Personality and Social Psychology Review
1–26

DOI: 10.1177/1088868315574461
pspr.sagepub.com

Explaining Extremity in Evaluation of Group Members: Meta-Analytic Tests of Three Theories

B. Ann Bettencourt[1], Mark Manning[2], Lisa Molix[3], Rebecca Schlegel[4], Scott Eidelman[5], and Monica Biernat[6]

Abstract

A meta-analysis that included more than 1,100 effect sizes tested the predictions of three theoretical perspectives that explain evaluative extremity in social judgment: complexity-extremity theory, subjective group dynamics model, and expectancy-violation theory. The work seeks to understand the ways in which group-based information interacts with person-based information to influence extremity in evaluations. Together, these three theories point to the valence of person-based information, group membership of the evaluated targets relative to the evaluator, status of the evaluators' ingroup, norm consistency of the person-based information, and incongruency of person-based information with stereotype-based expectations as moderators. Considerable support, but some limiting conditions, were found for each theoretical perspective. Implications of the results are discussed.

Keywords

Complexity extremity, subjective group dynamics, expectancy violation, impression formation, stereotyping, evaluative extremity

Many models of social judgment suggest that both categorical (group-based) and individuating (person-based) information influence our impressions of people. For example, knowing someone's race along with their grades in school may interact to affect judgments of competence. Prominent dual-process models (Brewer, 1988; Fiske & Neuberg, 1990) highlight the primacy of group-based information in such situations, unless the perceiver is motivated to individuate the target or the target does not fit category-based expectations. Connectionist or parallel process models consider the ways in which the two types of information (and their facilitatory and inhibitory associates) mutually constrain each other's meaning to affect the overall judgment (Kunda & Thagard, 1996).

These models address the question of when categorical versus individuating information dominates impressions and points to specific contexts that maximize the influence of one or the other. For example, outcome dependency and accuracy motives may increase the likelihood of using individuating information in judgment (Neuberg & Fiske, 1987), as may the presence of unambiguous individuating information (Locksley, Borgida, Brekke, & Hepburn, 1980). A meta-analysis of studies that manipulated both categorical and individuating information found that, overall, individuating information carried more weight in judgment (Kunda & Thagard, 1996).

Our meta-analysis focuses on a different question: Given identical individuating information (e.g., strong academic credentials), how does a person's category membership modify judgment? Qualifications may matter more than race in judging competence, but are highly qualified Blacks and Whites evaluated the same? Are poorly qualified Blacks and Whites judged equally incompetent? The research literature suggests answers of "no" to these kinds of questions: A member of one group may be judged more extremely than a member of another group, even given identical individuating information (e.g., Bettencourt, Dill, Greathouse, Charlton, & Mulholland, 1997; Jackson, Hymes, & Sullivan, 1987; Jussim, Coleman, & Lerch, 1987; Linville, 1982; Marques & Yzerbyt, 1988).

[1]University of Missouri, Columbia, USA
[2]Karmanos Cancer Institute, Detroit, MI, USA
[3]Tulane University, New Orleans, LA, USA
[4]Texas A&M University, College Station, USA
[5]University of Arkansas, Fayetteville, USA
[6]University of Kansas, Lawrence, USA

Corresponding Author:
B. Ann Bettencourt, University of Missouri, 210 McAlester Hall, Columbia, MO 65211, USA.
Email: Bettencourta@missouri.edu

Three theories have guided research on this type of evaluative extremity, and each predicts a different pattern as well as different moderators and mediators of extremity effects. These perspectives are complexity-extremity theory (e.g., Linville, 1982; Linville & Jones, 1980), the subjective group dynamics model (e.g., Marques, Abrams, Paez, & Martinez-Taboada, 1998; Marques & Yzerbyt, 1988; Marques, Yzerbyt, & Leyens, 1988), and expectancy-violation theory (e.g., Bettencourt et al., 1997; Jussim et al., 1987). A number of researchers have considered these theories in concert to understand patterns of evaluative extremity (e.g., Bettencourt et al., 1997; Jackson et al., 1987; Jackson, Sullivan, & Hodge, 1993; Jussim et al., 1987; Marques, Robalo, & Rocha, 1992), but these were single-lab investigations. Other meta-analysts have examined contextual moderators of the effects of social categories on judgment, but these have typically focused on only one social category (e.g., only gender), only one evaluative dimension, and have not been guided by theories of evaluative extremity (e.g., Dean, Roth, & Bobko, 2008; Eagly, Makhijani, & Klonsky, 1992; Finkelstein, Burke, & Raju, 1995; J. K. Ford, Kraiger, & Schechtman, 1986; Kite, Stockdale, Whitley, & Johnson, 2005; Roth, Huffcutt, & Bobko, 2003; Sackett & DuBois, 1991; Swim, Borgida, Maruyama, & Myers, 1989).

In our work, we consider these three theories in the context of the broader extant literature on impression formation, including studies of many social groups and a variety of evaluative dimensions. The purpose of our meta-analysis is to consider the theoretical perspectives of complexity-extremity, subjective group dynamics, and expectancy violation, examining the conditions under which each theory's hypotheses are supported in the broader social judgment literature. Focusing on theoretically based moderators, we direct our attention toward the ways in which a person's group membership combines with personal information to produce evaluative extremity. We do not compare the effects of group-based and person-based information on judgment (cf. Kunda & Thagard, 1996) or assess the overall direction of group bias (see Dean et al., 2008; Eagly et al., 1992; Finkelstein et al., 1995; J. K. Ford et al., 1986; Swim et al., 1989). Instead, we ask when identical person-based information is evaluated *differently* depending on the target's group membership.

Understanding evaluative extremity is an important goal, because it moves the field beyond either/or thinking about sources of impression formation. There is little doubt that group-based information and person-based information both matter for social judgment. But it is not the case that high status or positively stereotyped targets are always evaluated more favorably than low status or negatively stereotyped targets, nor is it the case that person-based information always overrides the effects of stereotypes (cf. F. J. Landy, 2008). Instead, social judgment is complex: Valence and other properties of person-based information may matter *differently* depending on group-based information. By examining the conditions under which group- and person-based information interact to affect judgment—leading to evaluative extremity for some targets—we can better understand when bias will emerge and when it will not.

The three theoretical perspectives that guide our meta-analysis suggest that the pattern of evaluative extremity that emerges will depend on whether the target shares group membership with the evaluator, whether the target's actions are positively or negatively valenced, whether the target violates ingroup norms, and whether the target violates stereotype-based expectations. In what follows, we describe the predicted effect of these variables in the context of the three theoretical perspectives.

Complexity-Extremity Theory

Complexity-extremity theory predicts more extreme evaluations of outgroup than ingroup targets and posits relative cognitive complexity as the mechanism responsible for this effect (Linville, 1982; Linville & Jones, 1980).[1] From this perspective, group membership of the target relative to the evaluator is important because of the assumption that people's cognitive representations of ingroups tend to be more complex than their representations of outgroups. Linville (1982) theorized that, because of this low complexity for outgroups, each new "bit" of individuating information carries more weight, pushing evaluations in the direction of the information. As such, the direction of judgmental extremity depends on the valence of the person-based information: "Positive information leads to more favorable ratings of an outgroup than an ingroup; negative information, however, leads to less favorable ratings of an outgroup member" (Linville, 1982, p. 193).

Several studies have demonstrated this judgmental pattern. For example, Linville and Jones (1980) found that White participants judged a high-quality Black law school applicant *more* favorably than an identical White law school applicant, but judged a low-quality Black law school applicant *less* favorably than an equivalent White law school applicant. Likewise, Hart and Morry (1997) found that Black confederates who delivered a speech with persuasive nonverbal behaviors (e.g., eye contact, frequent gesturing) were evaluated more positively by White judges than White confederates who acted similarly. By contrast, when Black confederates with nonpersuasive nonverbal behaviors (e.g., gaze aversion, pauses, fillers) delivered a speech, they were judged less favorably than their White counterparts.

In a test of whether the complexity of cognitive representations explained evaluative polarization, Linville (1982) asked college-aged men to read vignettes about either two college-aged men or two older men. For each pair of targets, person-based information about one man was positive and the other was negative. This within-subjects design allowed for a calculation of the magnitude of evaluative polarization. The results showed that the difference in evaluations between

the positive and negative targets was larger for the outgroup (older men) than for the ingroup (younger men) and that participants' level of cognitive complexity for the outgroup was negatively correlated with evaluative polarization, suggesting that limited complexity explains judgmental extremity.[2]

Although Linville (1982) initially predicted that people's cognitive representations for ingroups should be more complex than those for outgroups, Linville, Fischer, and Salovey (1989) found that male and female college students had similarly complex cognitive representations for their gender ingroup and outgroup (also see Linville, Fischer, & Yoon, 1996). In a similar vein, Fiske (1993) suggested that, compared with members of high status groups, members of low status groups have relatively complex cognitive representations for their respective high status outgroups. Based on Linville et al.'s and Fiske's perspectives, our meta-analysis takes into account (a) whether outgroups were defined by gender or by other group categories for which evaluators are known to have low cognitive complexity (e.g., race, nationality, age; Linville et al., 1989) and (b) whether the perceivers were members of high status groups (i.e., expected lower cognitive complexity) or low status groups (i.e., expected higher cognitive complexity).

To test complexity-extremity theory meta-analytically, we calculated an effect size that compared the evaluation of an ingroup target with that of an outgroup target, while keeping the valence of person-based information (i.e., positive, negative, or neutral) constant. We analyzed the effect sizes using target valence, type of group perceiver (i.e., expected low or high complexity for outgroup), and status of perceiver (i.e., high or low status) as moderators. Consistent with Linville (1982), we expected evaluators' judgments of outgroup targets to be more extreme in the direction of the valence of the information, but only when evaluators were expected to have limited cognitive complexity for their outgroups (Linville et al., 1989).

Black Sheep Effect and the Theory of Subjective Group Dynamics

In contrast to complexity-extremity theory, the black sheep effect predicts extremity in the evaluation of *ingroup* targets relative to outgroup targets and emphasizes group-protective motivation as the mechanism (e.g., Biernat, Vescio, & Billings, 1999; Marques et al., 1992; Marques & Yzerbyt, 1988; Marques et al., 1988). That is, the black sheep effect is a pattern of judgment in which positively depicted ingroup targets are judged more favorably than similar outgroup targets, but negatively depicted ingroup targets (i.e., "black sheep") are judged less favorably than similar outgroup targets.

The explanation for this evaluative pattern is predicated upon social identity theory's (Tajfel, 1982; Tajfel & Turner, 1986; Turner, 1975) emphasis on ingroup favoritism as a means of preserving positive differentiation between one's ingroup relative to an outgroup. Unfavorable ingroup members threaten the ingroup's claim to be positive and distinct from the outgroup, and their devaluation is thought to be a form of symbolic exclusion that attempts to reestablish the positivity of the ingroup (Eidelman, Silvia, & Biernat, 2006; Marques & Paez, 1994). Favoritism toward positive ingroup relative to outgroup targets is a more direct form of ingroup favoritism. From the point of view of an evaluator, a positively valenced target behaves in ways that are consistent with ingroup standards, whereas a negatively valenced target violates ingroup standards (Marques et al., 1988). Providing early support for the black sheep effect, Marques et al. (1988) found that Belgian students judged "likeable" Belgian students more positively than "likeable" North African students, and the reverse was observed when the students were "unlikable."

This idea of upholding ingroup standards was extended and refined in the subjective group dynamics model (Abrams, Marques, Brown, & Henson, 2000; Marques et al., 1998), which proposes that perceivers are motivated to differentiate the ingroup from the outgroup (i.e., category differentiation) as well as maintain support for ingroup norms (i.e., normative differentiation). Marques et al. (1998) manipulated participants' category membership by assigning them to a group, ostensibly based on a ranking task. Participants were then asked to evaluate ingroup targets that violated or did not violate ingroup norms. Consistent with the subjective group dynamics model, norm-consistent ingroup targets were evaluated more favorably than norm-violating ingroup targets. These findings suggest that positive social identity is particularly threatened when an ingroup member undermines group-based norms, resulting in derogation of this deviant. Ingroup members who undermine "generic" norms (norms applying to ingroups and outgroups) are also derogated. In this case, deviation from a norm undermines the ingroup's unique claim to do what is right.

To test predictions of this perspective, we used the ingroup–outgroup effect sizes, described previously (ingroup–outgroup evaluative differences at each level of person-based information). The analyses included studies in which person-based information could be coded as negative, positive, or neutral, and orthogonally coded as norm-consistent or norm-violating, thereby providing the opportunity to separate the effects of valence and norm consistency of person-based information. Based on the subjective group dynamics model, we expected ingroup targets to be evaluated more negatively than outgroup targets when person-based information was negative *and* violated ingroup norms. Positive norm violations are also deviant, but support the ingroup by underscoring its relative validity (Abrams et al., 2000). Extrapolating from this idea, we predicted that ingroup members who violated a norm in the positive direction would be evaluated more positively than the outgroup counterparts. In the absence of norm violation, we expected ingroup targets to be evaluated more favorably than outgroup

targets when person-based information was positive or neutral (a direct manifestation of ingroup favoritism) and less favorably when person-based information was negative (an indirect form of ingroup favoritism; that is, the black sheep effect). Based on the theory, when person-based information was neutral but otherwise violated norms, we expected that ingroup members would be evaluated more negatively than outgroup targets.

Expectancy-Violation Theory

Expectancy-violation theory (e.g., Bettencourt et al., 1997; Jussim et al., 1987) does not emphasize the evaluator's group membership vis-à-vis the target's group membership. Instead, it predicts that extremity in target evaluations is influenced by the interaction between preexisting stereotyped expectations about a target's group and personal information about the target. The theory suggests that individuals will be evaluated more extremely when their information violates stereotyped expectations for their salient groups (e.g., Bettencourt et al., 1997; Jussim et al., 1987; Jussim, Fleming, Coleman, & Kohberger, 1996; Kernahan, Bartholow, & Bettencourt, 2000) and that the valence of the information determines the direction of the evaluative extremity. This pattern of outcomes is expected, regardless of whether or not the evaluator shares group membership with the target. These predictions are based on the assumption that stereotype violations are surprising (see Olson, Roese, & Zanna, 1996), garnering more attention from the evaluator, and that the affect associated with the information valence drives evaluative extremity.

Supporting expectancy-violation theory, Jussim et al. (1987) found that White participants evaluated Black job applicants who appeared skillful (in the manner of speech and dress) more favorably than otherwise identical White job applicants, whereas participants evaluated White job applicants who appeared unskillful more negatively than otherwise identical Black job applicants. Using an identical study design but including Black participants, Coleman, Jussim, and Kelley (1995) found similar patterns of evaluative extremity toward expectancy-violating Black targets and White targets, suggesting that the group membership of the evaluator does not modify the effects.

Tests of this theory include targets from high and low status groups and vary the valence of person-based information in ways that violate stereotypes for one of these groups. To examine the predictions of expectancy-violation theory, we calculated an effect size that compared evaluations of high status and low status targets, while maintaining person-based information constant. We analyzed these effect sizes using target valence and stereotype incongruency (incongruent or congruent) as moderators. Following expectancy-violation theory, we expected more extreme positive evaluations of targets whose positive person-based information violated group-based stereotypes, relative to positive targets whose information did not violate group-based stereotypes. Also, we expected more extreme negative evaluations of targets whose negative person-based information violated group-based stereotypes, relative to negative targets whose information did not.

The Current Meta-Analysis

Our meta-analysis incorporates the broader impression formation literature, with the goal of testing theoretical predictions about when and how group-based and person-based information interact to produce evaluative extremity. To test the relevant three theoretical frameworks, we calculated two types of primary effect sizes; each used the social category of the target as a basis of the effect size, while holding person-based information constant. One compared evaluations of ingroup targets with evaluations of outgroup targets (i.e., ingroup–outgroup effect size), and the other compared evaluations of high status targets with evaluations of low status targets (i.e., high status–low status effect size). Also, we coded studies for theoretically relevant moderators: the valence of person-based information, the status of the evaluators' ingroup (high or low status), the norm consistency of the person-based information, and stereotype incongruence of the person-based information. We examined whether the meta-analytic data supported each theoretical model by analyzing the relevant type of effect size and the theoretically specified moderators. Our intent was not to pronounce which theory would be revealed as the best; rather, our goal was to focus on the broader literature to examine the conditions under which the predictions derived from each would be supported.

