norms, or cultural traditions. In the present research, we identify stigma as a group-level process that moderates (and, in fact, fully accounts for) the relationship between group-level conservatism and intergroup evaluations among disadvantaged groups. There are at least two possible explanations for this finding.

One possibility is that the measures of conservatism and stigma may constitute different operationalizations of the same latent construct, with stigma being the superior measure of the latent variable. However, we deem this explanation less likely, given the conceptual differences between stigma—operationalized here as the negative group evaluations by others—and ideological self-placement (cf. Koch, Imhoff, Dotsch, Unkelbach, & Alves, 2016). Another possibility is that stigma might influence conservatism in disadvantaged groups. From this perspective, we propose that the relative frequency of experiences of stigmatization (which can include individuals personally experiencing stigma, as well as hearing reports of stigmatization from fellow group members) may lead disadvantaged groups as a whole to lean more or less conservative over time.

Consistent with the possibility that stigma increases conservatism among disadvantaged groups, members of disadvantaged groups who internalize negative stereotypes and evaluations of their group tend to assimilate to the dominant culture (David, Schroeder, & Fernandez, 2019). Furthermore, some members of disadvantaged

[16] Note that samples leaned fairly liberal on average, as indexed by their negative sign in Table 2, fourth column.

groups cope with stigma by gravitating towards ideological beliefs that provide a sense of safety, such as authoritarianism (Henry, 2011). Thus, the relationships between stigma and ideological beliefs in disadvantaged groups might reflect a self-regulatory strategy. Similarly, SJT proposed that system justifying beliefs would serve a "palliative function" (Jost & Hunyady, 2002), helping members of disadvantaged groups who are stigmatized to cope with negative emotions. Taken together, these ideas about relationships between stigma and ideological beliefs were formulated at the individual level, but could provide a framework for future theorizing at the social group level. Based on this framework, conservatism as a group-level construct may not primarily (or necessarily) reflect individual political ideology, but rather a group-based cultural adaptation process, by which members of disadvantaged groups adhere to a more or less conservative group norm in response to the level of stigma faced by their group. That said, such causal claims remain speculative, given the correlational nature of the present research. Future research—ideally longitudinal—is necessary to investigate a causal effect of stigma on group conservatism, as well as potential cultural or social processes that may mediate and/or moderate such effect(s).

In addition to stigma, future research might investigate other group-level processes to explain disparate relationship between conservatism and intergroup evaluation at the individual versus social group levels. One such group-level process might be *group consciousness* (e.g., Duncan, 1999; Gurin, Miller, & Gurin, 1980; see Ashmore, Deaux, & McLaughlin-Volpe, 2004, for a related concept), the tendency to reflect on the ingroup's relative position in society. Group consciousness can vary between people but also between groups (Gurin et al., 1980). Our finding that conservatism is more strongly related to intergroup evaluations at the social group versus individual level might suggest that disadvantaged groups differ in levels of group consciousness more so than do disadvantaged individuals. Other group-level processes that relate to intergroup evaluations, such as entitativity (**???**) or the permeability of group boundaries (Bettencourt, Charlton, Dorr, &

Hume, 2001), might also help to explain the diverging findings observed here. By continuing to investigate the relationship between group status and intergroup evaluations at multiple levels of analysis, future research may build on the present research to more fully develop SJT as a group-level theory.

Stigma. In the present research, conservatism explained substantial variance in intergroup evaluation between social groups, but this variance was fully accounted for by stigma. Importantly, this relationship was not moderated by measurement type: at the group level, both implicit and explicit intergroup evaluations in disadvantaged groups were related to stigma. One possible interpretation of these results is that stigma influences intergroup evaluations—and with this speculation go the usual caveats about causal claims and correlational data. While future, experimental work is necessary to support this claim, this pattern of results is nevertheless consistent with SJT's proposed "internalization of inferiority" (Jost et al., 2004; Jost & van der Toorn, 2012), as well as with the more general notion that intergroup evaluation depends on how social groups are evaluated by society (Allport, 1954; Dasgupta, 2004; Lane et al., 2005; Livingston, 2002).

The present research not only supports the existing literature on intergroup evaluation among disadvantaged groups, but also extends it with novel findings. For example, prior individual-level research has usually operationalized stigma subjectively, in terms of disadvantaged group members' perceptions of how their social group is evaluated by others (Livingston, 2002; Rudman et al., 2002). In contrast, the present research treats stigma as an objective cultural phenomenon, operationalized as the rest of society's measured evaluations of the disadvantaged group. Consequently, our work offers a novel perspective on why disadvantaged groups sometimes display outgroup favoritism: Disadvantaged groups' evaluations of their own groups appear to align with everyone else's evaluations.

Our finding that stigma moderated disadvantaged groups' intergroup evaluation would seem to be consistent with the Bias of Crowds model (BoC; Payne et al., 2017).

