Group Processes & Intergroup Relations

Attribution and Categorization Effects in the Representation of Gender Stereotypes

Joachim I. Krueger, Julie H. Hall, Paola Villano and Meredith C. Jones

The online version of this article can be found at:

Published by:

$SAGE

http://www.sagepublications.com

Additional services and information for *Group Processes & Intergroup Relations* can be found at:

Email Alerts: http://gpi.sagepub.com/cgi/alerts

Subscriptions: http://gpi.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.co.uk/journalsPermissions.nav

Citations http://gpi.sagepub.com/cgi/content/refs/11/3/401

Attribution and Categorization Effects in the Representation of Gender Stereotypes

Joachim I. Krueger
Brown University

Julie H. Hall
University at Buffalo

Paola Villano
Università di Bologna

Meredith C. Jones
University of Denver

Social stereotypes involve judgments of how typical certain personality traits are of a group. According to the *attribution hypothesis*, judgments of trait typicality depend on the perceived prevalence of the trait in the target group. According to the *categorization hypothesis*, such judgments depend on the degree to which a trait is thought to be more or less prevalent in the target group than in a relevant comparison group. A study conducted with women and men as target groups showed that the attribution hypothesis fit the data best when typicality ratings were made in an absolute format. When, however, typicality ratings were made in a comparative format (how typical is the trait of women as compared with men?), both hypotheses received support. Analytical derivation, supported by empirical evidence, showed an inverse relationship between the size of perceived group differences and their weight given in stereotyping. Implications for stereotype measurement and the rationality of social perception are discussed.

KEYWORDS   accentuation, bias, gender stereotypes, rational

HUMAN judgment, be it psychophysical or social, occurs in absolute and comparative modes. In the absolute mode, a judge directly assesses the properties of an object and maps them on an existing scale. In the comparative mode, a judge evaluates a target object in relation to another object, which provides a standard for comparison. Neither one of these modes is entirely pure. The ability to make absolute judgments requires past experience with other objects from the same domain, and the ability to make comparative judgments requires some

Author's note
Address correspondence to Joachim I. Krueger, Department of Psychology, Brown University, Box 1853, 89 Waterman St., Providence, RI 02912, USA [email: Joachim_Krueger@Brown.edu]

sense of how each object—the target and the referent—is alike.

The dialectic between absolute and comparative modes of judgment pervades many areas of social perception, including the study of social stereotypes. Social stereotypes are often assessed by having people rate how typical or characteristic various traits are of a target group. By one account, people only care about the prevalence of a trait in the group. The more prevalent it is, the more typical it must be. This hypothesis assumes a simple associative mechanism linking typicality judgments to prevalence estimates. We refer to this idea as the *attribution hypothesis*. By another account, people also care about prevalence of a trait in a comparison group. The larger—and the more positive—the difference in trait prevalence is, the more typical the trait must be. This hypothesis assumes a more complex psychological mechanism. Two prevalence estimates must be generated and compared before trait typicality can be judged. We refer to this account as the *categorization hypothesis*.

Both hypotheses have a long history. Zawadski (1948) and Allport (1954) recognized their theoretical appeal, but did not express a preference. Empirical work initially favored the associationism embodied by the attribution hypothesis (e.g., Brigham, 1971; Mann, 1967), but later emphasized the idea that stereotypes involve perceptual differentiations between groups (e.g., Ford & Stangor, 1992; Judd & Park, 1993; Martell & DeSmet, 2001; McCauley & Stitt, 1978). The way stereotypes were measured yielded corresponding definitions of what stereotypes were. Both types of definition can be found in the literature. Reflecting the attribution point of view, Hilton and von Hippel (1996) define stereotypes as 'beliefs about the characteristics, attributes, and behaviors of members of certain groups' (p. 240). Applying this perspective to gender stereotypes, Eagly, Mladinic, and Otto (1991) suggest that 'the percentage of people in each group (e.g., women) who have each characteristic (e.g., who are "warm") [captures the] association between the group and each characteristic' (p. 208). Conversely, Kunda and Thagard (1996) take the categorization view, suggesting that any 'stereotype is noteworthy only inasmuch as it implies that members of the stereotyped group differ significantly from ordinary people' (p. 297). With regard to gender, Chiu and colleagues suggest that 'knowledge about gender differences [is] organized into a network in a person's long-term memory' (Chiu, Hong, Lam, Fu, Tong, & Lee, 1998, p. 82).