Method

Sample of Studies

We sought studies in which participants evaluated individual targets whose group-based and/or person-based information was manipulated. We used several strategies to locate relevant studies published prior to and including the year 2009. First, computer literature searches were conducted using the PsycINFO database. In addition to key terminology from each theory (e.g., "complexity-extremity," "black sheep," "subjective group dynamics," "expectancy violation"), the following keywords, in various combinations, were used for all searches: *person perception, person impression, social perception, impression formation, race, race and ethnic discrimination, racial and ethnic attitudes, racial and ethnic differences, sex-role attitudes, gender, status, stigma, stereotyped attitudes, stereotyped behavior, stereotypes, ingroup–outgroup, and evaluation*. This search yielded more than 9,000 abstracts.

Two of the authors read each abstract to determine whether the study was eligible for inclusion. When uncertain, a third

author independently examined the abstract. If still uncertain, the report was obtained and a final determination was made. Also, the reference sections of relevant reviews of the impression formation literature were scanned to identify additional eligible reports, as were the reference sections of all articles in the present meta-analysis. Finally, we contacted all researchers who had at least two reports in the meta-analysis to request relevant published and unpublished reports. Through these procedures, 755 reports were identified as likely eligible, and each was thoroughly examined to determine whether it met our full inclusion criteria. In total, 167 eligible reports met the criteria, yielding 201 separate studies and 1,153 effect sizes. The number of studies and effect sizes included in the present meta-analysis exceeds the average number of studies (M = 72, median = 64, range = 24-167) and the average number of effect sizes (M = 188, median = 124, range = 50-476) included in meta-analyses published in *Personality and Social Psychology Review* during the years 2010 to 2013.

We used the random effects method because it assumes that the results are reflective of the "random sample of the relevant distribution of effects" (i.e., available in studies) but not reflective of the entire population of possible studies (Borenstein, Hedges, & Rothstein, 2007, p. 86). Although the number of studies included in the present meta-analysis is quite large, we make no claim that the results go beyond the assumptions of the random effects model. This method is preferred over the fixed effect method and typically provides a more conservative estimate of significance (Borenstein et al., 2007).

Inclusion Criteria

Studies were included if the target of evaluation was an individual (evaluations of groups as a whole were not included). Information provided about the target had to include a cue that made salient both a social category and person-based information. We included a wide array of social categories, including race, ethnicity, gender, sexual orientation, socioeconomic status, disability status, occupation, university affiliation, political party, college major, and experimentally created groups. The types of person-based information were equally varied and inclusive, and included resumes, job performance, school applications, court records, and trait listings. It should be noted that, in some cases, seemingly group-based information could be considered person-based information. A key determinate of when a feature will represent person-versus group-based information is context (e.g., McGuire, McGuire, Child, & Fujioka, 1978; Turner, Oakes, Haslam, & McGarty, 1994). For the purposes of the present meta-analysis, if additional information was presented to participants that could be defined as person-based information, such as college performance, then a category that could be relevant to the current meta-analysis (ingroup vs. outgroup or high status vs. low status) was considered the group category and the additional information (poor performing student) was considered person-based (i.e., the categories were determined by the context). It should be noted, too, that some studies included two category distinctions, such that targets could be described as, for example, Black and male, White and male, Black and female, or White and female. In this case, two separate effect sizes were calculated, one that compared Black versus White targets and a second that compared female versus male targets. Of course, calculating more than one effect size from a single study creates problems with dependency among effect sizes. As will be described in the analyses section, we used hierarchical linear modeling (HLM) to account for within-study dependence.

All reports included in the meta-analysis were in English, but there were no restrictions with respect to the country in which the study was conducted. All studies were experiments in the sense that they either manipulated both group-based and person-based information or manipulated only one type of information, while maintaining the other constant. Both field and laboratory studies were included. Participants included children of at least grade/primary school age, high school students, college or university students, and adults from the community at large.

Because our focus was on person evaluation, we limited effect sizes to those that were measures of social attraction, likability, affective evaluations, global evaluations, composite or single-item trait evaluations (i.e., trustworthiness, aggressiveness, competence, etc.), or composites of traits. We excluded idiosyncratic measures that could not be categorized within these types of evaluative measures (e.g., clinical psychologists evaluating the mental health of a hypothetical target).

Any single study could contribute more than one evaluative measure. As recommended by Cooper (1989), if a study reported results for separate samples of participants, the samples contributed separate effect sizes. A very small percentage (less than 3%) of studies included additional conditions, such as manipulations of mood, accuracy, power, and interdependence—all of which have been shown to affect judgmental bias in the impression formation literature (e.g., Biernat et al., 1999; R. W. Doherty, 1998; Rudman & Fairchild, 2004). This small number prevented us from examining whether they moderated the effects of the theoretically relevant moderators. For studies that included these variables, we retained the conditions that did not use these inductions. One experimental condition known to influence evaluative outcomes was retained because of its importance for testing the subjective group dynamics model. Specifically, ingroup norm violation conditions are designed to moderate the influence of the effect of group membership on target judgments. For this condition (i.e., ingroup norm violation condition), effect sizes were calculated, but they were introduced into the meta-analysis only when the subjective group dynamics theory was tested.

Variables Retrieved and Coded

All data were independently retrieved by trained coders. Methods for coding attributes of studies and extracting effect sizes were formulated through discussion among the authors. A subset of the eligible reports (25%) were then coded, during which issues were discussed thoroughly and further adjustments were made to the coding strategy as part of the coder training process. There was a strong agreement among the coders (average %agreement = 95%; range = 82%-100%). Following this training process, coders continued to meet regularly with the first author to discuss any questions and resolve discrepancies.

Attributes of group- and person-based information were coded for each effect size. For group-based information, we recorded the social groups compared (i.e., Black and White, female and male, etc.), the social category of the evaluator (i.e., Black, male, etc.), and the status of the targets' ingroup (i.e., high status, low status, no status difference; for example, if the two groups examined in a study were Blacks and Whites and the context was the United States, Black targets' ingroup was the low status group and White targets' ingroup was the high status group). We also coded the status of the evaluators' ingroup (i.e., low status, high status, no status difference).

For person-based information, we coded the valence of person-based information (i.e., positive, negative, unvalenced/neutral; for example, if person-based information indicated that a job applicant was skillful, it was coded as positive; an unskilled job applicant was coded as negative). The unvalenced category subsumed person-based information that was neither clearly positive nor clearly negative, including person-based information that was unvalenced or neutral (e.g., a student liked or disliked science), ambiguous (e.g., advice about seeking learning assistance that could be either construed as helpful or critical), and mixed (e.g., job applicants with above average skills and no work experience). We also coded the incongruence of person-based information (i.e., incongruent with the low status group, incongruent with high status group, incongruent with neither). Finally, we coded whether the study that provided the effect size specifically included a norm violation manipulation, and separately, whether the targets' person-based information violated a group norm (i.e., violation of general norm, violation of ingroup norm, not norm violating).[3]

Effect Size Calculation

Relevant to complexity-extremity and subjective group dynamics perspectives, we calculated an effect size that compared an evaluation of an ingroup target with that of an outgroup target. It was calculated using a mean evaluation of ingroup target minus a mean evaluation of an outgroup target while holding person-based information constant (i.e., *ingroup–outgroup* effect size). Relevant to expectancy-violation theory, we calculated an effect size that compared an evaluation of a high status target with that of a low status target: the mean evaluation of a high status target minus a mean evaluation of a low status target, while holding person-based information constant (i.e., *high status–low status* effect size).

The effect-size metric used in this report is the *d* index (Hedges, 1981; Hedges & Olkin, 1985). The *d* index is the difference between two group means, standardized by the pooled standard deviation, and it corrects for the fact that the *g* statistic (J. Cohen, 1988) overestimates the population effect size, especially for small sample sizes (Hedges & Olkin, 1985). Eligible studies contained sufficient statistical information (e.g., cell means and standard deviations, *F* ratios, *t* statistics) to calculate an effect size. The relevant means for the dependent variables of interest from each study were retrieved and the respective standard deviations were recorded. If standard deviations were not available, a pooled error term was estimated from inferential statistics (Cooper, 1998; R. Rosenthal, 1991). When the dependent variables were frequencies or proportions, appropriate formulas were used to calculate the effect size from resulting contingency tables. When means were drawn from a within-subjects design (e.g., the same participant made judgments about intelligence for a male and a female), a between-condition correlation for the dependent variable of .5 was introduced for the calculation of *d* and its variance, v_d (see Morris & DeShon, 2002; also Lipsey & Wilson, 2001, for general discussion on calculation of effect sizes). The correlation of .5 was chosen based on research that allowed an examination of the likely correlation between dependent variables from within-subjects designs (e.g., Eidelman et al., 2006), and is consistent with the recommendations of Morris and DeShon (2002).

Analyses

Utilizing HLM 6.0 (Raudenbush, Bryk, & Congdon, 2000), multilevel models (MLMs) were used to examine the meta-analytic hypotheses (Kalaian & Raudenbush, 1996). While relatively few meta-analyses have employed this technique (e.g., Dickerson & Kemeny, 2004; Lopez-Duran, Kovacs, & George, 2009; Nowak & Heinrichs, 2008; Schneider et al., 2010), mixed-effects models offer several advantages. First, the model incorporates both fixed effects coefficients that can be used to estimate effect sizes given the effects of included moderators, and random effects coefficients, with associated significance tests, thus allowing the results to be generalized to a population of studies in contrast to being restricted to the studies included in our meta-analysis. Second, the model accounts for the hierarchical structure of our data described in a subsequent paragraph, thus allowing each study to contribute multiple effect sizes while accounting for the dependencies in the data. Third, the model allows for inclusion of moderators at the appropriate "level" of analysis. In calculating the population effect sizes, each individual effect size was weighted by the inverse of its variance.

Because the presence of extreme values in a dataset seriously distorts the outcomes of analyses (Tabachnick & Fidell, 2007; Wilcox, 1995), including meta-analyses (see Bettencourt & Miller, 1996, for discussion), the distributions of the high status–low status and ingroup–outgroup effect sizes were examined separately and separately for between-subjects and within-subjects designs. Effect size outliers were identified following the technique recommended by Hedges and Olkin (1985). Outlying effect sizes were defined as those with standardized residuals that were 3.5 *SD* above or below a mean effect size, with that effect size being excluded.[4] Given that the analyses included a relatively large number of effect sizes, we chose 3.5 *SD* (instead of the more typically adopted 3.0) in an effort to be more conservative in the identification of outliers. Among ingroup–outgroup effect sizes used for the test of complexity-extremity hypotheses, 2.2% were identified as outliers, and 1.2% of the positive–negative effect sizes were identified as outliers. Among ingroup–outgroup effect sizes used for tests of subjective group dynamics, 2.4% were identified as outliers. Finally, 2.3% of the high status–low status effect sizes for tests of expectancy violation were identified as outliers. Outlying values in a distribution may be either excluded or modified (Tabachnick & Fidell, 2007). To retain all of the effect sizes in the analyses, we modified the outlying effect sizes to the value of the next nonoutlying effect size, plus or minus 0.01 depending on whether the effect size was positive or negative, separately for between-subjects and within-subjects studies for each type of effect size.

Due to the complex nature of our dataset, three-level models using full maximum likelihood as the method of estimation were used to simultaneously account for (Level 1) sample variance associated with each effect size, (Level 2) variance within studies due to the type of measure and type of evaluative target, and (Level 3) variance between studies.[5] Sample variance for each individual effect size was a known value (v_d), the inverse of which was used to weight the effect sizes at Level 1 of the model.

Given that all moderators in our analyses were categorical (e.g., person-based information was either positive, negative, or neutral), the HLM model utilized contrast codes to examine hypotheses associated with each theory and to estimate cell-specific mean effect sizes. Specifically, (a) we defined contrasts to examine theoretically meaningful main effects and interactions of the moderators on the population effect sizes and (b) we used the fixed effects coefficients associated with the contrasts in each model to calculate a mean effect size for each "cell" defined by the categorical variables in the analysis. We used general linear hypothesis tests and resulting Wald's test statistics within HLM to examine whether the cell-specific mean effect sizes were significantly greater than 0. Given the unbalanced cells associated with the assignment of contrast coefficients, the resulting model intercept is interpreted as the unweighted sample mean (i.e., the mean of all the cell means) in contrast to the grand mean (the mean of all the effect sizes).

As previously stated, we included three levels in the HLM model—mean effect sizes (Level 1) are nested within types of measures and types of comparisons (e.g., norm consistent vs. norm violating, Level 2), which are nested within the studies (Level 3). The intercept in each model (π_0) is interpreted as the population effect size. Level 1 of the model takes the following general form:

$$d_{\mathrm{ijk}} = \pi_0 + \pi_1 \mathrm{A} + \pi_2 \mathrm{B} + \pi_3 \mathrm{A} \times \mathrm{B} + \mathrm{e}_{ijk},$$

where π_1 and π_2 are differences in mean effect sizes defined by the contrasted levels of the relevant moderators, and π_3 indicates whether there is a significant interaction between the moderators.

For tests of the complexity-extremity hypothesis, we included cross-level interactions with a contrast-coded Level 2 variable indicating whether the effect size was for a comparison between males and females or other groups. Thus, Level 2 of the model took the following form:

$$\pi_0 = \beta_{00} + \beta_{01}\mathrm{L2Variable} + r_{0jk},$$

$$\pi_1 = \beta_{10} + \beta_{11}\mathrm{L2Variable},$$

$$\pi_2 = \beta_{20} + \beta_{21}\mathrm{L2Variable},$$

$$\pi_3 = \beta_{30} + \beta_{31}\mathrm{L2Variable}.$$

$\beta_{.0}$ (where "." = 0, 1, 2, or 3) is the mean value for each Level 1 coefficient, and β_1 indicates whether there is a significant effect of the Level 2 variable, as defined by the contrast codes, on the respective Level 1 coefficient. The inclusion of the Level 2 random effect, r_{0jk}, on the Level 1 intercept indicates that there may be additional within-study variance in the mean effect size that is not accounted for by the Level 2 predictor in the model.

Level 3 of the model represents the between-study level. While there were no between-study moderators included in our substantive analyses, random effects are included on the intercept. Note that the complex error term that represents the sources of unexplained variance in effect sizes in the current model ($\mathrm{e}_{ijk} + r_{0jk} + U_{00k}$) may be considered an expansion of the error term in a random effects model (sample variance + between-studies variance).

Results

Complexity-Extremity Theory

To test complexity-extremity theory, we used the effect sizes that subtracted an evaluation of an outgroup target from that of an ingroup target (ingroup–outgroup effect sizes).