Based on the principle of concept accessibility, the BoC model suggests that implicit bias does not merely reflect personal attitudes but, instead, reflects reflects context-related attitude accessibility. Consequently, implicit bias should be stronger in contexts where people are more frequently exposed to direct or indirect expressions of intergroup bias. According to the BoC model, implicit bias is best understood to reflect concepts that are activated by contextual cues and briefly pass through peoples' minds. From this perspective, implicit bias is better conceptualized as a stable property of places and situations rather than a stable property of people. Thus, the BoC model would seem to suggest that the strong relationship between disadvantaged groups' intergroup evaluations and the rest of society's intergroup evaluations (i.e., stigma) observed in the present research reflects a common context-related cause, such as structural inequality. Moreover, our findings that stigma moderates intergroup evaluations across both implicit and explicit measures may reflect an extension of BoC which, to date, is only articulated in terms of implicit bias—which, in turn, may suggest that the group as unit of analysis is more relevant to BoC than is the measurement approach.

That said, the BoC perspective does not perfectly explain the pattern of results reported here. For example, disadvantaged religious and sexual minority groups consistently demonstrated ingroup favoritism across all measures of intergroup bias. However, these groups are objectively stigmatized by the rest of society, in that non-disadvantaged groups' evaluations reflect preferences for the advantaged over disadvantaged groups. Thus, in at least some cases, the intergroup biases of disadvantaged groups do not perfectly correspond to the intergroup biases of the rest of society, which suggests either that certain biases do not reflect a common (e.g., structural) source, or that the influence of this common source is moderated by other processes (e.g., other individual differences or group processes; Branscombe, Schmitt, & Harvey, 1999; Cadinu & Rothbart, 1996; Jost et al., 2004; Tajfel & Turner, 1979). Future research should continue to examine this.

Open Questions and Future Directions

Conservatism and system justifying beliefs. Our analyses provide support for the prediction that as political conservatism increases, outgroup favoritism becomes more likely among disadvantaged groups (Jost et al., 2004). Nevertheless, the present research points to the need for more theorizing about *how* conservatism translates into intergroup evaluation in disadvantaged groups because the moderating effects of conservatism were inconsistent across attitude measures: the relationships between conservatism and intergroup evaluations observed on the IAT and one-item preference measure did not persist for the feeling thermometer. Research has so far primarily tried to explain links between conservatism and outgroup attitudes in dominant groups (see Hodson & Dhont, 2015, for a review), but has not yet articulated the mechanism by which conservative ideology might shape intergroup evaluation among disadvantaged groups. Future work on the underlying psychological processes and mechanisms of this relationship might benefit from considering the following three perspectives.

A first important step towards a more process-oriented understanding would be to investigate whether higher conservatism among disadvantaged groups is related to more positive evaluations of advantaged outgroups, more negative evaluations of the disadvantaged ingroup, or both. A variety of theoretical perspectives propose that ingroup favoritism primarily reflects positive ingroup evaluations rather than negative outgroup evaluations (e.g., Allport, 1954; Brewer, 1999; Brewer & Campbell, 1976; Greenwald & Pettigrew, 2014; Mummendey & Otten, 1998). However, to date, no perspective makes clear predictions about the relative contributions of positive and negative evaluations to outgroup favoritism in general or among disadvantaged groups specifically, nor are there clear predictions about how conservatism might moderate these evaluations. Developing more refined theories about the relationship between conservatism and outgroup favoritism among disadvantaged groups will help to advance research in this domain.

Second, further research is needed to determine which aspects of conservatism are related to intergroup evaluation among disadvantaged groups. For example, Jost and colleagues (2003) proposed that conservatism is based on two core ideologies: the opposition to change and preference for inequality. In the present research, we relied on a single-item political orientation measure (ranging from conservative to liberal), so we were not able to disentangle the contributions of these two core ideologies. Future research might employ scales assessing sub-components of conservatism, such as acceptance of inequality (e.g., Ho et al., 2015) or opposition to change (e.g., White, Kinney, Danek, Smith, & Harben, 2020), in order to better understand how conservatism moderates intergroup evaluation among disadvantaged groups.

Third, SJT predicts that system justifying tendencies are related to intergroup evaluations (Jost et al., 2004), but conservatism is only one possible manifestation of system justifying beliefs (see Jost & Hunyady, 2005). For example, system justifying beliefs have also been operationalized in terms of economic system justification (Jost & Thompson, 2000), general levels of system justification (Kay & Jost, 2003), and social dominance orientation (SDO; Pratto et al., 1994). Supporting the predictions of SJT (Jost et al., 2004), economic system justification was positively related to outgroup favoritism in South Italians (a disadvantaged group, see Jost et al., 2002). Yet, other studies have found no relationship between economic system justification and outgroup favoritism (Jost & Thompson, 2000, Study 4), or between SDO and intergroup evaluation (Ashburn-Nardo et al., 2003), among Black participants. Complementing previous findings, we conducted auxiliary analyses on a small subset of Project Implicit datasets, which suggest that relationships between system justifying beliefs and outgroup favoritism may depend on the specific measure of system justifying beliefs. These analyses (and their limitations) are described in greater detail in the Supplement. Still, more research and theorizing are needed to clarify which ideologies and belief systems are related to intergroup evaluations among members of disadvantaged groups.