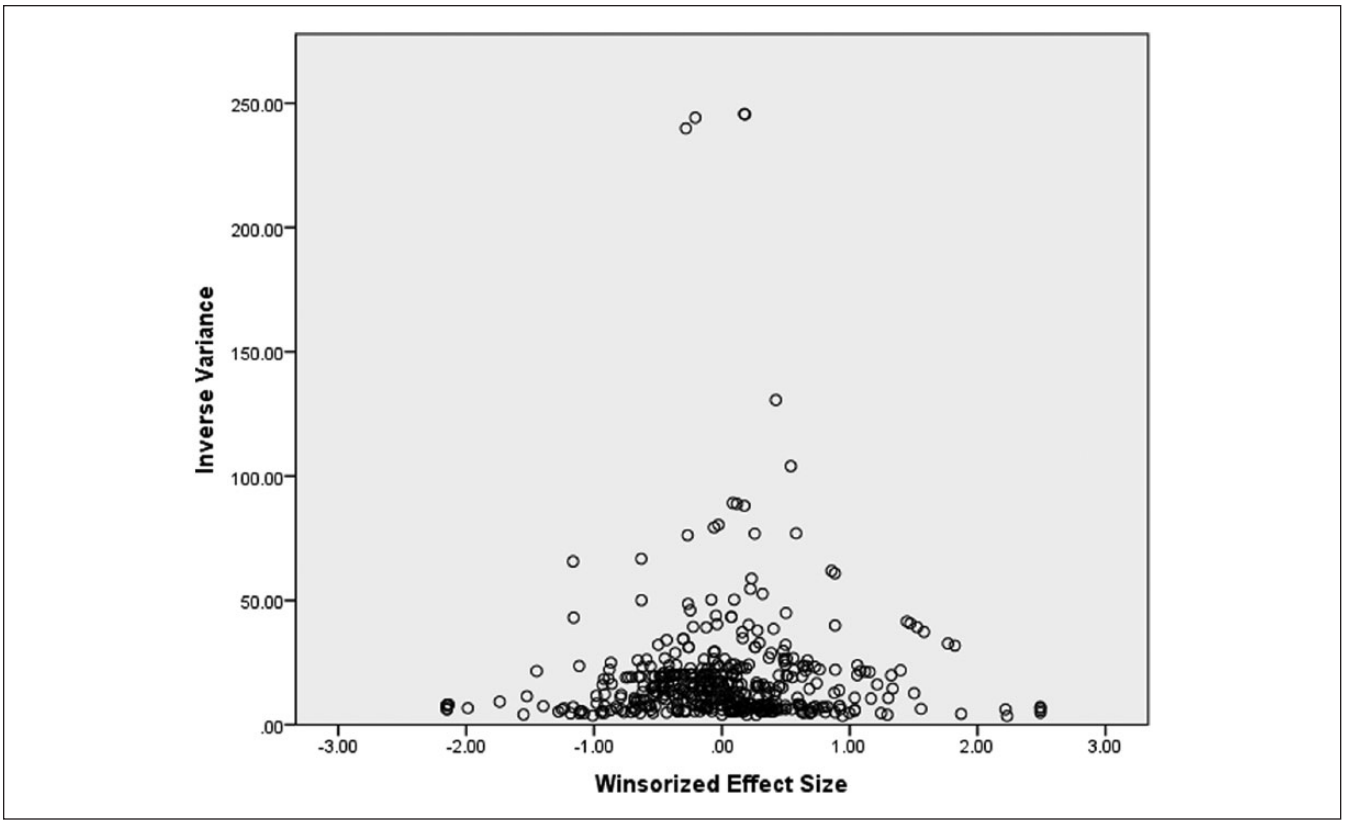

Figure 1. Funnel plots showing the ingroup–outgroup effect sizes as a function of the inverse of the variance for the effect sizes used to test complexity-extremity theory.

To examine the possibility of publication bias, we plotted the effect sizes against the inverse of the effect size variance to examine their distribution. As shown in Figure 1, the funnel plot indicates that effect sizes associated with larger inverse variances (i.e., indicators of sample size) tend to cluster around the center of the distribution. The symmetrical shape of the plot suggests little cause for concern regarding publication bias.

Based on the complexity-extremity theory, we examined the moderating influences of the evaluator's group membership (high vs. low status or male vs. female) and the valence of person-based information (positive, negative, neutral) on these ingroup–outgroup effect sizes. To test the hypotheses, 389 effect sizes from 98 studies were included in a three-level MLM.

Following complexity-extremity theory, it was predicted that high status perceivers with limited cognitive complexity (i.e., high status groups not defined by gender) would evaluate outgroup targets more favorably than ingroup targets when information was positive but less favorably when information was negative. That is, it was expected that these evaluations of outgroup targets would be more extreme, in the direction of information valence. We did not expect this pattern of outgroup polarization among low status evaluators, because theoretically (Fiske, 1993), they should have relatively complex cognitive representations of their high status outgroups.

We specified a contrast to examine whether there was a significant difference between effect sizes for positively (.5) and negatively (−.5) valenced individuating information (unvalenced information was coded as 0). We specified a separate contrast to examine whether the effect size for unvalenced individuating information (.67) was significantly different from positively and negatively valenced individuating information (−.33). We also included contrasts to examine whether the effect sizes were significantly different when the perceiver was from the high status group (.5) versus the low status group (−.5). Finally, to examine whether the effect of valence was influenced by group membership of the perceiver, we included interaction between the positive–negative contrasts and the high versus low status contrasts, as well as interactions between the valenced versus unvalenced and the high versus low contrasts. We included the type of groups used as stimuli as a within-study (Level 2) moderator of the effect size moderator (i.e., the Level 1 effects).

The MLM results are presented in Table 1. The intercept indicated that the mean effect size across all conditions (d_+ = 0.04) was not significantly different from 0, $t(97) = 0.70$, *ns*. There was a significant difference between positively and negatively valenced effect sizes, $\gamma_{100} = .18$, $t(151) = 2.99$, $p <$

Table 1. Fixed and Random Effects for Complexity-Extremity Theory.

Fixed effects	γ	*SE*
Intercept	.04	0.05
M–F vs. Other groups	−.07	0.09
Pos vs. Neg	.18*	0.06
Pos vs. Neg × M–F vs. Other groups	−.20	0.12
Valenced vs. Neutral	−.01	0.08
Valenced vs. Neutral × M–F vs. Other groups	−.11	0.16
High vs. Low status	−.15*	0.05
High vs. Low status × M–F vs. Other groups	.10	0.11
Pos vs. Neg × High vs. Low status	−.59*	0.12
Pos vs. Neg × High vs. Low status × Other groups	1.00*	0.23
Valenced vs. Neutral × High vs. Low status	.20	0.10
Valenced vs. Neutral × High vs. Low status × Other groups	−.46*	0.20
Random effects	**Variance**	
Level 2		
Intercept	0.09*	
Level 3		
Intercept	0.06*	

Note. M–F = males and females; other groups = other high and low status groups; pos = positively valenced; neg = negatively valenced; high = high status groups; low = low status groups.
*$p < .05$.

.01, and a significant difference between high status and low status effect sizes, $\gamma_{300} = -.15$, $t(151) = -2.76$, $p < .01$. A significant coefficient for the relevant interaction term, $\gamma_{400} = -.59$, $t(151) = -5.01$, $p < .001$, indicated that the differences between positively and negatively valenced effect sizes were qualified by whether the perceiver was from the high status or low status group. The interaction was further qualified by whether the study was comparing males and females or not, $\gamma_{410} = 1.00$, $t(151) = 4.27$, $p < .001$. Finally, there was a marginally significant valenced versus unvalenced information by high status versus low status perceiver interaction, $\gamma_{500} = .20$, $t(151) = 1.97$, $p = .05$, which was further qualified by whether the study was comparing males and females or not, $\gamma_{510} = -.46$, $t(151) = -2.28$, $p < .05$. After accounting for the male–female versus other group comparison, the random effects indicate that there was a significant within-study and between-study variance in the effect sizes. We used general linear hypothesis tests to determine whether the model implied mean effect sizes for each cell defined by the contrasts was significantly different from 0.

Figure 2 depicts the mean of the ingroup–outgroup effect sizes for targets with memberships in high status and low status groups, other than gender groups. As shown in the left side of Figure 1, the results revealed that high status perceivers evaluated outgroup targets more positively than identical ingroup targets, $d_+ = -0.24$, $\chi^2(1) = 13.33$, $p < .001$, when person-based information was positive. This evidence of outgroup extremity is consistent with the complexity-extremity hypothesis. When the information was negative, the results revealed an ingroup–outgroup effect size equivalent to 0, $d_+ = +0.02$, $\chi^2(1) = .05$, *ns*. Thus, evaluative extremity toward the outgroup target was not observed in this case.[6]

Evaluative extremity toward outgroups was *not* predicted for low status perceivers, because they were expected to have relatively complex schemas for their high status outgroups. As shown in the right side of Figure 1, the results revealed that, when person-based information was positive, low status perceivers evaluated ingroup targets more favorably than outgroup targets, $d_+ = +0.64$, $\chi^2(1) = 13.41$, $p < .01$. When the information was negative, low status perceivers evaluated ingroup and outgroup targets similarly, $d_+ = -0.19$, $\chi^2(1) = 1.06$, *ns*.

Consistent with complexity-extremity theory, the results did not reveal more extreme evaluations of outgroup targets. But complexity-extremity theory does not predict ingroup polarization in the case of positively valenced targets; this pattern is more consistent with the theory of subjective group dynamics.

As noted previously, some effect sizes were associated with person-based information that was neutral. Complexity-extremity theory does not make a prediction about targets depicted with neutral information, but extrapolating from studies revealing ingroup bias (e.g., Tajfel & Turner, 1986), evaluators might be expected to favor ingroup targets over outgroup targets. As shown in Figure 1, when individuating information was neutral, there was a trend toward ingroup favoritism among high status perceivers, $d_+ = +0.15$, $\chi^2(1) = 2.87$, $p < .10$, and although the mean effect size for low status perceivers was in a similar direction of favoring the high status group, it was equivalent to 0, $d_+ = +0.06$, $\chi^2(1) = .22$, *ns*.

Figure 3 prevents comparable mean effect sizes for gender groups only, the case in which Linville et al. (1989) have explicitly argued that there is likely to be little difference in the complexity of representations for ingroups and outgroups. Consistent with this premise, when both evaluators and targets were defined by gender, there was no evidence of outgroup polarization (i.e., all mean effect sizes equivalent to 0).

Subjective Group Dynamics Theory

As shown in Figure 4, the funnel plot for the ingroup–outgroup effect sizes used to test subjective group dynamics theory indicates that effect sizes associated with larger inverse variances tend to cluster around the center of the distribution, and the symmetrical shape of the plot suggests that publication bias is not a concern.

We examined the modifying influences of the valence of person-based information and whether the information was consistent with or violated group norms. Based on subjective group dynamics theory, we predicted that, when person-based

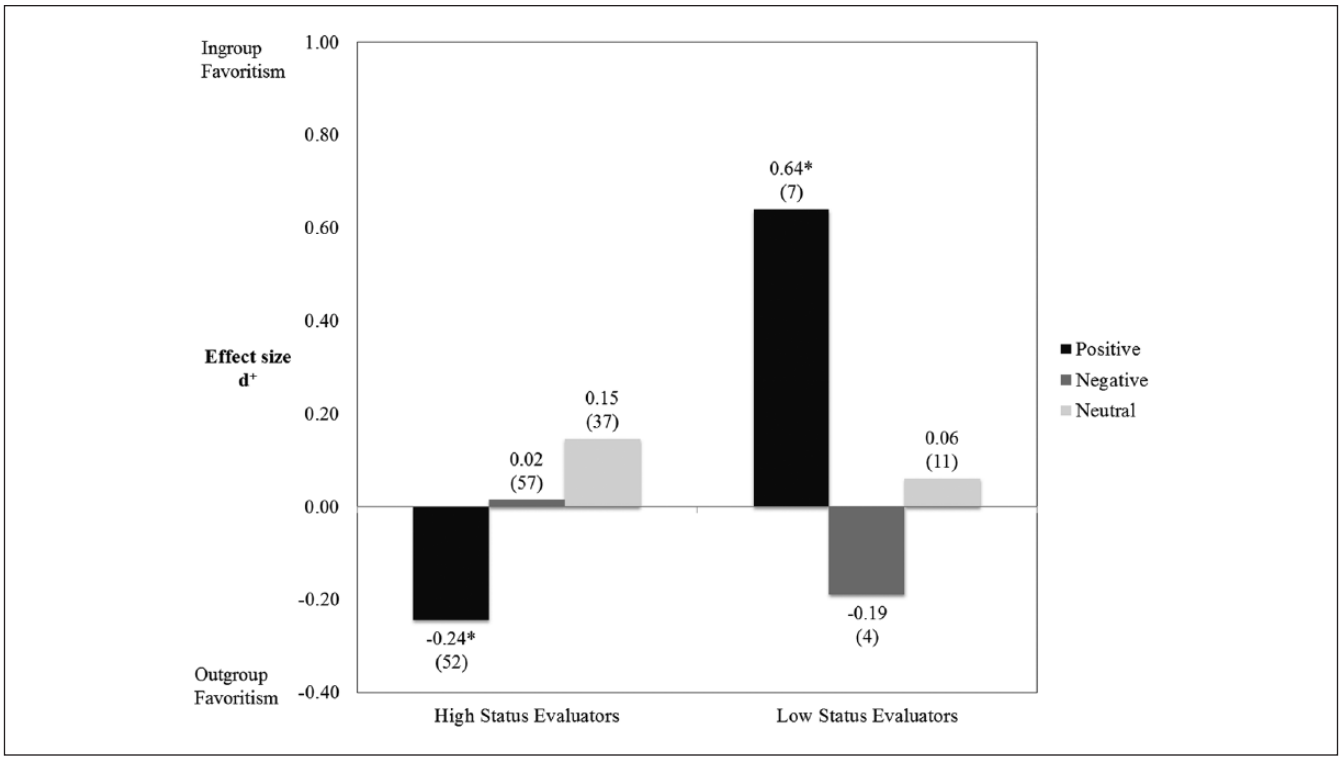

Figure 2. Mean ingroup–outgroup effect sizes moderated by perceivers' group status and valence of person-based information.
Note. A negative ingroup–outgroup effect size reveals more positive evaluations of outgroup targets, and a positive effect size reveals more positive evaluations of ingroup targets. Values in parentheses indicate the number of effect sizes, *k*.
*Significantly different from 0.

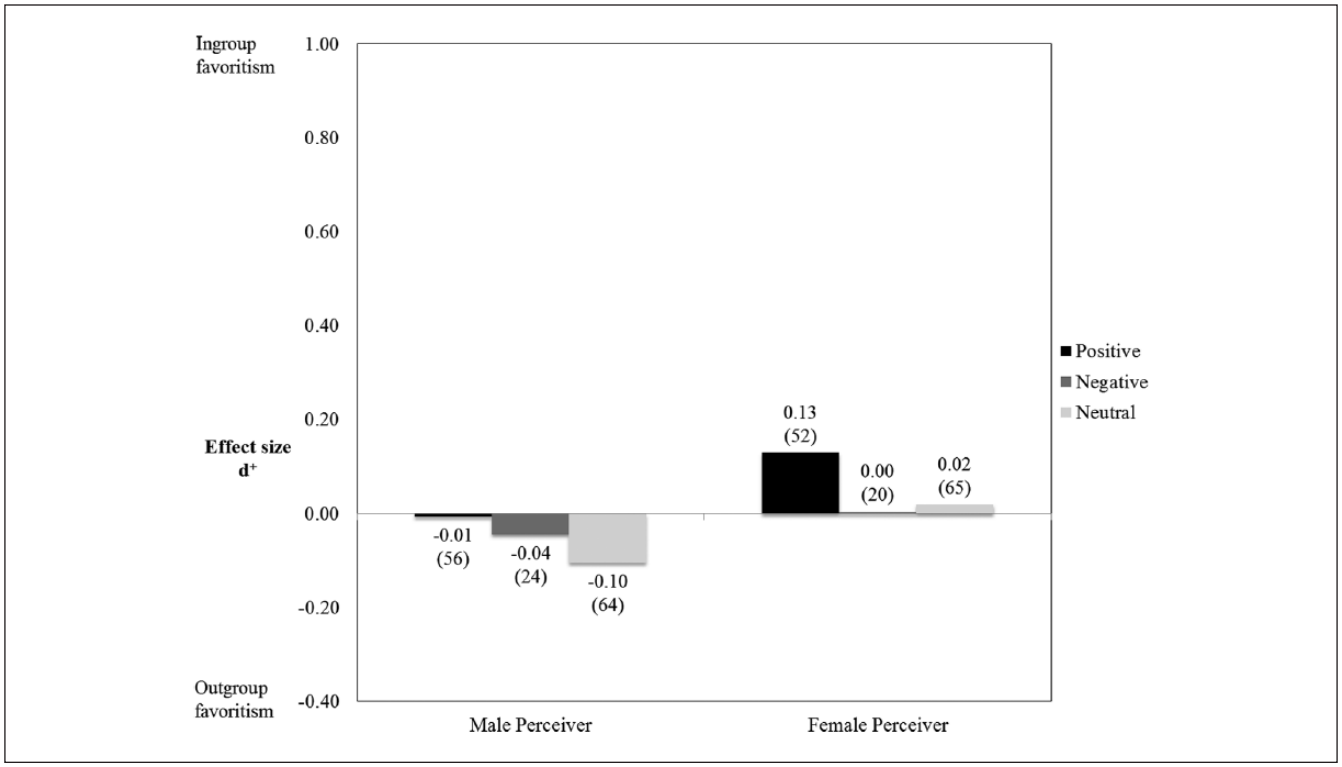

Figure 3. Mean ingroup–outgroup effect sizes moderated by perceivers' gender status and valence of person-based information.
Note. None of the mean effect sizes are significantly different from 0. Values in parentheses indicate the number of effect sizes, *k*.

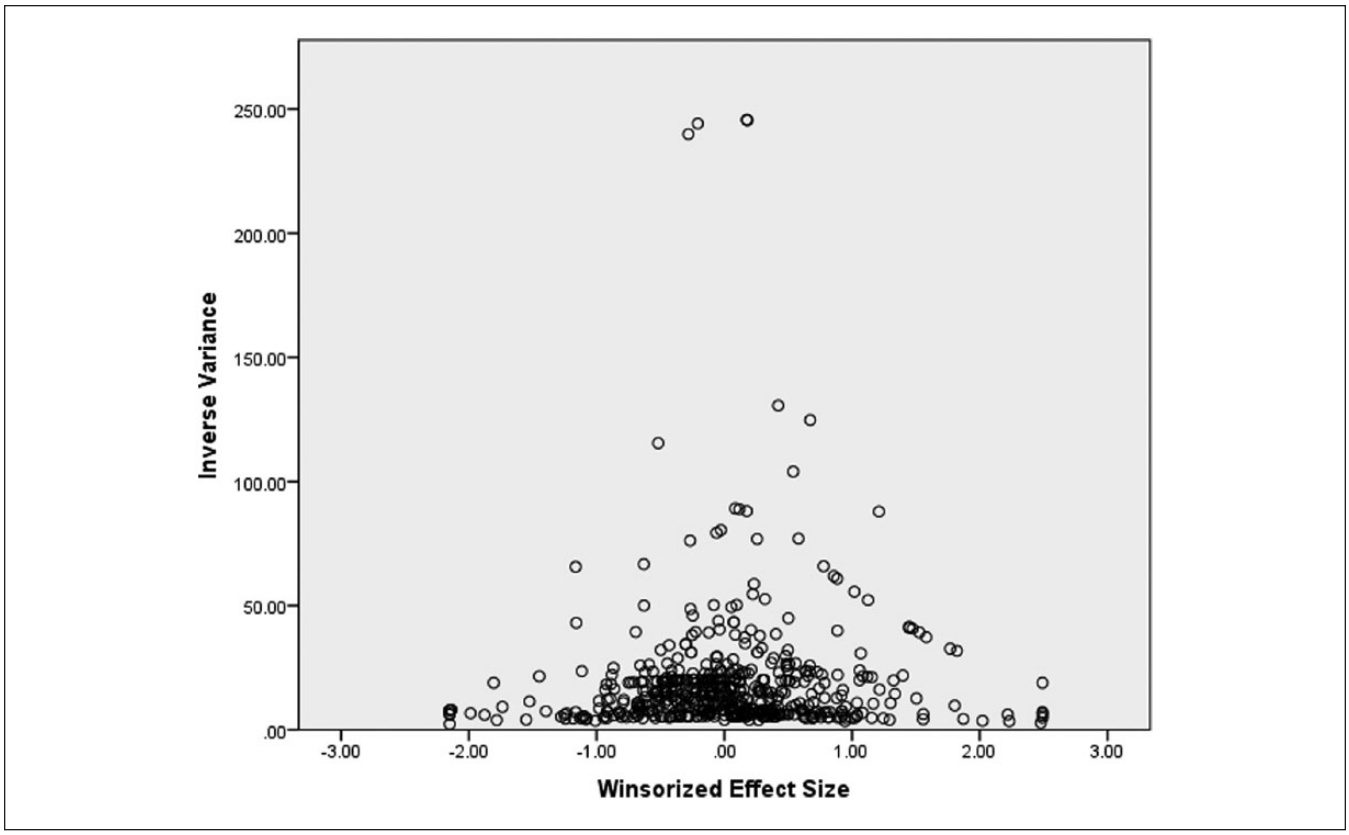

Figure 4. Funnel plots showing the ingroup–outgroup effect sizes as a function of the inverse of the variance for the effect sizes used to test subjective group dynamics theory.

information was negative and violated group norms, perceivers would evaluate ingroup targets more negatively than outgroup targets (e.g., black sheep effect). By contrast, we expected that when person-based information was positive and violated group norms, ingroup targets would be evaluated more positively than outgroup targets. In the absence of clearly valenced person-based information, we expected that ingroup targets that violated group norms would be evaluated more negatively than those who did not. To test the hypotheses, 536 effect sizes from 136 studies were included in the model.

To examine the effects of valence of individuating information, we specified Level 1 contrasts similar to the ones used for tests of the complexity-extremity theory (i.e., positive vs. negative and valenced vs. unvalenced contrast coefficients). Also, we included contrast coefficients to examine whether effect sizes that were used for information that violated group norm (0.5) were significantly different from those that were norm consistent (−0.5). Finally, we included an interaction between the positive versus negative information and the norm violation versus norm-consistent contrasts as well as an interaction between the valenced versus unvalenced information and the norm violation versus norm-consistent contrasts.

The MLM results are presented in Table 2. The mean ingroup–outgroup effect size across all conditions was d_{+} = −0.19, $t(135) = -3.71$, $p < .001$. This finding suggests that, overall, perceivers evaluated outgroup targets more favorably than ingroup targets, but the interactions between conditions reveal that it is most important to consider the mean effect sizes within the various theory-relevant conditions. The results also showed that there were significant main effects of positive versus negative individuating information, $\gamma_{100} = .29$, $t(220) = 3.21$, $p < .01$, and of norm-violating versus norm-consistent individuating information, $\gamma_{200} = -.47$, $t(220) = 5.16$, $p < .001$. More importantly, the interaction between the two was significant, $\gamma_{400} = .70$, $t(220) = 3.96$, $p < .001$. Finally, whereas there was no main effect of valenced (negative and positive) versus unvalenced individuating information, there was a significant interaction between valenced versus unvalenced information and norm-violating versus norm-consistent individuating information, $\gamma_{500} = -.67$, $t(220) = -4.04$, $p < .001$. The random effects indicate that there was a significant within-study and between-study variance in the effect sizes.

We used general linear hypothesis tests to estimate the model implied mean effect sizes for each cell. As shown in Figure 5, when person-based information was negative and

Table 2. Fixed and Random Effects for Subjective Group Dynamics Theory.

Fixed effect	γ	*SE*
Intercept	−.19*	0.05
Pos vs. Neg	.29*	0.09
Norm vio vs. Cons	−.47*	0.09
Val vs. Unv	−.09	0.09
Pos vs. Neg × Norm vio vs. Cons	.70*	0.18
Val vs. Unv × Norm vio vs. Cons	−.67*	0.17
Random effect	**Variance**	
Level 1 intercept	0.05*	
Level 2 intercept	0.11*	

Note. Pos = positive valence; neg = negative valence; norm vio = norm violation; norm cons = norm consistent; val = valenced; neut = neutral.
*$p < .05$.

violated ingroup norms, ingroup targets were evaluated less favorably than outgroup targets, $d_+ = -0.61$, $\chi^2(1) = 18.11$, $p < .01$. By comparison, when individuating information was positive, regardless of whether it violated or was consistent with group norms, perceivers evaluated ingroup and outgroup targets similarly, $d_+ = +0.004$, $\chi^2(1) = .007$, $d_+ = 0.028$, $\chi^2(1) = .03$, respectively, *ns.* The results, depicted on the far right side of Figure 5, show that when information was relatively neutral but nevertheless violated group norms, perceivers evaluated ingroup targets more negatively than outgroup targets (i.e., ingroup derogation), $d_+ = -0.72$, $\chi^2(1) = 48.08$, $p < .01$. By contrast, when information was neutral and was consistent with group norms, perceivers evaluated ingroup targets more favorably than outgroup targets, $d_+ = 0.20$, $\chi^2(1) = 17.94$, $p < .01$.

Expectancy-Violation Theory

As noted previously, expectancy-violation theory predicts effects of information valence and whether that information violates stereotyped expectations for high and low status groups. To test this theory, we used the high status–low status effect sizes, which represent the difference in the evaluations of a high status target and a low status target. As shown in Figure 6, the funnel plot for the distribution of the high status–low status effect sizes indicates that studies associated with larger inverse variances tend to cluster around the center of the distribution, and the symmetrical shape of the plot suggests little, if any, publication bias.

The relevant moderators were valence of the person-based information and whether this information was inconsistent with or consistent with stereotyped expectations for either the low status or high status group. Four hundred twenty-nine effect sizes were available from 109 studies to test these hypotheses.

Based on expectancy-violation theory, we hypothesized that targets would be evaluated more extremely when their information violates stereotyped expectations for their salient ingroups and that the information valence determines the direction of the evaluative extremity. That is, it was expected that, when the person-based information was positive and incongruent with the stereotyped expectations for the low status group, perceivers would evaluate low status targets more favorably than high status targets depicted with the same positive but congruent information. But when the positive information was incongruent with the stereotypes about the high status group, it was expected that high status targets would be evaluated more favorably than low status targets that had identically positive but congruent information. By contrast, it was expected that, when information was *negative* and incongruent with the stereotyped expectations for the low status group, low status targets would be evaluated more negatively than similar but congruent negative high status targets, but when negative information was incongruent with stereotypes about the high status group, high status targets would be evaluated more negatively than similar but congruent negative low status targets. Expectancy-violation theory is silent with regard to instances in which person-based information is relatively neutral. Nevertheless, we expected that perceivers would be less favorable toward targets that behaved in ways that violated stereotyped expectations (Figure 7).

To examine the effects of the valence of individuating information, we specified Level 1 contrasts similar to the ones used in the previous analyses. We also included contrast coefficients to examine whether effect sizes that were used for individuating information that was incongruent with the low status group (0.5) were significantly different from those that were incongruent with the high status group (−0.5). Finally, we included an interaction between the positive–negative contrasts and the incongruency contrasts as well as an interaction between the valenced versus unvalenced information and the incongruency contrasts. Results of the MLM are presented in Table 3.

The results indicated that the mean effect size (−0.07) across all conditions was significantly different from 0, $t(108) = -2.01$, $p < .05$. This small, but reliable, effect size suggests that, overall, perceivers evaluated low status targets more positively than high status targets, but the significant interactions among the theoretical moderators suggest that it is important to consider the ways in which they influence the magnitude of the effect sizes. The results also revealed a main effect of positive versus negative valenced individuating information, $\gamma_{100} = -.21$, $t(171) = -4.36$, $p < .001$, and a main effect of valenced (positive and negative) versus neutral individuating information, $\gamma_{200} = .16$, $t(171) = 2.32$, $p < .05$. Effect sizes associated with information that was incongruent with the low status group were significantly different from those associated with information incongruent with the high status group, $\gamma_{300} = .14$, $t(171) = 4.47$, $p < .001$. Finally, there was a significant interaction between valenced versus unvalenced and the incongruence of individuating information,

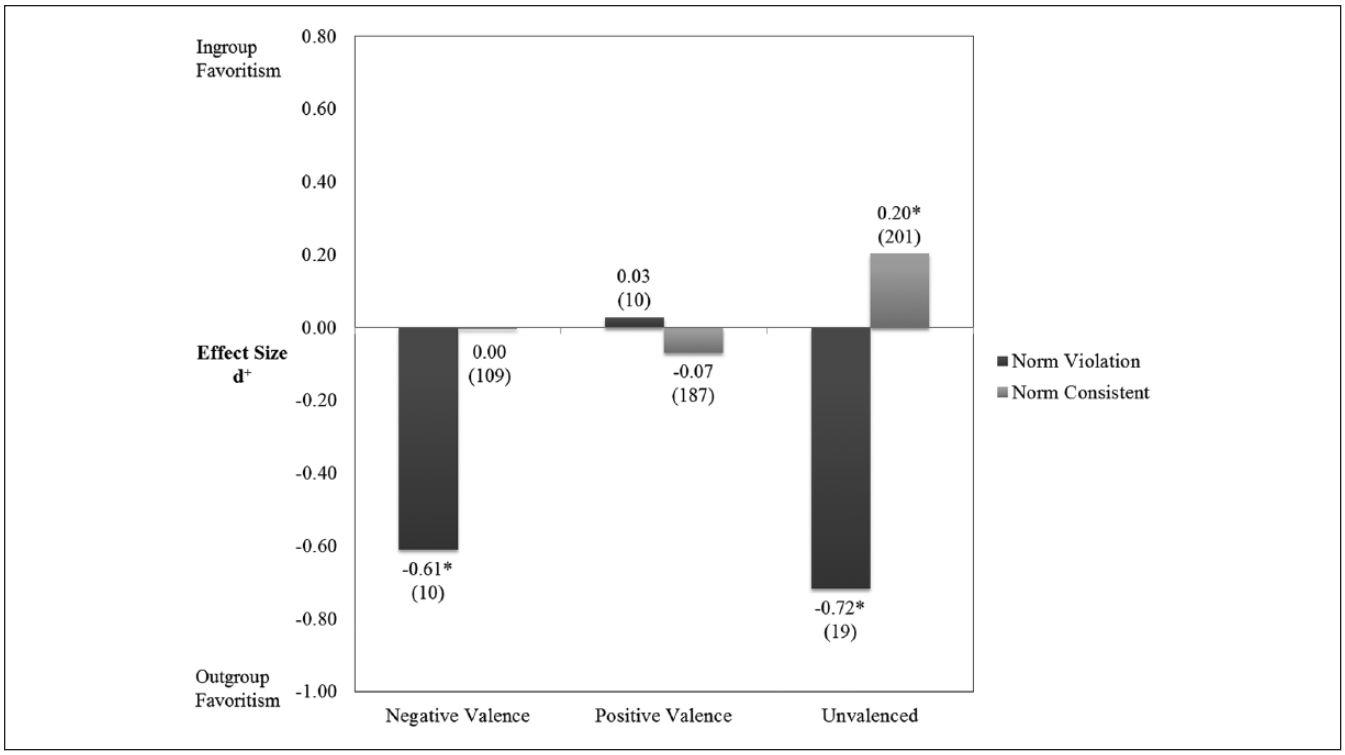

Figure 5. Mean ingroup–outgroup effect sizes moderated by valence of person-based information and whether the information violated or was consistent with group norms.
Note. A negative ingroup–outgroup effect size reveals more positive evaluations of outgroup targets, and a positive effect size reveals more positive evaluations of ingroup targets. Values in parentheses indicate the number of effect sizes, *k*.
*Significantly different from 0.

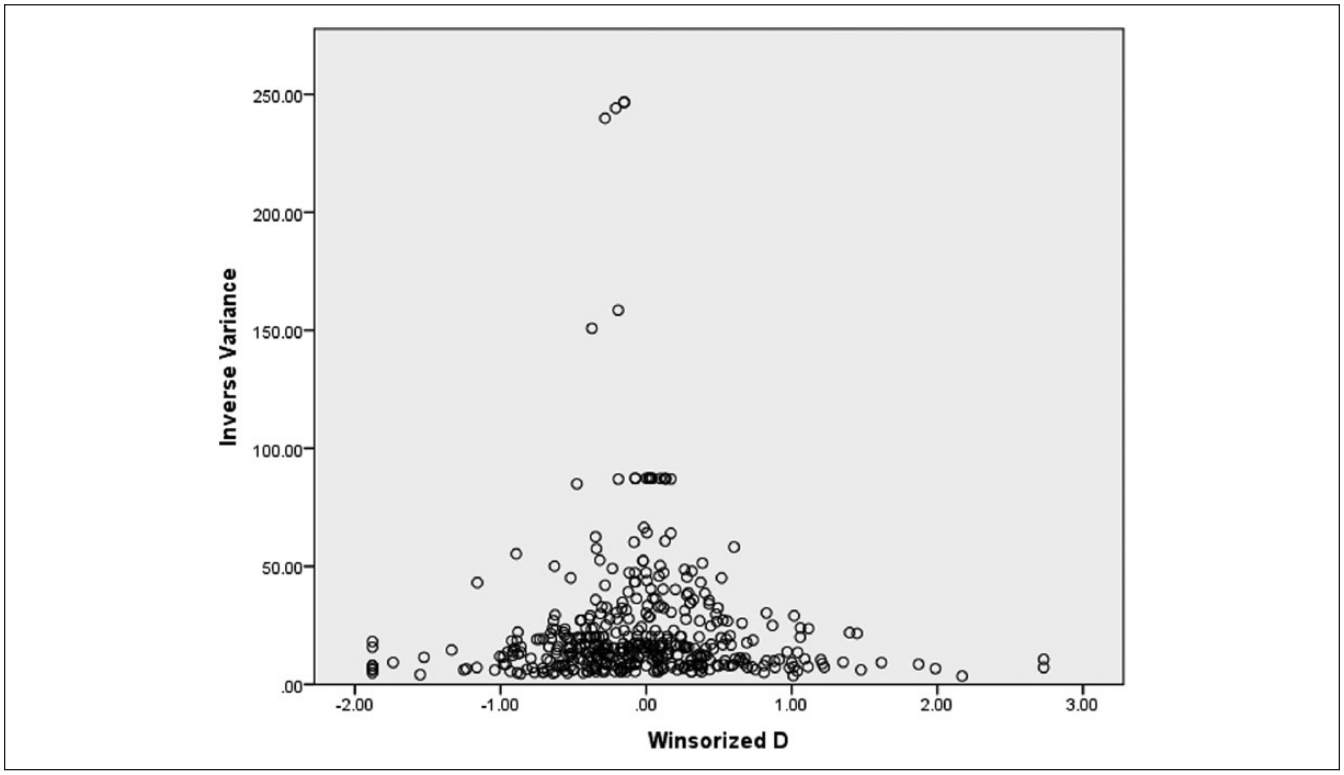

Figure 6. Funnel plots showing the high status–low status effect sizes as a function of the inverse of the variance for the effect sizes used to test expectancy-violation theory.