Lastly, and more broadly, the motivational processes underlying intergroup evaluation among members of disadvantaged groups will be better understood to the extent that each hypothesized motivation is measured directly. Specifically, SJT posits that members of disadvantaged groups' motives to see themselves and their ingroup positively are often in conflict with their motives to justify and defend the social systems in which they live (Jost & Banaji, 1994; Jost et al., 2004). The present research relied on a measure of system justification motives that does not distinguish among these three motivations (i.e., political conservatism), so our findings do not provide clear insights regarding this motivational conflict. We thus strongly encourage researchers to directly investigate these proposed motivational structures underlying intergroup evaluations by separately measuring system, social group, and ego motives (cf. Kay & Jost, 2014). Future research will benefit from using measures that provide sufficient granularity to differentiate between motives, and perhaps provide insight into whether motivational conflicts underlie variations in intergroup evaluations among members of disadvantaged groups.

Intergroup domain. Exploratory analyses using social group as a moderator indicated that intergroup evaluation in disadvantaged groups was characterized by a high degree of variability. Some disadvantaged groups always displayed outgroup favoritism on both implicit and explicit measures (e.g., overweight or older participants), whereas other disadvantaged groups always displayed ingroup favoritism on both types of measures (e.g., religious or Gay and Lesbian participants), and others displayed no preference on the implicit measure but ingroup favoritism on the explicit measure (e.g., Black participants). This pattern of results is not easily explained by SJT as it is currently articulated (Jost & Banaji, 1994; Jost et al., 2004), and thus, seems to suggest a boundary condition. We speculate here about why intergroup evaluation might vary across disadvantaged groups.

One recently-raised idea focuses on the role of societal discourses in shaping intergroup biases. Charlesworth and Banaji (2019) proposed that the extent to which society prioritizes issues might account for patterns of intergroup biases. From this

perspective, our finding that disadvantaged groups' intergroup evaluations systematically aligned with the rest of society's intergroup evaluations (i.e., stigma) seems to suggest that both group's evaluations reflect a common influence of social priorities.

Another explanation for why intergroup evaluation might vary across disadvantaged groups focuses on social norms and other meta-evaluations. For example, age-related bias is not prohibited by strong egalitarian norms in Western culture: Bias against older people and in favor of younger people is among the largest and most consensual of biases against a social group in the United States (Levy & Banaji, 2002; Nosek et al., 2002, 2007). Similarly, social norms about the suppression of weight-related prejudice have been shown to be weaker compared to other forms of prejudice (e.g., Crandall & Eshleman, 2003; Crandall, Eshleman, & O'Brien, 2002; Degner & Wentura, 2009). Previous research has also highlighted weight-related bias as a pervasive and unique form of prejudice in that body weight is often regarded as controllable (e.g., Crandall, 1994). Consequently, overweight individuals are frequently seen as responsible (e.g., Tiggemann & Rothblum, 1997) and blamed for their weight and associated stigma—with overweight individuals often sharing these assumptions (e.g., Crandall, 1994; see Crandall, Merman, & Hebl, 2009, for a review). Given that norms and other meta-evaluations vary across groups, future research should investigate the extent to which they moderate intergroup evaluation in disadvantaged groups.

Effects of intergroup domain versus country. Building on the present research's focus on intergroup evaluations as they vary across social groups, future research might also investigate the extent to which intergroup evaluations vary across countries. The distinction between social groups and countries as unit of analysis is important because, on the one hand, some disadvantaged groups might generally be stigmatized more than other disadvantaged groups. However, on the other hand, some countries might be characterized by higher levels of overall stigma than others (e.g., Marini et al., 2013). To test the latter, we report in the Supplement a series of analyses in which country is

included as a categorical moderator. We observed only inconsistent country-level moderation effects on intergroup evaluations in disadvantaged groups. Specifically, country moderated intergroup evaluations on implicit measures, but only inconsistently moderated intergroup evaluations on explicit measures. Furthermore, auxiliary moderator analyses using country-level indices of cultural value dimensions did not reveal consistent relationships between cultural values and intergroup evaluations in disadvantaged groups (see Supplement). Thus, country-level differences do not seem to be consistently related to intergroup evaluations in the context of the present research.

That said, these country-level analyses are complicated by two issues. First, given the rather low power of these analyses ($n = 14$ countries), they are not strong tests of the relationship between country-level factors and disadvantaged groups' intergroup evaluations. Second, different intergroup domains were studied in different countries. Many datasets were available for some countries (e.g., United States, United Kingdom), but relatively fewer datasets were available for other countries (e.g., Korea, Russia). Consequently, an analysis that treats country as a moderator is inherently biased because the moderator variable "country" is not independent of the moderator variable "intergroup domain". Thus, a moderation effect by country might reflect the fact that specific intergroup domains were examined in some countries but not others (i.e., biased selection). Taken together, our auxiliary analyses do not provide consistent evidence for country-level effects on intergroup evaluations in disadvantaged groups. Nevertheless, this issue should be addressed in future theory-driven research.