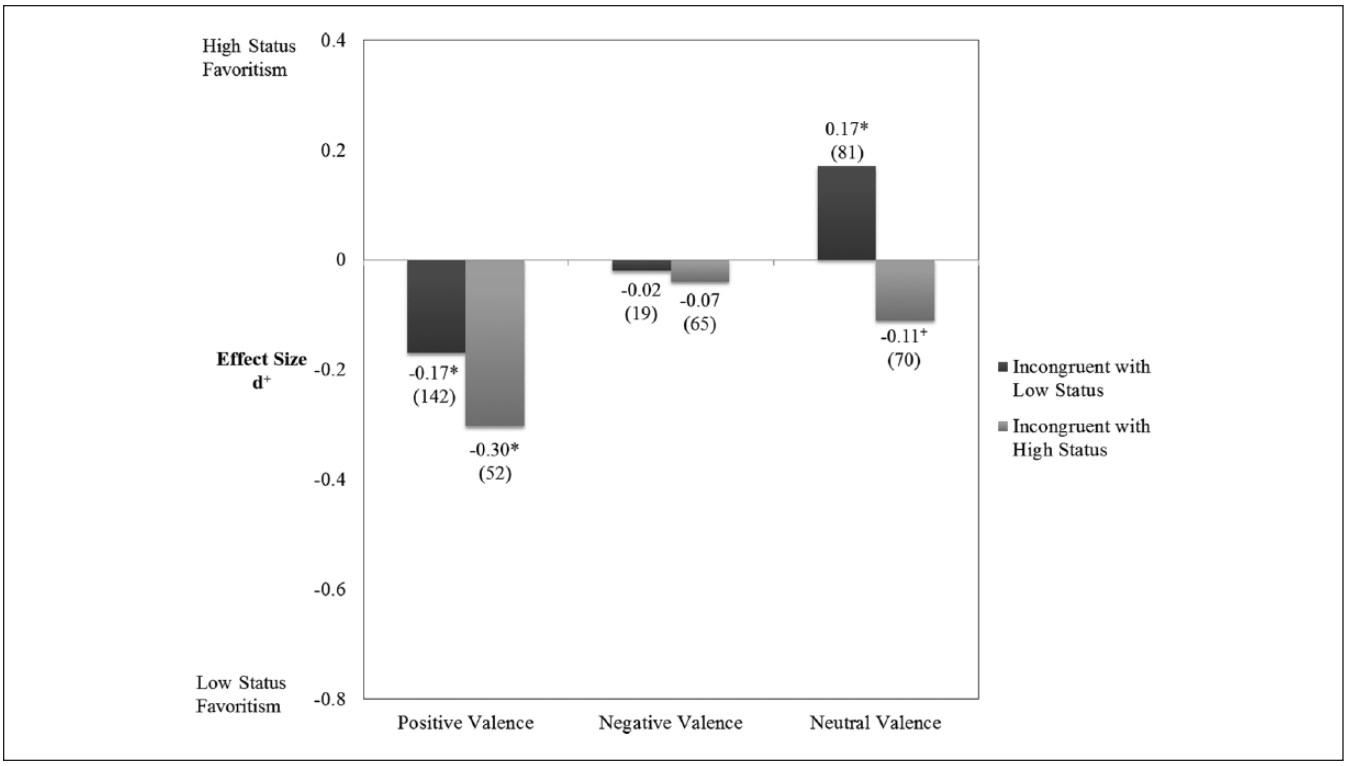

Figure 7. Mean high status–low status effect sizes moderated by valence of person-based information and whether the information was incongruent with either the high status group or the low status group.

Note. The high status–low status effect size is representative of evaluations of low status targets subtracted from evaluations of high status targets. A negative value reveals more positive evaluations of low status targets, and a positive value reveals more positive evaluations of high status target. Values within parentheses indicate the number of effect sizes, *k*.

*Significantly different from 0. +Different from 0, $p = .06$.

Table 3. Fixed and Random Effects for Expectancy-Violation Theory.

Fixed effects	γ	*SE*
Intercept	−.08*	0.04
Pos vs. Neg	−.21*	0.05
Val vs. Neut	.16*	0.07
Incong low vs. Incong high	.14*	0.03
Pos vs. Neg × Incong low vs. Incong high	.11	0.11
Val vs. Neut × Incong low vs. Incong high	.21*	0.06
Random effects	**Variance**	
Level 1 intercept	0.05*	
Level 2 intercept	0.08*	

Note. Pos = positive valence; neg = negative valence; val = valenced; neut = neutral; incong low = incongruent with low status; incong high = incongruent with high status.

*$p < .05$.

$\gamma_{500} = .21$, $t(171) = 3.68$, $p < .001$. The random effects indicate that there was a significant within-study and between-study variance in the effect sizes.

The model implied effect sizes associated with the high status–low status effect sizes are depicted in Figure 4. For the high status–low status effect sizes, a negative sign (e.g., −0.25) indicates more favorable evaluations of low status targets and a positive sign (e.g., +0.25) indicates more favorable evaluations of high status targets. As shown in the left side of the figure, when person-based information was positive and incongruent with the *low* status group, low status targets were evaluated more favorably than high status targets, $d_+ = -0.17$, $\chi^2(1) = 13.20$, $p < .01$. This is consistent with the predictions of the theory. Contrary to the predictions, however, when information was positive and incongruent with the *high status group*, perceivers also evaluated low status targets more favorably than high status targets, $d_+ = -0.30$, $\chi^2(1) = 26.20$, $p \leq .01$.

As shown in the middle bars of Figure 4, when person-based information was negative perceivers evaluated high status and low status targets similarly, regardless of whether the information was incongruent with the low status group, $d_+ = -0.04$, $\chi^2(1) = .42$, *ns*, or incongruent with the high status group, $d_+ = -0.02$, $\chi^2(1) = .03$, *ns*.

Finally, when neutral person-based information was incongruent with the low status group, perceivers evaluated low status targets more negatively than high status targets, $d_+ = 0.17$, $\chi^2(1) = 5.05$, $p < .05$. Conversely, when information was neutral and incongruent with the high status group, perceivers

Table 4. Summary of Empirical Support of Complexity-Extremity, Subjective Group Dynamics, and Expectancy-Violation Theories.

Theory	Empirical support	Nonsupport	Findings that are irrelevant to theory specifics
Complexity-extremity	• Evaluative extremity for positive outgroup targets among high status perceivers, with low cognitive complexity for outgroup. • No evaluative extremity toward negative outgroup targets among low status perceivers, with high cognitive complexity for outgroup.[a] • No outgroup extremity, either in positive or in negative direction, among gender groups, with high cognitive complexity for outgroup.[a] • Greater outgroup polarization among high status perceivers, with low cognitive complexity for outgroup. • No outgroup polarization among gender groups, with high cognitive complexity for outgroup.	• No evaluative extremity toward negative outgroup targets among high status perceivers, with low cognitive complexity for outgroup. • Ingroup polarization among low status perceivers (but not gender groups).	• Ingroup favoritism toward neutral targets among high status perceivers, with low cognitive complexity for outgroups, but when status is defined by gender.
Subjective group dynamics	• Derogation of negative, norm-violating ingroup targets ("Black-sheep effect"). • Derogation toward neutral, norm-violating ingroup targets.[a] • Ingroup favoritism toward neutral, norm-consistent targets.	• No ingroup favoritism toward positive pro-norm target. • No ingroup favoritism toward positive, norm-consistent targets (inconsistent with black sheep effect).	
Expectancy-violation	• Evaluative extremity toward low status targets whose positive attributes violate stereotyped expectations.	• No extremity toward high status targets whose positive attributes violate expectations.[a] • No evaluative extremity toward targets when individuating information was negative.[a] • Nonsignificant tendency to evaluate low status targets whose negative attributes violate expectations more negatively.[a]	• Low status targets evaluated more favorably when information was neutral and congruent with stereotyped expectancies for low status group.[a] • High status targets evaluated more favorably when information was neutral and congruent with stereotyped expectancies for high status group.[a]

[a]Findings not previously revealed by complexity-extremity, subjective group dynamics, or expectancy-violation studies.

tended to evaluate high status targets more negatively than low status targets, $d_+ = -0.11$, $\chi^2(1) = 3.25$, $p = .06$.

Discussion

Both categorical information (group-based) and individuating information (person-based) about targets matter for social judgment. This is a theme in a number of models of impression formation, including dual-process (Brewer, 1988; Fiske & Neuberg, 1990) and connectionist (Kunda & Thagard, 1996) perspectives. But rather than highlighting when each component may dominate impressions, the current meta-analytic findings point to a variety of factors that affect whether group-based information interacts with

person-based information to produce evaluative extremity. The findings suggest that, given equivalent person-based information, targets from different groups are not necessarily evaluated the same. Rather, evaluative extremity toward members of particular groups may result, depending on the cognitions, emotions, and motivations that are elicited by particular combinations of group-based and person-based information.

This meta-analysis was guided by three theoretical perspectives: complexity-extremity (Linville, 1982; Linville & Jones, 1980), subjective group dynamics (Marques et al., 1998; Marques et al., 1988), and expectancy-violation theories (Bettencourt et al., 1997; Jussim et al., 1987). Each covers different ground regarding why and what pattern of evaluative extremity should be observed. Overall, the meta-analysis showed support for many, but not all, of the theoretical predictions and uncovered new findings of theoretical significance that go beyond these perspectives (see Table 4).

In addition to reviewing the specific implications for these theories, we consider how the findings speak to several broad, general points about person perception and the relationship between group-based and person-based information in forming impressions. First, social category information tended to matter when target information was *positive*. Generally, both high status and low status evaluators favored low status targets when personal information was positive. Such favoritism toward low status targets who convey positive information is surprising, especially in light of the notion of double standards—that members of negatively stereotyped groups need to "work twice as hard" to be perceived as "half as good" (Foschi, 1998, 2000). These data suggest, instead, that evidence of positivity can overcome initial negative group-based expectations. Still, this pattern could be driven by *shifting standards* or by the tendency to judge group members on stereotyped dimensions relative to within-group expectations (Biernat, Manis, & Nelson, 1991). With low expectations, targets seem "better," especially on subjective rating scales that are typical of the studies in our meta-analysis. Such positivity could also be driven by social desirability concerns, political correctness norms (Crandall & Eshleman, 2003), or motivations to appear unprejudiced (Harber, 1998; Plant & Devine, 1998). It may be particularly easy for perceivers to meet goals of nonprejudice when they are called upon to judge negatively stereotyped targets who have positive qualities. Additional work is needed to examine the viability of these accounts.

Next, most of our findings suggested that negative information renders group-based information irrelevant. This finding is consistent with Kunda and Thagard's (1996) suggestion that diagnostic person-based information overrides the influence of group-based information, but our findings modify this claim by highlighting that diagnostic *negative* information can override the effects of group-based information. This finding may also reflect the strong general influence of negative information on interactions, emotions, and events (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Notably, there was one case, however, in which negative person-based information did not override the effects of group-based information. In particular, when negative information violated group-based norms, ingroup members were judged more negatively than their outgroup counterparts. In this case, negative information led to greater derogation of ingroup norm violators (i.e., black sheep).

A third general finding from the meta-analysis was revealed for targets that were depicted neutrally. Evaluators preferred persons that fit either group norms or stereotyped expectations. Specifically, ingroup targets depicted with neutral information were judged more favorably than comparable outgroup targets, but this pattern reversed to ingroup derogation when the neutral information violated group norms. And regardless of whether targets were members of high or low status groups, evaluators appeared to favor stereotype-confirming targets in the absence of clear valence of person-based information. These findings suggest a general preference for the expected (Olson et al., 1996). In the absence of clearly valenced information, favoritism toward group members who meet expectations may be the rule.

These general findings touch on our three guiding theories and ultimately suggest that none provides a complete picture of how and when group- and person-based information combine to affect judgment. In what follows, we consider specific support for each of the perspectives, before returning to some general conclusions.

Complexity-Extremity Theory

Supporting complexity-extremity theory, we found that high status perceivers evaluated positive outgroup targets more favorably than positive ingroup targets. These more positive evaluations of outgroup targets were found only for high status evaluators who were expected to have less complex cognitive schemas for their outgroups. Contrary to the theory, negative outgroup targets were not evaluated more negatively than otherwise identical ingroup targets. In their original research, Linville (1982, Study 2) and Linville and Jones (1980, Study 2) also found that negatively valenced outgroup targets were judged only marginally more negatively than comparable ingroup targets. Outgroup polarization seems to be largely one sided, occurring when person-based information is positive.

Also, the meta-analysis revealed several findings not previously shown in the literature but that are consistent with the complexity-extremity perspective. When evaluators were expected to have relatively complex cognitive schemas for their outgroups (when groups were based on gender, or when evaluators were members of low status groups), there was no evidence of evaluative extremity toward either positively or negatively depicted outgroup targets. Although suggested by Linville et al. (1996), this finding provides the first clear demonstration that complex cognitive representations of

outgroups moderate the polarization of positively valenced outgroup targets. This pattern also qualifies those of Finkelstein et al. (1995), who reported that high status evaluators express more ingroup bias than low status evaluators. Our findings suggest that this is true only when high status evaluators have relatively simple cognitive representations of their lower status outgroups.

In addition, when person-based information was relatively neutral, high status perceivers tended to evaluate ingroup targets *more positively* than outgroup targets (i.e., ingroup bias), but only when these high status perceivers were expected to have limited cognitive complexity for their outgroups. Evidence of ingroup bias is certainly not new to the literature, but these findings provide the first suggestion that, when person-based information is neutral, less cognitive complexity about the outgroup may facilitate ingroup bias and more cognitive complexity may reduce it.

Subjective Group Dynamics Theory

Supporting the subjective group dynamics model, we found that ingroup targets were evaluated more negatively than similarly depicted outgroup targets when person-based information was negative *and* norm violating. Contrary to the predictions, however, perceivers did not evaluate ingroup targets more positively than outgroup targets when person-based information was positive and norm violating. Also, in the absence of norm violation, ingroup and outgroup targets were judged similarly whether person-based information was positive or negative.

The meta-analysis revealed a pattern of findings that, although consistent with subjective group dynamics, had not been revealed in tests of the model. Specifically, when person-based information was neutral and norm violating, ingroup targets were evaluated more negatively than outgroup targets. This finding suggests that relatively neutral behaviors that violate norms may be perceived as negative by ingroup evaluators. Overall, the meta-analytic findings suggest that group-based motives are relevant to evaluation, and that most norm-violating behaviors engender negative evaluations of ingroup members. These findings illustrate the strong pressure group members likely feel to comply with ingroup norms (e.g., Festinger & Thibaut, 1951; Schachter, 1951).

Expectancy-Violation Theory

The pattern of mean effect sizes showed only minimal support for expectancy-violation theory. Consistent with the theory, low status targets were evaluated more favorably when the individuating information was positive and incongruent with stereotypes for their group, but it was also the case that low status targets were evaluated more favorably when the positive information was incongruent with stereotypes for the high status group. Said in another way, the latter finding suggests that perceivers favor low status targets when positive individuating information confirms stereotypes about the low status group. Also, the findings failed to show that negative stereotype incongruency differentially affected evaluations of low status and high status targets. Instead, the findings suggested that the negative individuating information overrode any effects of stereotype incongruency. Lack of support for expectancy violation, in this meta-analysis, may reflect the fact that most of the targets were depicted as merely incongruent with stereotypes. To observe evaluative extremity toward low and high status targets, it may be necessary that individuating information clearly violates stereotypes.

A final novel finding of the meta-analyses was that, when person-based information was neutral and incongruent with stereotypes, targets were evaluated more negatively. That is, perceivers evaluated both low status and high status targets more negatively when they were depicted in ways that were incongruent with stereotyped expectancies. This pair of findings is consistent with theorizing that expectancy violations, in general, should engender negative affect (Olson et al., 1996), and suggest that perceivers, may in some instances, prefer stereotype-consistent targets.

Limitations

Our meta-analysis included studies that were derived from a large variety of theoretical perspectives. As such, the vast majority of eligible studies were not designed to test the three theoretical models that were our focus. This approach is a plus for generalizability and avoidance of bias introduced by lab-specific effects, but it may be less than ideal for specifically testing the three theoretical hypotheses. Nevertheless, by summarizing a large number of studies, the meta-analytic investigation had the capacity to provide an understanding of when evaluative extremity is *most* likely and under which circumstances.

Another methodological limitation of the meta-analysis is that we were unable to test the influence of the proposed mediators (i.e., cognitive complexity, social identity motives, expectancy violation). Only a small handful of studies have measured these mediators, making it impossible to meta-analytically test them. Although some meta-analysts ask judges to render their perceptions about theoretical variables of this sort, this methodology would have been difficult for our purposes (e.g., it would have been difficult for judges to accurately determine the cognitive complexity or social motivation of participants).

Finally, too few studies in our meta-analysis included manipulations of other moderators (e.g., cognitive busyness, power, accountability, accuracy, mood, interdependence, type of judgment scale, etc.) that have been demonstrated to influence evaluations. Thus, we were unable to test the ways in which these variables interacted with the moderators relevant to our theoretical perspectives (e.g., valence, group

membership, status). Integrating such motivational and cognitive moderators with the markers of evaluative extremity could prove useful providing a more full picture of when judgments are biased by group-based information.

Conclusion

Social judgment is a part of everyday life, and as such, it is important to understand whether and how social category information interacts with individuating information to influence judgment. The current meta-analytic findings suggest that group-based information can bias judgment, in either evaluative direction, depending on information valence, ingroup/outgroup status, norm violation, and stereotypes. These outcomes have implications for a variety of settings in which evaluations are made about people, including informal judgments in everyday encounters with others as well as formal evaluations in the school or the workplace. The meta-analytic findings point to the likely reality that, all else being equal, people from different groups are not necessarily evaluated equally. Based on the cognitions, affect, and motivations elicited by a particular combination of group-based and person-based information, evaluative extremity may result.

Our findings suggest that positive information can be particularly beneficial to low status, outgroup, and negatively stereotyped targets. Whether this positive evaluative extremity is genuine, driven by the use of low standards, or a by-product of concerns about appearing prejudiced (e.g., Harber, 1998) is unclear. But negative group-based expectations, which are then violated by positive information, seem to be at the heart of this pattern. We also found that negative information generally overrode the influence of group-based effects, except when ingroup members violated group-based norms. Neutral information generally prompted positivity toward the *expected*: stereotype-confirming targets and ingroup norm supporters.

What do these findings mean for our three theoretical perspectives? We suggest that all of these models have value, but none provide a full account of when and why evaluative extremity occurs. Ingroup status, valence of individuating information, group expectancies, and group norms all matter for social judgment, though they interact in complex ways, and undoubtedly in combination with variables outside the purview of our meta-analysis. Rather than dismiss the models for their shortcomings and offer a new alternative model, we suggest that future researchers consider the moderators identified here, plan research that includes the "small k" conditions identified in our meta-analysis, and incorporate measures of potential mediating mechanisms, including cognitive complexity, group-protective motives, and perceived expectancy violation. Expanding our knowledgebase in these ways will help the field move toward a fuller, more nuanced understanding of social judgment.

Acknowledgment

The authors are grateful to the action editor and the reviewers for their positive contributions to the work reported herein. Also, the authors acknowledge Ed Merkle for offering his statistical advice for the betterment of the work.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by a grant from the National Science Foundation provided to B. Ann Bettencourt (NSF SBR 0345241) and Monica Biernat (NSF SBR 0345685).

Notes

1. One additional model, ambivalence-amplification theory, also predicts outgroup polarization but offers a motivational rather than cognitive explanation for these effects. This model suggests that perceivers generally have ambivalent feelings toward members of stigmatized groups, which derive, in turn, from the combination of "sympathy for the underdog" and hostility regarding the perceived deviation from dominant society's norms and values (Katz, 1981; Katz & Hass, 1988; Katz, Wackenhut, & Hass, 1986; Katz, Glass, & Cohen, 1973). Ambivalence is a tension state that can be reduced by amplifying responses to individual members of stigmatized groups; the direction of the amplification is cued by the valence of the individuating information. We do not focus on this model here because of the difficulty of ascertaining whether ambivalence characterizes attitudes toward the variety of outgroups used in the research in this meta-analysis. Still, patterns of outgroup polarization may be viewed as consistent with this model.
2. An ideal meta-analytic test of complexity-extremity predictions would involve within-subjects study designs that include categorical and person-based information, measure cognitive complexity, and test the association between judgmental polarization and cognitive complexity. Unfortunately, very few studies using within-subjects designs have measured the cognitive complexity of the evaluators. For this reason, we were unable to test the specific mediating mechanism of cognitive complexity in this meta-analysis, though we did make assumptions about whether or not evaluators were likely to have complex representations of outgroups. Similar assumptions guided analyses of the subjective group dynamics and expectancy-violation models.
3. We also coded for several methodological features of the studies. Some of these codes might be interpreted as representing varying degrees of the salience of group-based or person-based information. We coded if the cover story and materials were hypothetical, used deception (i.e., made the targets seem real), or involved real targets. Also, we coded the medium in which the stimuli were presented (i.e., on paper, audiovisual, or in person), the type of stimuli (scenario, application materials, other detailed materials, or interaction of some type), and age

of the participants (grade/high school students, college students, older adults/community). For both the ingroup–outgroup and high status–low status effect sizes, these methodological variables had no significant effect with only two exceptions. For the ingroup/outgroup effect sizes, only, college students tended to express more favorable attitudes toward outgroup targets, and for the high status–low status effect sizes, only, high status targets were evaluated more favorably than low status targets when targets were seemingly real, via deception, or were actually real, compared with when targets were merely hypothetical. It was impossible to test these methodological variables along with the theoretical moderators (e.g., valence, stereotype congruency, evaluator status) because of low numbers of studies (k) or no studies available in some of the resulting cells in the analysis.

4. For each effect size (d_i), an estimate of the mean effect size with d_i excluded ($d_{+(i)}$) was calculated using a random effects model method of estimation separately by type of effect size and between-subjects versus within-subjects studies. A random effects model was used as it incorporates estimates of variance randomly distributed across studies in addition to the within-study sampling variance; note that this was the only population effect size calculation for which the mixed-effects model was not used. A residual, $d_{+(i)} - d_i$, was calculated for each effect size. Because $d_{+(i)}$ is independent of d_i, we can use this residual to examine the extent to which a particular effect size deviates from the independent estimate of the population effect size without that effect size, d_i, influencing the population effect size. It follows that the effect sizes that were furthest from the mean (i.e., the outliers) would have the largest residual. The residuals were standardized and we defined conservatively outliers as those effect sizes with a standardized residual that was greater than 3.5 *SD* above or below $d_{+(i)}$.
5. Note that, whereas Kalaian and Raudenbush (1996) used two-level models to model their effect sizes, we have chosen to organize our effect sizes with three-level models. To illustrate the complexities that lead to the three-level conceptualization, we examine Study 1 from Linville and Jones's (1980) report: "Polarized appraisals of out-group members." In this study, the researchers examined the effects of person-based information on positively valenced evaluations (ability, motivation, activeness, and liking) of law school applicants in a 2 (Gender) by 2 (Race: Black or White) between-subjects design. Using these data, we were able to calculate effect sizes that were characterized by the valence of the individuating information (positively valenced), their congruence with the high (White and/or male) or low (Black or female) status group in the context of the study, and group membership of the person making the evaluation (high or low status group member). The characterization of these effect sizes is nested within evaluations (i.e., ability, motivation, etc.) that are selected from a "universe" of positively valenced evaluative measures, and simultaneously nested within between-group comparisons (male vs. female; Black vs. White) selected from a "universe" of group comparisons (e.g., mentally ill vs. mentally healthy, heterosexual vs. homosexual, etc.). Finally, the choices for evaluations and comparisons are nested within each study. Note that the choice of a three-level model required the use of full-level maximum-likelihood estimation in which both the variance components and fixed effects coefficients are estimated by means of maximum likelihood.
6. To capture the degree of evaluative polarization, we calculated a supplemental effect size that compared positive person-based information with negative person-based information (i.e., positive–negative), while holding the target's category constant. A larger positive–negative effect size reveals greater evaluative polarization. At the effect size level (Level 1), we specified contrasts to indicate whether the target was from a high (0.5) or low (−0.5) status group, and separate contrasts to indicate whether the evaluator was from a high (0.5) or low (−0.5) status group. We also included the interaction between the two. Finally, at the within-study level (Level 2), we included the contrasts to examine whether the comparison was between males and females or other groups as moderators of the Level 1 effects. The mean positive–negative effect size (1.68) across all conditions was significant, $t(41) = 7.15$, $p < .001$. There was a significant main effect of target group status on the effect size, $\gamma_{100} = -.36$, $t(64) = -4.65$, $p < .001$, which was qualified by whether the groups being compared were male–female or not, $\gamma_{110} = .33$, $t(64) = 2.14$, $p < .05$. There was a significant main effect of evaluator group status on the effect size, $\gamma_{200} = -.27$, $t(64) = -2.97$, $p < .01$. Finally, there was a significant target by evaluator group status interaction, $\gamma_{300} = .33$, $t(64) = 2.08$, $p < .05$. Random effects indicated that there remained significant within-study and between-study variance in the effect sizes. Somewhat consistent with complexity-extremity theory, for high status evaluators (i.e., not gender groups), the magnitude of the positive–negative effect size revealed polarization for outgroup targets, $d_+ = 1.87$, $\chi^2(1) = 39.68$, $p < .01$. Although, the positive–negative effect size for ingroup targets was relatively large, $d_+ = 1.58$, $\chi^2(1) = 20.91$, $p < .01$. Consistent with the theory, outgroup polarization was not revealed for low status perceivers. But somewhat unexpectedly, a pattern of *ingroup* polarization emerged: The positive–negative effect size was larger for *ingroup* targets, $d_+ = 2.33$, $\chi^2(1) = 60.32$, $p < .01$, than that for outgroup targets, $d_+ = 1.56$, $\chi^2(1) = 54.83$, $p < .01$. For male and female evaluators, judging gender ingroup and outgroup targets, the positive–negative effect sizes merely reflected that, regardless of the type of target, positively depicted targets were judged more favorably.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*Abbey, A., & Melby, C. (1986). The effects of nonverbal cues on gender differences in perceptions of sexual intent. *Sex Roles*, *15*, 283-298.

*Abrams, D., & Hogg, M. A. (1987). Language attitudes, frames of reference, and social identity: A Scottish dimension. *Journal of Language and Social Psychology*, *6*, 201-213.

*Abrams, D., Marques, J., Brown, N., & Dougill, M. (2002). Anti-norm and pro-norm deviance in the bank and on the campus: Two experiments on subjective group dynamics. *Group Processes & Intergroup Relations*, *5*, 163-182. doi:10.1177/1368430202005002922

Abrams, D., Marques, J. M., Brown, N. J., & Henson, M. (2000). Pro-norm and anti-norm deviance. *Journal of Personality and Social Psychology*, *78*, 906-912.

*Abrams, D., Randsley de Moura, G., Marques, J. M., & Hutchison, P. (2008). Innovation credit: When can leaders oppose their

group's norms?*Journal of Personality and Social Psychology*, *95*, 662-678.

*Abrams, D., Rutland, A., & Cameron, L. (2003). The development of subjective group dynamics: Children's judgments of normative and deviant in-group and out-group individuals. *Child Development*, *74*, 1840-1856.

*Abrams, D., Rutland, A., Cameron, L., & Marques, J. M. (2003). The development of subjective group dynamics: When in-group bias gets specific. *British Journal of Developmental Psychology*, *21*, 155-176.

*Abramson, P. R., Goldberg, P. A., Greenberg, J. H., & Abramson, L. M. (1977). The talking platypus phenomenon: Competency ratings as a function of sex and professional status. *Psychology of Women Quarterly*, *2*, 114-124.

*Arkkelin, D., & O'Connor, R. (1992). The "good" professional: Effects of trait-profile gender type, androgyny, and likableness on impressions of incumbents of sex-typed occupations. *Sex Roles*, *27*, 517-532.

*Atkinson, D. R., Maruyama, M., & Matsui, S. (1978). Effect of counselor race and counseling approach on Asian Americans' perceptions of counselor credibility and utility. *Journal of Counseling Psychology*, *25*, 76-83.

*Banikiotes, P. G., & Merluzzi, T. V. (1981). Impact of counselor gender and counselor sex role orientation on perceived counselor characteristics. *Journal of Counseling Psychology*, *28*, 342-348.

*Barnett, M. A., McMinimy, V., Flouer, G., & Masbad, I. (1987). Adolescents' evaluations of peers' motives for helping. *Journal of Youth and Adolescence*, *16*, 579-586.

*Bar Tal, D., Raviv, A., & Arad, M. (1989). Effects of information of student-teachers' stereotypic perception of pupils. *British Journal of Educational Psychology*, *59*, 143-154.

*Bartol, K. M., & Butterfield, D. (1976). Sex effects in evaluating leaders. *Journal of Applied Psychology*, *61*, 446-454.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323-370.

*Bean, L. D. (1977). Interpersonal perception among college students as a function of sex and level of assertiveness. *Dissertation Abstracts International*, *38*, 1292.

*Beauvais, C., & Spence, J. T. (1987). Gender, prejudice, and categorization. *Sex Roles*, *16*, 89-100.

*Bègue, L. (2001). Social judgment of abortion: A black-sheep effect in a catholic sheepfold. *Journal of Social Psychology*, *141*, 640-649.

*Benedict, M. E., & Levine, E. L. (1988). Delay and distortion: Tacit influences on performance appraisal effectiveness. *Journal of Applied Psychology*, *73*, 507-514.

*Bernstein, B. L., & Figiolo, S. W. (1983). Gender and credibility introduction effects on perceived counselor characteristics. *Journal of Counseling Psychology*, *30*, 506-513.

*Bettencourt, B. A., Dill, K. E., Greathouse, S. A., Charlton, K., & Mulholland, A. (1997). Evaluations of ingroup and outgroup members: The role of category-based expectancy violation. *Journal of Experimental Social Psychology*, *33*, 244-275. doi:10.1006/jesp.1996.1323

Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin*, *119*, 422-447. doi:10.1037/0033-2909.119.3.422

*Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*, 5-20. doi:10.1037/0022-3514.66.1.5

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, *60*, 485-499. doi:10.1037/0022-3514.60.4.485

*Biernat, M., & Vescio, T. K. (2002). She swings, she hits, she's great, she's benched: Implications of gender-based shifting standards for judgment and behavior. *Personality and Social Psychology Bulletin*, *28*, 66-77.

*Biernat, M., Vescio, T. K., & Billings, L. S. (1999). Black sheep and expectancy violation: Integrating two models of social judgment. *European Journal of Social Psychology*, *29*, 523-542. doi:10.1002/(sici)1099-0992(199906)29:4<523::aid-ejsp944>3.0.co;2-j

*Blanchard, F. A., Weigel, R. H., & Cook, S. W. (1975). The effect of relative competence of group members upon interpersonal attraction in cooperating interracial groups. *Journal of Personality and Social Psychology*, *32*, 519-530.

*Bodenhausen, G. V. (1987). *Effects of social stereotypes on evidence processing: The cognitive basis of discrimination in juridic decision making* (Doctoral dissertation). University of Illinois at Urbana-Champaign.

Borenstein, M., Hedges, L., & Rothstein, H. (2007). *Introduction to meta-analysis*. Available from www.meta-analysis.com

*Braithwaite, V. A. (1986). Old age stereotypes: Reconciling contradictions. *Journal of Gerontology*, *41*, 353-360.