In contrast to the possibility that some countries might be characterized by higher levels of overall stigma than others, we consistently observed that some disadvantaged groups always displayed ingroup favoritism and others always displayed outgroup favoritism, regardless of their country of origin. One potential interpretation of these moderation effects by intergroup domain is that, regardless of the societal context, certain disadvantaged groups are generally stigmatized more, whereas other disadvantaged groups

are generally stigmatized less. In other words, differences in stigma associated with different social groups might (at least partly) explain the effects of intergroup domain on intergroup evaluations. As an illustration, we have restructured the caterpillar plot of IAT *D* Score effect sizes according to both country and effect size (Figure D1; Supplement). Additionally, in this restructured figure we have plotted the stigma estimates for each intergroup domain for each country. This reconfigured figure illustrates two main takeaways. First, a series of thumbnail copies emerge depicting a consistent pattern of effects across intergroup domains: Within each country's cluster, Gay and Lesbian participants always demonstrate the highest degree of ingroup favoritism, and overweight and older participants always demonstrate the highest degree of outgroup favoritism. Second, the pattern of intergroup evaluations among disadvantaged groups closely aligns with stigma: Sexual identities are always associated with the lowest levels of stigma, and weight and age identities with the highest levels of stigma, leaving stigma levels related to ethnic identities in the middle of the distribution. Taken together, across countries, similar hierarchies emerge for both intergroup evaluations and stigma. Future research might build upon these observations to investigate the extent to which intergroup evaluations among disadvantaged groups are moderated by level of analysis (i.e., social group, country).

Internalization of inferiority. When taken at face value, the observed strong relations between stigma and intergroup evaluations in disadvantaged groups appear consistent with SJT's proposed "internalization of inferiority" (Jost et al., 2004, p. 881). Based on the concept of false consciousness, SJT proposes that disadvantaged groups internalize negative evaluations of and stereotypes about the ingroup (Jost & Banaji, 1994). "[I]nterpreting outgroup favoritism as an indicator of internalization" (Jost et al., 2004, p. 894) is thus a straightforward deduction from this idea. This conclusion, however, would rely on the assumption that measures of intergroup evaluations are (direct or indirect) indicators of internalized attitudes. This assumption is tentative because SJT does not articulate a clear conceptualization of internalization.

The associative-propositional evaluation model (APE; Gawronski & Bodenhausen, 2006) offers two possible conceptualizations of internalization. On the one hand, internalization may refer to any process of associative learning that results in the formation of an internal representation (e.g., Gawronski & Bodenhausen, 2018). The APE model postulates that associative learning reflects spatio-temporal contingencies: the more often members of disadvantaged groups experience information about their group paired with negative evaluation, the more likely they are to form negative associations in long-term memory. Consequently, members of disadvantaged groups will have negative group-relevant experiences in proportion to the extent that their group is stigmatized. From this perspective, associative learning offers a parsimonious explanation of the relationship between stigma and intergroup evaluations observed in the present research—and especially the finding that stigma correlations are strongest for the implicit measure may be interpreted as supporting this rationale.

That said, we are not convinced that this associative learning account captures the gist of the internalization processes presumed to underlie the formation of "false consciousness" in system justification theory. Instead, "consciousness" suggests a degree of introspective awareness and/or deliberation on the subjective truth-value of information. This characterization corresponds to propositional learning, as articulated by the APE model. From this perspective, we would conceptualize a negative group evaluation to be internalized only when the internal representation is based on some degree of awareness and/or subjective acceptance of validity of the negative evaluation of one's own social group.

To the extent that the internalization of inferiority can be conceptualized as an internal representation of associations between one's own group and negative evaluations, a number of further questions emerge. For example, internalized group evaluations might be understood as a structure of relatively stable associations in memory, formed either through associative or propositional learning processes in the course of socialization (Gawronski &

Bodenhausen, 2018). In this case, the strength of the association determines the chronic accessibility of intergroup evaluations within the individual which, in turn, is reflected in responses on implicit measures such as the IAT. However, recent theoretical developments offer an alternative interpretation. For example, the BoC model conceptualizes implicit bias as a "social phenomenon that passes through individual minds" (Payne et al., 2017, p. 236) that does not require the presumption of internalized stable evaluations. In that sense, the effects of outgroup favoritism observed in the present research can be interpreted to reflect cultural contexts where stigmatizing associations are more readily activated than others.

The present research cannot address whether intergroup evaluation in disadvantaged groups reflects internalized stable evaluations, propositions, or situationally-dependent concepts. Instead, the questions we raise here about how internalization of inferiority is conceptualized by SJT (Jost et al., 2004; Jost & van der Toorn, 2012) illustrate that the theory might benefit from more clearly articulating assumptions about the psychological processes underlying intergroup evaluations in disadvantaged groups.

The use of feeling thermometers. In the present research, we observed consistent moderating effects of both conservatism and stigma on IATs and one-item preference measures, but inconsistent effects on feeling thermometer difference scores. Feeling thermometers have been labeled "notoriously unreliable" (Broockman, Kalla, & Aronow, 2015, p. 3), and we think that measurement error could have played a role in these inconsistent findings. Moreover, the calculation of difference scores from feeling thermometers likely increased measurement error, further reducing statistical power in the moderator analyses (Edwards, 1995; Hunter & Schmidt, 2007; Overall & Woodward, 1975). Additionally, IATs and one-item preference measures are structurally similar to one another, in that both target groups are evaluated relative to one another, whereas responses on feeling thermometers are made in the context of one target group at a time. We cannot discern the extent to which measurement error, structural differences, or other conceptual differences among measures contribute to the observed discrepancies. However,

these discrepancies point to the need for further research into measurement properties and the psychological concepts underlying feeling thermometers.

Limitations of the Present Research

One limitation of the present research is that our analyses relied on only one implicit measure: the IAT. As such, the present research is not poised to answer questions regarding implicit measures in general. Moreover, previous research suggests that different implicit measures often do not correlate strongly and, thus, might assess different constructs (e.g., Degner & Wentura, 2009, 2010). Thus, it is unclear whether and to what extent the magnitude and direction of intergroup evaluations among members of disadvantaged groups might depend on the types of implicit measures. Consequently, this limitation highlights the need for a meta-analysis synthesizing research on intergroup evaluations among disadvantaged groups across different implicit measures, which would allow for generalizations beyond one operationalization of the construct (see Essien & Degner, 2020).

Another limitation of the present research is that it relies solely on data from Project Implicit. Participants visit the demonstration website voluntarily and, consequently, are neither random nor representative samples of the general population, or any specific population at all. That said, this limitation is, in part, offset by the size and diversity of the Project Implicit samples, relative to what could reasonably be expected from samples of university undergraduates typically employed in psychological research (Henrich, Heine, & Norenzayan, 2010; Sears, 1986). Consequently, the Project Implicit datasets (Xu et al., 2017, 2018) used in the present research help to advance our understanding of SJT, and intergroup relations more generally, by providing very good statistical power and internal validity to examine a wide variety of social identities, some of which may be rare or otherwise hard to sample.

Concluding Remarks

Using large samples of online participants, the present research examined hypotheses proposed by SJT (Jost et al., 2004) regarding intergroup evaluation in disadvantaged groups. The present research advances the literature by testing these hypotheses at both the individual and social group levels. Across nations and social identities, we found that disadvantaged groups generally displayed outgroup favoritism on the implicit measure, but ingroup favoritism or no intergroup preference on explicit measures. Exploratory analyses revealed that intergroup evaluation in disadvantaged groups was moderated by the intergroup domain. Furthermore, implicit and explicit intergroup evaluations were moderated by political ideology and stigmatization. Taken together, these findings generally support SJT, but at the same time highlight the need for more theory-driven research into the boundary conditions of SJT.

References

Allport, G. W. (1954). *The nature of prejudice*. Cambridge, Mass.: Addison-Wesley.

Ashburn-Nardo, L., Knowles, M. L., & Monteith, M. J. (2003). Black americans' implicit racial associations and their implications for intergroup judgment. *Social Cognition*, *21*(1), 61–87. https://doi.org/10.1521/soco.21.1.61.21192

Ashmore, R. D., Deaux, K., & McLaughlin-Volpe, T. (2004). An organizing framework for collective identity: Articulation and significance of multidimensionality. *Psychological Bulletin*, *130*(1), 80–114. https://doi.org/10.1037/0033-2909.130.1.80

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

Bettencourt, B. A., Charlton, K., Dorr, N., & Hume, D. L. (2001). Status differences and in-group bias: A meta-analytic examination of the effects of status stability, status legitimacy, and group permeability. *Psychological Bulletin*, *127*(4), 520–542. https://doi.org/10.1037/0033-2909.127.4.520

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. https://doi.org/10.1002/jrsm.12

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

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate?

Journal of Social Issues, *55*(3), 429–444. https://doi.org/10.1111/0022-4537.00126

Brewer, M. B., & Campbell, D. T. (1976). *Ethnocentrism and intergroup attitudes: East african evidence*. Oxford, England: Sage.

Broockman, D., Kalla, J., & Aronow, P. (2015). Irregularities in lacour (2014). *Working Paper*.

Cadinu, M. R., & Rothbart, M. (1996). Self-anchoring and differentiation processes in the minimal group setting. *Journal of Personality and Social Psychology*, *70*(4), 661–677. https://doi.org/10.1037/0022-3514.70.4.661

Charlesworth, T. E. S., & Banaji, M. R. (2019). Patterns of implicit and explicit attitudes: I. Long-term change and stability from 2007 to 2016. *Psychological Science*, *30*(2), 174–192. https://doi.org/10.1177/0956797618813087

Cheung, M. W.-L. (2015). *Meta-analysis: A structural equation modeling approach*. Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118957813

Clark, K. B., & Clark, M. P. (1950). Emotional factors in racial identification and preference in negro children. *The Journal of Negro Education*, *19*(3), 341. https://doi.org/10.2307/2966491

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*, 1–21. 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

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, *129*(3), 414–446. https://doi.org/10.1037/0033-2909.129.3.414

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*(3), 359–378. https://doi.org/10.1037/0022-3514.82.3.359

Crandall, C. S., Merman, A., & Hebl, M. (2009). Anti-fat prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 469–487). New York, NY: Psychology Press; US.