*Breakwell, G. M., Vignoles, V. L., & Robertson, T. (2003). Stereotypes and crossed-category evaluations: The case of gender and science education. *British Journal of Psychology*, *94*, 437-455.

Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds), *A dual process model of impression formation* (pp. 1-36). Hillsdale, NJ: Lawrence Erlbaum.

*Brewer, M. B., Ho, H. K., Lee, J. Y., & Miller, N. (1987). Social identity and social distance among Hong Kong schoolchildren. *Personality and Social Psychology Bulletin*, *13*, 156-165.

*Bridges, J. S., & del Ciampo, J. (1981). Children's perception of the competence of boys and girls. *Perceptual and Motor Skills*, *52*, 503-506.

*Brischetto, C. S., & Merluzzi, T. V. (1981). Client perceptions in an initial interview as a function of therapist sex and expertness. *Journal of Clinical Psychology*, *37*, 82-87.

*Brown, N. J., & Abrams, D. (2003). Despicability in the workplace: Effects of behavioral deviance and unlikeability on the evaluation of in-group and out-group members. *Journal of Applied Social Psychology*, *33*, 2413-2426.

*Buckly, F. (1979). The acceptability of the stereotypic and non-stereotypic male and female. *Dissertation Abstracts International*, *40*, 1949-1950.

*Byrne, D., London, O., & Reeves, K. (1968). The effects of physical attractiveness, sex and attitude similarity on interpersonal attraction. *Journal of Personality*, *36*, 259-271.

*Byrne, D., & McGraw, C. (1964). Interpersonal attraction toward Negroes. *Human Relations*, *17*, 201-213.

*Carranza, M. A., & Ryan, E. B. (1975). Evaluative reactions of bilingual Anglo and Mexican American adolescents towards speakers of English and Spanish. *International Journal of the Sociology of Language*, *6*, 83-104.

*Carter, J. A. (1978). Impressions of counselors as a function of counselor physical attractiveness. *Journal of Counseling Psychology*, *25*, 28-34.

*Carver, C. S., Glass, D. C., Snyder, M. L., & Katz, I. (1977). Favorable evaluations of stigmatized others. *Personality and Social Psychology Bulletin*, *3*, 232-235.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

*Coleman, L. M., Jussim, L., & Kelley, S. H. (1995). A study of stereotyping: Testing three models with a sample of Blacks. *Journal of Black Psychology*, *21*, 332-356. doi:10.1177/00957984950214004

Cooper, H. M. (1989). *Integrating research: A guide for literature reviews* (2nd ed.). Thousand Oaks, CA: SAGE.

Cooper, H. M. (1998). *Synthesizing research: A guide for literature reviews* (3rd ed.). Thousand Oaks, CA: SAGE.

*Cooper, H. M., Baron, R. M., & Lowe, C. A. (1975). The importance of race and social class information in the formation of expectancies about academic performance. *Journal of Educational Psychology*, *67*, 312-319.

Crandall, C. S., & Eshleman, A. (2003). A justification -suppression model of the expression and experience of prejudice. *Psychological Bulletin*, *129*, 414-446.

Dean, M. A., Roth, P. L., & Bobko, P. (2008). Ethnic and gender subgroup differences in assessment center ratings: A meta-analysis. *Journal of Applied Psychology*, *93*, 685-691. doi:10.1037/0021-9010.93.3.6852008-05281-015

*Deaux, K. (1972). To err is humanizing. But sex makes a differences. *Representative Research in Social Psychology*, *3*, 20-28.

*Deaux, K., & Taynor, J. (1973). Evaluation of male and female ability: Bias works two ways. *Psychological Reports*, *32*, 261-262.

*Decker, W. H. (1986). Occupation and impressions: Stereotypes of males and females in three professions. *Social Behavior and Personality*, *14*, 69-75.

*De Cremer, D., & Vanbeselaere, N. (1999). I am deviant, because . . . : The impact of situational factors upon the black sheep effect. *Psychologica Belgica*, *39*, 71-79.

*DeJong, W. (1980). The stigma of obesity: The consequences of naïve assumptions concerning the causes of physical deviance. *Journal of Health and Social Behavior*, *21*, 75-87.

*De La Cruz-Barrera, A. (1984). Gender, race, expectancy, and interpersonal behavior. *Dissertation Abstracts International*, *45*, 701.

*del Boca, F. K., & Ashmore, R. D. (1980). Sex stereotypes and implicit personality theory. II. A trait-inference approach to the assessment of sex stereotypes. *Sex Roles*, *6*, 519-535.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*, 355-391. doi:10.1037/0033-2909.130.3.355

*Dienstbier, R. A. (1970). Positive and negative prejudice: Interactions of prejudice with race and social desirability. *Journal of Personality*, *38*, 198-215.

*Dietz, S. M., & Purkey, W. W. (1969). Teacher expectation of performance based on race of student. *Psychological Reports*, *24*, 694.

*Doherty, K. T. (1998). A mind of her own: Effects of need for closure and gender on reactions to nonconformity. *Sex Roles*, *38*, 801-819.

Doherty, R. W. (1998). Emotional contagion and social judgment. *Motivation and Emotion*, *22*, 187-209. doi:10.1023/a:1022368805803

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but . . . : A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*, 109-128. doi:10.1037/0033-2909.110.1.109

Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*, 3-22. doi:10.1037/0033-2909.111.1.3

*Eidelman, S., & Biernat, M. (2003). Derogating black sheep: Individual or group protection?*Journal of Experimental Social Psychology*, *39*, 602-609.

*Eidelman, S., Silvia, P. J., & Biernat, M. (2006). Responding to deviance: Target exclusion and differential devaluation. *Personality and Social Psychology Bulletin*, *32*, 1153-1164. doi:10.1177/0146167206288720

*Erber, J. T., Caiola, M. A., & Pupo, F. A. (1994). Age and forgetfulness: Managing perceivers' impressions of targets' capability. *Psychology and Aging*, *9*, 554-561.

*Erber, J. T., & Long, B. (2006). Perceptions of forgetful and slow employees: Does age matter?*The Journals of Gerontology*, *61B*, 333-339.

*Erber, J. T., Szuchman, L. T., & Etheart, M. E. (1993). Age and forgetfulness: Young perceivers' impressions of young and older neighbors. *International Journal of Aging & Human Development*, *37*, 91-103.

*Etaugh, C., & Folger, D. (1998). Perceptions of parents whose work and parenting behaviors deviate from role expectations. *Sex Roles*, *39*, 215-223.

*Etaugh, C., & Kasley, H. C. (1981). Evaluating competence: Effects of sex, marital status, and parental status. *Psychology of Women Quarterly*, *62*, 196-203.

Festinger, L., & Thibaut, J. (1951). Interpersonal communication in small groups. *The Journal of Abnormal and Social Psychology*, *46*, 92-99.

Finkelstein, L. M., Burke, M. J., & Raju, M. S. (1995). Age discrimination in simulated employment contexts: An integrative analysis. *Journal of Applied Psychology*, *80*, 652-663. doi:10.1037/0021-9010.80.6.652

*Fishbein, C. (1979). An investigation of sex-role stereotypes and assertive behavior in men and women. *Dissertation Abstracts International*, *40*, 912.

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

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. *23*, pp. 1-74). New York: Academic Press.

Ford, J. K., Kraiger, K., & Schechtman, S. L. (1986). Study of race effects in objective indices and subjective evaluations of performance: A meta-analysis of performance criteria. *Psychological Bulletin*, *99*, 330-337. doi:10.1037/0033-2909.99.3.330

Foschi, M. (1998). Double standards: Types, conditions, and consequences. *Advances in Group Processes*, *15*, 59-80.

Foschi, M. (2000). Double standards for competence: Theory and research. *Annual Review of Sociology*, *26*, 21-42.

*Furlong, M. J., Atkinson, D. R., & Casas, J. M. (1979). Effect of counselor ethnicity and attitudinal similarity on Chicano students' perceptions of the counselor credibility and attractiveness. *Hispanic Journal of Behavioral Sciences*, *1*, 41-53.

*Gallois, C., Callan, V., & Johnstone, M. (1984). Personality judgements of Australian Aborigine and White speakers: Ethnicity, sex, and context. *Journal of Language and Social Psychology*, *3*, 39-57.

*Garcia, D. M., Reser, A. H., Amo, R. B., Redersdorff, S., & Branscombe, N. R. (2005). Perceivers' responses to in-group and out-group members who blame a negative outcome on discrimination. *Personality and Social Psychology Bulletin*, *31*, 769-780.

*Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation: Associations influence the construal of person-based information. *European Journal of Social Psychology*, *33*, 573-589.

*Gerdes, E. P., & Kelman, J. H. (1981). Sex discrimination: Effects of sex-role incongruence, evaluator sex, and stereotypes. *Basic and Applied Social Psychology*, *2*, 219-226.

*Gordon, R. A., Rozelle, R. M., & Baxter, J. C. (1989). The effect of applicant age, job level, and accountability on perceptions of female job applicants. *Journal of Psychology*, *123*, 59-68.

*Gurwitz, S. B., & Marcus, M. (1978). Effects of anticipated interaction, sex, and homosexual stereotypes on first impressions. *Journal of Applied Social Psychology*, *8*, 47-56.

*Guttmann, J., Lazar, A., & Moran, K. (2008). Teachers' and school children's stereotypic perception of the child of divorce: 20 years later. *Journal of Divorce & Remarriage*, *42*, 131-141.

*Hagen, R. L., & Kahn, A. (1975). Discrimination against competent women. *Journal of Applied Social Psychology*, *5*, 362-376.

*Hammen, C. L., & Peters, S. D. (1977). Differential responses to male and female depressive reactions. *Journal of Consulting and Clinical Psychology*, *45*, 994-1001.

Harber, K. D. (1998). Feedback to minorities: Evidence of a positive bias. *Journal of Personality and Social Psychology*, *74*, 622-628.

*Harber, K. D. (2004). The positive feedback bias as a response to out-group unfriendliness. *Journal of Applied Social Psychology*, *34*, 2272-2297.

*Hart, A. J., & Morry, M. M. (1997). Trait inferences based on racial and behavioral cues. *Basic and Applied Social Psychology*, *19*, 33-48. doi:10.1207/s15324834basp1901_3

Hedges, L. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107-128. doi:10.2307/1164588

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

*Heilman, M. E., Amato, M. A., Kaplow, S., & Stathatos, P. (1993). When similarity is a liability: The effects of sex-based preferential selection on reactions to like and different sex others. *Journal of Applied Psychology*, *78*, 917-927.

*Heilman, M. E., Block, C. J., & Lucas, J. A. (1992). Presumed incompetent? Stigmatization and affirmative action efforts. *Journal of Applied Psychology*, *77*, 536-544.

*Heilman, M. E., Block, C. J., & Stathatos, P. (1997). The affirmative action stigma of incompetence: Effects of performance information. *Academy of Management Journal*, *40*, 603-625.

*Heilman, M. E., & Chen, J. J. (2005). Same behavior, different consequences: Reactions to men's and women's altruistic citizenship behavior. *Journal of Applied Psychology*, *90*, 431-441.

*Heilman, M. E., & Haynes, M. C. (2005). No credit where credit is due: Attributional rationalization of women's success in mixed-sex teams. *Journal of Applied Psychology*, *90*, 905-916.

*Heilman, M. E., Martell, R. F., & Simon, M. C. (1988). The vagaries of sex bias: Conditions regulating the undervaluation, equivaluation, and overvaluation of women job applicants. *Organizational Behavior and Human Performance*, *23*, 360-372.

*Heilman, M. E., & Okimoto, T. G. (2007). Why are women penalized for success at male tasks?: The implied communality deficit. *Journal of Applied Psychology*, *92*, 81-92.

*Heilman, M. E., Wallen, A. S., Fuchs, D., & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology*, *89*, 416-427.

*Hewstone, M., Jaspars, J., & Lalljee, M. (1982). Social representations, social attribution and social identity: The intergroup images of "public" and "comprehensive" schoolboys. *European Journal of Social Psychology*, *12*, 241-269.

*Isaacs, M. B. (1981). Sex role stereotyping and the evaluation of the performance of women: Changing trends. *Psychology of Women Quarterly*, *6*, 187-195.

*Izraeli, D. N. (1987). Sex effects in the evaluation of influence tactics. *Journal of Occupational Behaviour*, *8*, 79-86.

*Jackson, L. A. (1983). The influence of sex, physical attractiveness, sex role, and occupational sex-linkage on perceptions of occupational suitability. *Journal of Applied Social Psychology*, *13*, 31-44.

*Jackson, L. A., & Cash, T. F. (1985). Components of gender stereotypes: Their implications for inferences on stereotypic and nonstereotypic dimensions. *Personality and Social Psychology Bulletin*, *11*, 326-344.

*Jackson, L. A., Hymes, R. W., & Sullivan, L. A. (1987). The effects of positive information on evaluations of Black and White targets by Black and White subjects. *Journal of Social Psychology*, *127*, 309-316.

*Jackson, L. A., MacCoun, R. J., & Kerr, N. L. (1987). Stereotypes and nonstereotypic judgments: The effects of gender role attitudes on ratings of likability, adjustment, and occupational potential. *Personality and Social Psychology Bulletin*, *13*, 45-52.

*Jackson, L. A., & Sullivan, L. A. (1989). Cognition and affect in evaluations of stereotyped group members. *Journal of Social Psychology*, *129*, 689-672.

*Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects of attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology*, *65*, 69-84.

*Jenkins-Hall, K., & Sacco, W. P. (1967). Effect of client race and depression on evaluations by White therapists. *Journal of Social & Clinical Psychology*, *10*, 322-333.

*Jones, M. (1991). Stereotyping Hispanics and Whites: Perceived differences in social roles as a determinant of ethnic stereotypes. *Journal of Social Psychology*, *131*, 469-476.

*Jones, R. H. (1970). Sex prejudice: Effects of the inferential process of judging hireability. *Dissertation Abstracts International*, *31*, 1013A.

*Juodvalkis, J. L., Grefe, B. A., Hogue, M., Svyantek, D. J., & DeLamarter, W. (2003). The effects of job stereotype, applicant gender, and communication style on ratings in screening interviews. *International Journal of Organizational Analysis*, *11*, 67-84.

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

*Jussim, L., Fleming, C. J., Coleman, L., & Kohberger, C. (1996). The nature of stereotypes: II. A multiple-process model of evaluations. *Journal of Applied Social Psychology*, *26*, 283-312.

Kalaian, H. A., & Raudenbush, S. W. (1996). A multivariate mixed linear model for meta-analysis. *Psychological Methods*, *1*, 227-235. doi:10.1037/1082-989x.1.3.227

Katz, I. (1981). *Stigma: A social psychological analysis*. Hillsdale, NJ: Erlbaum.

Katz, I., Glass, D. C., & Cohen, S. (1973). Ambivalence, guilt, and the scapegoating of minority group victims. *Journal of Experimental Social Psychology*, *9*, 423-436.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*, 893-905.

Katz, I., Wackenhut, J., & Hass, R. G. (1986). Racial ambivalence, value duality, and behavior. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 35-60). Newark: Academic Press.

*Kelly, J. A., Kern, J. M., Kirkley, B. G., Patterson, J. N., & Keane, T. M. (1980). Reactions to assertive versus unassertive behavior: Differential effects for males and females and implications for assertiveness training. *Behavior Therapy*, *11*, 670-682.

*Kenyon, C., & Hewitt, J. (1989). Reaction to positive and negative behavior in same-sex vs. opposite-sex others. *Perceptual and Motor Skills*, *69*, 931-934.

Kernahan, C., Bartholow, B. D., & Bettencourt, B. A. (2000). Effects of category-based expectancy violation on affect-related evaluations: Toward a comprehensive model. *Basic and Applied Social Psychology*, *22*, 85-100. doi:10.1207/15324830051036171

*Khan, S. R., & Lambert, A. J. (1998). Ingroup favoritism versus black sheep effects in observations of informal conversations. *Basic and Applied Social Psychology*, *20*, 263-269.

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

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, *103*, 284-308. doi:10.1037/0033-295x.103.2.284

*Kwong See, S. T., & Heller, R. B. (2004). Judging older targets' discourse: How do age stereotypes influence evaluations?*Experimental Aging Research*, *30*, 63-73.