Cunningham, W. A., Nezlek, J. B., & Banaji, M. R. (2004). Implicit and explicit ethnocentrism: Revisiting the ideologies of prejudice. *Personality and Social Psychology Bulletin*, *30*(10), 1332–1346. https://doi.org/10.1177/0146167204264654

Czellar, S. (2006). Self-presentational effects in the implicit association test. *Journal of Consumer Psychology*, *16*(1), 92–100. https://doi.org/10.1207/s15327663jcp1601%7B/textunderscore%20%7D11

Dahl, D. B. (2016). *Xtable: Export tables to latex or html*. Retrieved from https://CRAN.R-project.org/package=xtable

Dasgupta, N. (2004). Implicit ingroup favoritism, outgroup favoritism, and their behavioral manifestations. *Social Justice Research*, *17*(2), 143–169. https://doi.org/10.1023/B:SORE.0000027407.70241.15

David, E. J. R., Schroeder, T. M., & Fernandez, J. (2019). Internalized racism: A systematic review of the psychological literature on racism's most insidious consequence. *Journal of Social Issues*, *75*(4), 1057–1086. https://doi.org/10.1111/josi.12350

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.

Degner, J., & Wentura, D. (2010). Automatic prejudice in childhood and early adolescence.

Journal of Personality and Social Psychology, *98*(3), 356–374. https://doi.org/10.1037/a0017993

Devos, T., & Banaji, M. R. (2005). American = white? *Journal of Personality and Social Psychology*, *88*(3), 447–466. https://doi.org/10.1037/0022-3514.88.3.447

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

Duncan, L. E. (1999). Motivation for collective action: Group consciousness as mediator of personality, life experiences, and women's rights activism. *Political Psychology*, *20*(3), 611–635. https://doi.org/10.1111/0162-895X.00159

Edwards, J. R. (1995). Alternatives to difference scores as dependent variables in the study of congruence in organizational research. *Organizational Behavior and Human Decision Processes*, *64*(3), 307–324. https://doi.org/10.1006/obhd.1995.1108

Essien, I., & Degner, J. (2020). *Implicit intergroup evaluations in negatively stigmatized groups: A meta-analysis*. Unpublished manuscript.

Fiedler, K., & Bluemke, M. (2005). Faking the iat: Aided and unaided response control on the implicit association tests. *Basic and Applied Social Psychology*, *27*(4), 307–316. https://doi.org/10.1207/s15324834basp2704%7B/textunderscore%20%7D3

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*(5), 692–731. https://doi.org/10.1037/0033-2909.132.5.692

Gawronski, B., & Bodenhausen, G. V. (2018). Evaluative conditioning from the perspective of the associative-propositional evaluation model. *Social Psychological Bulletin*, *13*(3). https://doi.org/10.5964/spb.v13i3.28024

Gawronski, B., & Hahn, A. (2019). Implicit measures: Procedures, use, and interpretation.

In H. Blanton, J. M. LaCroix, & G. Webster (Eds.), *Measurement in social psychology* (pp. 29–55). New York, NY: Taylor & Francis.

Gawronski, B., Morrison, M., Phills, C. E., & Galdi, S. (2017). Temporal stability of implicit and explicit measures: A longitudinal analysis. *Personality and Social Psychology Bulletin*, *43*(3), 300–312. https://doi.org/10.1177/0146167216684131

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., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, *85*(2), 197–216. https://doi.org/10.1037/0022-3514.85.2.197

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

Gurin, P., Miller, A. H., & Gurin, G. (1980). Stratum identification and consciousness. *Social Psychology Quarterly*, *43*(1), 30. https://doi.org/10.2307/3033746

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

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2-3), 61–135. https://doi.org/10.1017/S0140525X0999152X

Henry, P. J. (2011). The role of stigma in understanding ethnicity differences in authoritarianism. *Political Psychology*, *32*(3), 419–438. https://doi.org/10.1111/j.1467-9221.2010.00816.x

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

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ: British Medical Journal*, *327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., … Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new sdo7 scale. *Journal of Personality and Social Psychology*, *109*(6), 1003–1028. https://doi.org/10.1037/pspi0000033

Hodson, G., & Dhont, K. (2015). The person-based nature of prejudice: Individual difference predictors of intergroup negativity. *European Review of Social Psychology*, *26*(1), 1–42. https://doi.org/10.1080/10463283.2015.1070018

Hoffarth, M. R., & Jost, J. T. (2017). When ideology contradicts self-interest: Conservative opposition to same-sex marriage among sexual minorities—a commentary on pinsof and haselton (2016). *Psychological Science*, *28*(10), 1521–1524. https://doi.org/10.1177/0956797617694866

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

Jost, J. T. (2019). A quarter century of system justification theory: Questions, answers, criticisms, and societal applications. *British Journal of Social Psychology*, *58*(2), 263–314. https://doi.org/10.1111/bjso.12297

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology*, *33*(1), 1–27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x

Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*(6), 881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x

Jost, J. T., Burgess, D., & Mosso, C. O. (2001). Conflicts of legitimation among self, group, and system: The integrative potential of system justification theory. In J. T. Jost & B. Major (Eds.), *The psychology of legitimacy* (pp. 363–388). New York, NY, US: Cambridge University Press.