Landy, F. J. (2008). Stereotypes, bias, and personnel decisions: Strange and stranger. *Industrial and Organizational Psychology*, *1*, 379-392.

*Lao, R. C., Upchurch, W. H., Corwin, B. J., & Grossnickle, W. F. (1975). Biased attitudes toward females as indicated by ratings of intelligence and likeability. *Psychological Reports*, *37*, 1315-1320.

*Larimer, G. S., Beatty, E., & Broadus, A. C. (1988). Indirect assessment of interracial prejudices. *Journal of Black Psychology*, *14*, 47-56.

*Larkin, E. C., & Pines, H. A. (1979). No fat persons need apply: Experimental studies of the overweight stereotype and hiring preference. *Sociology of Work and Occupations*, *6*, 312-327.

*Larwood, L., Rand, P., & der Hovanessian, A. (1979). Sex differences in response to simulated employee discipline cases. *Personnel Psychology*, *32*, 539-550.

*Lee, D. Y., Hallberg, E. T., Jones, L., & Haase, R. F. (1980). Effects of counselor gender on perceived credibility. *Journal of Counseling Psychology*, *27*, 71-75.

*Lenney, E., Mitchell, L., & Browning, C. (1983). The effect of clear evaluation criteria on sex bias in judgments of performance. *Psychology of Women Quarterly*, *7*, 313-328.

*Lifschitz, S. (1983). Male and female careers: Sex-role and occupational stereotypes among high school students. *Sex Roles*, *9*, 725-735.

*Linville, P. W. (1982). The complexity-extremity effect and age-based stereotyping. *Journal of Personality and Social Psychology*, *42*, 193-211. doi:10.1037/0022-3514.42.2.193

Linville, P. W., Fischer, G. W., & Salovey, P. (1989). Perceived distributions of the characteristics of in-group and out-group members: Empirical evidence and a computer simulation. *Journal of Personality and Social Psychology*, *57*, 165-188.

Linville, P. W., Fischer, G. W., & Yoon, C. (1996). Perceived covariation among the features of ingroup and outgroup members: The outgroup covariation effect. *Journal of Personality and Social Psychology*, *70*, 421-436.

*Linville, P. W., & Jones, E. E. (1980). Polarized appraisals of outgroup members. *Journal of Personality and Social Psychology*, *38*, 689-703. doi:10.1037/0022-3514.38.5.689

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE.

Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology*, *39*, 821-831.

Lopez-Duran, N. L., Kovacs, M., & George, C. J. (2009). Hypothalamic-pituitary-adrenal axis dysregulation in depressed children and adolescents: A meta-analysis. *Psychoneuroendocrinology*, *34*, 1272-1283. doi:10.1016/j.psyneuen.2009.03.016

*Lott, B., & Saxon, S. (2002). The influence of ethnicity, social class and context on judgments about U.S. women. *Journal of Social Psychology*, *142*, 481-499.

*Mai-Dalton, R. R., Feldman-Summers, S., & Mitchell, T. R. (1979). Effect of employee gender and behavioral style on the evaluations of male and female banking executives. *Journal of Applied Psychology*, *64*, 221-226.

*Marques, J. M., Abrams, D., Paez, D., & Martinez Taboada, C. (1998). The role of categorization and in-group norms in judgments of groups and their members. *Journal of Personality and Social Psychology*, *75*, 976-988. doi:10.1037/0022-3514.75.4.976

*Marques, J. M., Abrams, D., & Serôdio, R. G. (2001). Being better by being right: Subjective group dynamics and derogation of in-group deviants when generic norms are undermined. *Journal of Personality and Social Psychology*, *81*, 436-447.

Marques, J. M., & Paez, D. (1994). The black sheep effect: Social categorisation, rejection of ingroup deviates, and perception of group variability. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. *5*, pp. 37-68). Chichester, England: Wiley.

Marques, J. M., Robalo, E. M., & Rocha, S. A. (1992). Ingroup bias and the "black sheep" effect: Assessing the impact of social identification and perceived variability on group judgements. *European Journal of Social Psychology*, *22*, 331-352. doi:10.1002/ejsp.2420220403

*Marques, J. M., & Yzerbyt, V. Y. (1988). The black sheep effect: Judgmental extremity towards ingroup members in inter- and intra-group situations. *European Journal of Social Psychology*, *18*, 287-292. doi:10.1002/ejsp.2420180308

Marques, J. M., Yzerbyt, V. Y., & Leyens, J.-P. (1988). The "black sheep effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology*, *18*, 1-16. doi:10.1002/ejsp.2420180102

*Martell, R. F. (1991). Sex bias at work: The effects of attentional and memory demands on performance ratings of men and women. *Journal of Applied Social Psychology*, *21*, 1939-1960.

*McCabe, A. E., & Brannon, L. A. (2004). Effect of shifting standards on evaluations of white subtype members. *Journal of Social Psychology*, *144*, 91-92.

McGuire, W. J., McGuire, C. V., Child, P., & Fujioka, T. (1978). Salience of ethnicity in the spontaneous self-concept as a function of one's ethnic distinctiveness in the social environment. *Journal of Personality and Social Psychology*, *36*, 511-520.

*McKenzie, N. J. (1996). When she says what he says: The influence of gender and bragging on evaluations of communication competence. *Dissertation Abstracts International*, *56*, 4204.

*McKimmie, B. M., Newton, C. J., Terry, D. J., & Schuller, R. A. (2004). Jurors' responses to expert witness testimony: The effects of gender stereotypes. *Group Processes & Intergroup Relations*, *7*, 131-143.

*McKirnan, D. J., Smith, C. E., & Hamayan, E. V. (1983). A sociolinguistic approach to the belief-similarity model of racial attitudes. *Journal of Experimental Social Psychology*, *19*, 434-447.

*McRae, M. B. (1991). Sex and race bias in employment decisions: Black women considered. *Journal of Employment Counseling*, *28*, 91-98.

*McRae, M. B. (1994). Influence of sex role stereotypes on personnel decisions of Black managers. *Journal of Applied Psychology*, *79*, 306-309.

*Memon, A., & Shuman, D. W. (1998). Juror perception of experts in civil disputes: The role of race and gender. *Law & Psychology Review*, *22*, 179-197.

*Mettrick, J., & Cowan, G. (1996). Gender stereotypes and predictions of performance: Women in air combat. *Journal of Social Behavior & Personality*, *11*, 105-120.

*Mischel, H. N. (1974). Sex bias in the evaluation of professional achievements. *Journal of Educational Psychology*, *66*, 157-166.

*Morris, S. B., Callahan, C. M., & Lord, R. G. (1992, August). Gender bias in initial perceptions and subsequent hiring decisions. Paper presented at APA, Washington DC.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*, 105-125. doi:10.1037/1082-989x.7.1.105

Neuberg, S. L., & Fiske, S. T. (1987). Motivational influences on impression formation: Outcome dependency, accuracy-driven attention, and individuating processes. *Journal of Personality and Social Psychology*, *53*, 431-444.

Nowak, C., & Heinrichs, N. (2008). A comprehensive meta-analysis of Triple P-Positive Parenting Program using hierarchical linear modeling: Effectiveness and moderating variables. *Clinical Child and Family Psychology Review*, *11*, 114-144. doi:10.1007/s10567-008-0033-0

Olson, J. M., Roese, N. J., & Zanna, M. P. (1996). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 211-238). New York, NY: Guilford Press.

*Paludi, M. A., & Bauer, W. D. (1983). Goldberg revisited: What's in an author's name?*Sex Roles*, *9*, 387-390.

*Peck, T. (1978). When women evaluate women nothing succeeds like success! The differential effects of status upon evaluations of male and female professional ability. *Sex Roles*, *4*, 205-213.

*Phelan, J. E., Moss-Racusin, C. A., & Rudman, L. A. (2008). Competent yet out in the cold: Shifting criteria for hiring reflects backlash toward agentic women. *Psychology of Women*, *32*, 406-413.

*Pheterson, G. I., Kiesler, S. B., & Goldberg, P. A. (1971). Evaluation of the performance of women as a function of their sex, achievement, and personal history. *Journal of Personality and Social Psychology*, *19*, 114-118.

*Pirri, C., Eaton, E., & Durkin, K. (1995). Australian professional women's evaluations of male and female written products. *Sex Roles*, *32*, 691-697.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 812-832.

*Puckett, J. M., Petty, R. E., Cacioppo, J. T., & Fischer, D. L. (1983). The relative impact of age and attractiveness stereotypes on persuasion. *Journal of Gerontology*, *38*, 340-343.

Raudenbush, S., Bryk, T., & Congdon, R. (2000). *HLM 6 Hierarchical Linear and Nonlinear Modeling: Scientific Software International, Inc*. Retrieved from http://www.ssicentral.com/hlm/

*Rickabaugh, C. A. (1995). College students' stereotypes of gender and political activism. *Basic and Applied Social Psychology*, *16*, 319-331.

*Riggs, J. M. (1998). Social roles we choose and don't choose: Impressions of employed and unemployed parents. *Sex Roles*, *39*, 431-443.

*Rojahn, K., & Willemsen, T. M. (1994). The evaluation of effectiveness and likability of gender-role congruent and gender-role incongruent leaders. *Sex Roles*, *30*, 109-119.

*Rose, F. L. (1978). Sex effects on managerial hiring decisions. *Academy of Management Journal*, *21*, 104-112.

*Rosen, B., & Jerdee, T. H. (1974). The influence of sex-role stereotypes on personnel decisions. *Journal of Applied Psychology*, *59*, 9-14.

*Rosen, B., Jerdee, T. H., & Prestwich, T. L. (1975). Dual-career mutual adjustment: Potential effects of discriminatory managerial attitudes. *Journal of Marriage and the Family*, *37*, 565-572.

*Rosenthal, D. A. (2004). Effects of client race on clinical judgment of practicing European American vocational rehabilitation counselors. *Rehabilitation Counseling Bulletin*, *47*, 131-141.

*Rosenthal, D. A., & Berven, N. L. (1999). Effects of client race on clinical judgment. *Rehabilitation Counseling Bulletin*, *42*, 243-264.

*Rosenthal, D. A., & Chapman, D. C. (1980). Sex-role stereotypes: Children's perceptions of occupational competence. *Psychological Reports*, *46*, 135-139.

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (Rev. ed.). Thousand Oaks, CA: SAGE.

Roth, P. L., Huffcutt, A. I., & Bobko, P. (2003). Ethnic group differences in measures of job performance: A new meta-analysis. *Journal of Applied Psychology*, *88*, 694-706. doi:10.1037/0021-9010.88.4.694

*Ruby, C. L., & Brigham, J. C. (1996). A criminal schema: The role of chronicity, race, and socioeconomic status in law enforcement officials' perceptions of others. *Journal of Applied Social Psychology*, *26*, 95-111.

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

*Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, *5*, 1004-1010.

*Saal, F. E., & Moore, S. (1993). Perceptions of promotion fairness and promotion candidates' qualifications. *Journal of Applied Psychology*, *78*, 105-110.

Sackett, P. R., & DuBois, C. L. (1991). Rater-ratee race effects on performance evaluation: Challenging meta-analytic conclusions. *Journal of Applied Psychology*, *76*, 873-877. doi:10.1037/0021-9010.76.6.873

Schachter, S. (1951). Deviation, rejection, and communication. *The Journal of Abnormal and Social Psychology*, *46*, 190-207.

*Scheier, M., Carver, C., Schultz, R., Glass, D., & Katz, I. (1978). Sympathy, self-consciousness and reactions to the stigmatized. *Journal of Applied Social Psychology*, *8*, 270-282.

*Schimel, J., Simon, L., Greenberg, J., Pyszczynski, T., Solomon, S., Waxmonsky, J., & Arndt, J. (1999). Stereotypes and terror management: Evidence that mortality salience enhances stereotypic thinking and preferences. *Journal of Personality and Social Psychology*, *77*, 905-926.

Schneider, S., Moyer, A., Knapp-Oliver, S., Sohl, S., Cannella, D., & Targhetta, V. (2010). Pre-intervention distress moderates the efficacy of psychosocial treatment for cancer patients: A meta-analysis. *Journal of Behavioral Medicine*, *33*, 1-14. doi:10.1007/s10865-009-9227-2

*Schroeder, H. E., Rakos, R. F., & Moe, J. (1983). The social perception of assertive behavior as a function of response class and gender. *Behavior Therapy*, *14*, 534-544.

*Sczensy, S., Spreemann, S., & Stahlberg, D. (2006). Masculine = competent? Physical appearance and sex as sources of gender-stereotypic attributions. *Swiss Journal of Psychology*, *65*, 15-23.

*Shaffer, D. R., & Johnson, R. D. (1980). Effects of occupational choice and sex role preference on the attractiveness of competent men and women. *Journal of Personality*, *42*, 505-519.

*Sinclair, L., & Kunda, Z. (1999). Reactions to a Black professional: Motivated inhibition and activation of conflicting stereotypes. *Journal of Personality and Social Psychology*, *77*, 885-904.

*Sommers, S. R., & Ellsworth, P. C. (2000). Race in the courtroom: Perceptions of guilt and dispositional attributions. *Personality and Social Psychology Bulletin*, *26*, 1367-1379.

*Spence, J. T., Helmreich, R., & Stapp, J. (1975). Likability, sex-role congruence of interest, and competence: It all depends on how you ask. *Journal of Applied Social Psychology*, *5*, 93-109.

*Stroebe, W., Insko, C. A., Thompson, V. D., & Layton, B. D. (1971). Effects of physical attractiveness, attitude similarity, and sex on various aspects of interpersonal attraction. *Journal of Personality and Social Psychology*, *18*, 79-91.

*Suchner, R. W., & More, D. M. (1975). Stereotypes of males and females in two occupations. *Journal of Vocational Behavior*, *6*, 1-8.

Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations?*Psychological Bulletin*, *105*, 409-429. doi:10.1037/0033-2909.105.3.409

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*, 1-39. doi:10.1146/annurev.ps.33.020182.000245

Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worschel & W. G. Austin (Eds.), *Psychology of intergroup relations* (2nd ed., pp. 7-24). Chicago, IL: Nelson-Hall.

*Taynor, J., & Deaux, K. (1973). When women are more deserving than men: Equity, attribution and perceived sex differences. *Journal of Personality and Social Psychology*, *28*, 360-367.

*Taynor, J., & Deaux, K. (1975). Equity and perceived sex differences: Role behavior as defined by the task, the mode, and the actor. *Journal of Personality and Social Psychology*, *32*, 381-390.

Turner, J. C. (1975). Social comparison and social identity: Some prospects for intergroup behaviour. *European Journal of Social Psychology*, *5*, 5-34. doi:10.1002/ejsp.2420050102

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, *20*, 454-463.

*Vonk, R., & Konst, D. (1998). Intergroup bias and correspondence bias: People engage in situational correction when it suits them. *British Journal of Social Psychology*, *37*, 379-385.

*Vrugt, A., & Schabracq, M. (1996). Stereotypes with respect to elderly employees: The contribution of attribute information and representativeness. *Journal of Community & Applied Social Psychology*, *6*, 287-292.

*Walsh, R. P., & Connor, C. L. (1979). Old men and young women: How objectively are their skills assessed?*Journal of Gerontology*, *34*, 561-568.

*Wann, D. L., Koch, K., Knoth, T., Fox, D., Aljubaily, H., & Lantz, C. D. (2006). The impact of team identification on biased predictions of player performance. *Psychological Record*, *56*, 55-66.

*Wayne, J. H., & Cordeiro, B. L. (2003). Who is a good organizational citizen? Social perception of male and female employees who use family leave. *Sex Roles*, *49*, 233-246.

*Whitehead, G. I., Smith, S. H., & Eichhorn, J. A. (1982). The effect of subject's race and other's race on judgments of causality for success and failure. *Journal of Personality*, *50*, 193-202.

Wilcox, R. (1995). *Statistics for the social sciences*. Orlando, FL: Academic Press.

*Wiley, M. G., & Crittenden, K. S. (1992). By your attributions you shall be known: Consequences of attributional accounts for professional and gender identities. *Sex Roles*, *27*, 259-276.

*Wiley, M. G., & Eskilson, A. (1985). Speech style, gender stereotypes, and corporate success: What if women talk more like men?*Sex Roles*, *12*, 993-1007.

*Yoder, J. D., & Schleicher, T. L. (1996). Undergraduates regard deviation from occupational gender stereotypes as costly for women. *Sex Roles*, *34*, 171-188.