Jost, J. T., Gaucher, D., & Stern, C. (2015). The world isn't fair": A system justification perspective on social stratification and inequality. In M. Mikulincer & P. R. Shaver (Eds.), *APA handbook of personality and social psychology* (Vol. 2, pp. 317–340). Washington DC: American Psychological Association.

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin*, *129*(3), 339–375. https://doi.org/10.1037/0033-2909.129.3.339

Jost, J. T., & Hunyady, O. (2002). The psychology of system justification and the palliative function of ideology. *European Review of Social Psychology*, *13*(1), 111–153. https://doi.org/10.1080/10463280240000046

Jost, J. T., & Hunyady, O. (2005). Antecedents and consequences of system-justifying ideologies. *Current Directions in Psychological Science*, *14*(5), 260–265. https://doi.org/10.1111/j.0963-7214.2005.00377.x

Jost, J. T., Pelham, B. W., & Carvallo, M. R. (2002). Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology*, *38*(6), 586–602. https://doi.org/10.1016/s0022-1031(02)00505-x

Jost, J. T., & Thompson, E. P. (2000). Group-based dominance and opposition to equality

as independent predictors of self-esteem, ethnocentrism, and social policy attitudes among african americans and european americans. *Journal of Experimental Social Psychology*, *36*(3), 209–232. https://doi.org/10.1006/jesp.1999.1403

Jost, J. T., & van der Toorn, J. (2012). System justification theory. In Van Lange, Paul A. (Ed.), *Handbook of theories of social psychology* (p. 537). Thousand Oaks: Sage Publications Ltd.

Jost, J. T., Wakslak, C. J., & Tyler, T. R. (2008). System justification theory and the alleviation of emotional distress: Palliative effects of ideology in an arbitrary social hierarchy and in society. In *Justice* (Vol. 25, pp. 181–211). Emerald Group Publishing Limited. https://doi.org/10.1016/S0882-6145(08)25012-5

Kay, A. C., & Jost, J. T. (2003). Complementary justice: Effects of poor but happy and poor but honest stereotype exemplars on system justification and implicit activation of the justice motive. *Journal of Personality and Social Psychology*, *85*(5), 823–837. https://doi.org/10.1037/0022-3514.85.5.823

Kay, A. C., & Jost, J. T. (2014). Theoretical integration in motivational science: System justification as one of many autonomous motivational structures. *Behavioral and Brain Sciences*, *37*(02), 146–147. https://doi.org/10.1017/S0140525X13002057

Kite, M. E., Stockdale, G. D., Whitley, B. E., & Johnson, B. T. (2005). Attitudes toward younger and older adults: An updated meta-analytic review. *Journal of Social Issues*, *61*(2), 241–266. https://doi.org/10.1111/j.1540-4560.2005.00404.x

Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The abc of stereotypes about groups: Agency/socioeconomic success, conservative-progressive beliefs, and communion. *Journal of Personality and Social Psychology*, *110*(5), 675. https://doi.org/10.1037/pspa0000046

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and anovas. *Frontiers in Psychology*, *4*, 1–12.

https://doi.org/10.3389/fpsyg.2013.00863

Lane, K. A., Mitchell, J. P., & Banaji, M. R. (2005). Me and my group: Cultural status can disrupt cognitive consistency. *Social Cognition*, *23*(4), 353–386. https://doi.org/10.1521/soco.2005.23.4.353

Levy, B. R., & Banaji, M. R. (2002). Implicit ageism. In T. D. Nelson (Ed.), *Ageism* (pp. 49–75). Cambridge, MA: The MIT Press.

Link, B. G., & Phelan, J. C. (2001). Conceptualizing stigma. *Annual Review of Sociology*, *27*(1), 363–385. https://doi.org/10.1146/annurev.soc.27.1.363

Livingston, R. W. (2002). The role of perceived negativity in the moderation of african americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology*, *38*(4), 405–413. https://doi.org/10.1016/S0022-1031(02)00002-1

Marini, M., Sriram, N., Schnabel, K., Maliszewski, N., Devos, T., Ekehammar, B., … Nosek, B. A. (2013). Overweight people have low levels of implicit weight bias, but overweight nations have high levels of implicit weight bias. *PloS ONE*, *8*(12), e83543. https://doi.org/10.1371/journal.pone.0083543

Mentovich, A., & Jost, J. T. (2008). The ideological id? System justification and the unconscious perpetuation of inequality. *Conneticut Law Review*, *40*(4), 1095–1116.

Miller, D. T. (1999). The norm of self-interest. *American Psychologist*, *54*(12), 1053–1060. https://doi.org/10.1037/0003-066X.54.12.1053

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mummendey, A., & Otten, S. (1998). Positive–negative asymmetry in social discrimination. *European Review of Social Psychology*, *9*(1), 107–143.

https://doi.org/10.1080/14792779843000063

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

Neugarten, B. L. (1974). Age groups in american society and the rise of the young-old. *The ANNALS of the American Academy of Political and Social Science*, *415*(1), 187–198. https://doi.org/10.1177/000271627441500114

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

Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ranganath, K. A., … Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, *18*(1), 36–88. https://doi.org/10.1080/10463280701489053

Overall, J. E., & Woodward, J. A. (1975). Unreliability of difference scores: A paradox for measurement of change. *Psychological Bulletin*, *82*(1), 85–86. https://doi.org/10.1037/h0076158

Oxford Dictionaries. (2019). *Maghreb*. Retrieved from https://en.oxforddictionaries.com/definition/us/Maghreb

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, *28*(4), 233–248. https://doi.org/10.1080/1047840X.2017.1335568

Pinel, E. C. (1999). Stigma consciousness: The psychological legacy of social stereotypes. *Journal of Personality and Social Psychology*, *76*(1), 114–128. Retrieved from http://dx.doi.org/10.1037/0022-3514.76.1.114

Pratto, F., Sidanius, J., Stallworth, L. M., & Malle, B. F. (1994). Social dominance

orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, *67*(4), 741–763. https://doi.org/10.1037/0022-3514.67.4.741

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/

Rivers, A. M., Rees, H. R., Calanchini, J., & Sherman, J. W. (2017). Implicit bias reflects the personal and the social. *Psychological Inquiry*, *28*(4), 301–305. https://doi.org/10.1080/1047840X.2017.1373549

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, *15*(3), 351. https://doi.org/10.2307/2087176

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Thousand Oaks, CA, US: Sage Publications, Inc. https://doi.org/10.4135/9781412984997

Rudman, L. A., & Ashmore, R. D. (2007). Discrimination and the implicit association test. *Group Processes & Intergroup Relations*, *10*(3), 359–372. https://doi.org/10.1177/1368430207078696

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

Rushton, J. P., Brainerd, C. J., & Pressley, M. (1983). Behavioral development and construct validity: The principle of aggregation. *Psychological Bulletin*, *94*(1), 18–38. https://doi.org/10.1037/0033-2909.94.1.18

Sears, D. O. (1986). College sophomores in the laboratory: Influences of a narrow data base on social psychology's view of human nature. *Journal of Personality and Social Psychology*, *51*(3), 515–530. https://doi.org/10.1037/0022-3514.51.3.515

Selvin, H. C. (1958). Durkheim's suicide and problems of empirical research. *American Journal of Sociology*, *63*(6), 607–619. https://doi.org/10.1086/222356

Stanley, D. (2018). *ApaTables: Create american psychological association (apa) style tables*. Retrieved from https://CRAN.R-project.org/package=apaTables

Steffens, M. C. (2004). Is the implicit association test immune to faking? *Experimental Psychology*, *51*(3), 165–179. https://doi.org/10.1027/1618-3169.51.3.165

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; Calif.: Brooks/Cole.

Tiggemann, M., & Rothblum, E. d. (1997). Gender differences in internal beliefs about weight and negative attitudes towards self and others. *Psychology of Women Quarterly*, *21*(4), 581–593. https://doi.org/10.1111/j.1471-6402.1997.tb00132.x

Uhlmann, E., Dasgupta, N., Elgueta, A., Greenwald, A. G., & Swanson, J. (2002). Subgroup prejudice based on skin color among hispanics in the united states and latin america. *Social Cognition*, *20*(3), 198–226. https://doi.org/10.1521/soco.20.3.198.21104

Urbanek, S. (2013). *Png: Read and write png images*. Retrieved from https://CRAN.R-project.org/package=png

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. Retrieved from http://www.jstatsoft.org/v36/i03/

Viechtbauer, W., Lopez-Lopez, J. A., Sanchez-Meca, J., & Marin-Martinez, F. (2015). A comparison of procedures to test for moderators in mixed-effects meta-regression models. *Psychological Methods*, *20*(3), 360–374. https://doi.org/10.1037/met0000023

White, K. R. G., Kinney, D., Danek, R. H., Smith, B., & Harben, C. (2020). The resistance to change-beliefs scale: Validation of a new measure of conservative ideology. *Personality and Social Psychology Bulletin*, *46*(1), 20–35. https://doi.org/10.1177/0146167219841624

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

Wilke, C. O. (2017). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=cowplot

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

Xie, Y. (2016). *Bookdown: Authoring books and technical documents with R markdown*. Boca Raton, Florida: Chapman; Hall/CRC. Retrieved from https://github.com/rstudio/bookdown

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

Xu, F. K., Nosek, B. A., Greenwald, A. G., Ratliff, K. A., Umansky, E., & Lofarno, N. (2018). Project implicit international website datasets. Retrieved from osf.io/kaqi5

Click here to access/download
Supplemental Material
Appendix.pdf