The coexistence of divergent definitions and measurement approaches presents a challenge: how should a suitable definition be chosen among the available alternatives, and does it really matter? One option is to ask which definition is most consistent with the way people form judgments about social groups. Another option is to ask how empirical judgments map onto the underlying processes that give rise to these judgments. The appeal of the attribution hypothesis is its simplicity. This hypothesis assumes that people only need to learn and remember frequency information (Hasher & Zacks, 1984) and to transform this information into probabilistic prevalence estimates (Estes, 1976). These tasks require little attention and can yield high levels of accuracy (Anderson & Schooler, 1991), and this may explain why stereotypes are easily activated (Lambert, Payne, Jacoby, Shaffer, Chasteen, & Khan, 2003). In contrast, the categorization hypothesis requires the learning of probability differentials or the covariation between categories and features (Lewicki, Hill, & Czyzewska, 1992). This type of learning can also occur, but it is more easily disrupted (Stamos-Rossnagel, 2001).

Competitive hypothesis tests are complicated by the fact that the predictions are the same much of the time. Consider the perception that 60% and 40% of people in groups A and B respectively possess a certain trait. Both hypotheses predict that the trait will be seen as typical of group A and atypical of group B. Across traits, it can be expected that typicality ratings are correlated with both the simple prevalence estimates for the target group (e.g., A) and with the differences between the two prevalence estimates (e.g., A-B).

Tests of the attribution and the categorization hypotheses require assessments of their unique

effects. The unique attribution effect is given by the correlation between typicality ratings and prevalence estimates for the target group that control for the differences in the prevalence estimates for the two groups. The unique categorization effect is given by the correlation between typicality ratings and the differences between the two sets of prevalence estimates that control for the prevalence estimates for the target group. In a study of national stereotypes, American and Italian participants made prevalence estimates and typicality ratings for Americans, Italians, English, and Germans. Each of the four stereotypes was characterized by a strong unique attribution effect, whereas the unique categorization effects were negligible (Krueger, 1996). The superior fit of the attribution hypothesis was replicated in a study of gender stereotypes using American and Italian judges (Krueger, Hasman, Acevedo, & Villano, 2003).

Although suggestive, the evidence in favor of the attribution hypothesis has not been decisive (Kanahara, 2006). There are three possible reasons why the categorization hypothesis remains viable: one is conceptual, another is quantitative, and a third is a matter of research design. Consider all three in the context of gender stereotypes: the conceptual reason is that gender stereotypes, roles, and identities are inherently interrelated rather than isolated (Gawronski, Bodenhausen, & Banse, 2005; Goldstone, 1996). What it means to be male presupposes some idea of what it means to be female, and vice versa. The conceptual feature of gender stereotypes inevitably forms a background to all research on gender perception.

The quantitative reason is that when true differences between groups are small, the perceptions of these differences are often exaggerations (Krueger & DiDonato, 2008). Most gender-related differences are rather small (Hyde, 2005), while gender stereotypes suggest numerous perceived differences (Allen, 1995; Diekman, Eagly, & Kulesa, 2002; Martin, 1987; Rosenkrantz, Vogel, Bee, & Broverman, 1968; Spence & Buckner, 2000; but see also Swim, 1994). The accentuation of gender differences reflects a basic principle of categorical reasoning

(Eiser, 1996; Krueger, 1992; Krueger & Clement, 1994; Rothbart, Davis-Stitt, & Hill, 1997). Differences between categories are exaggerated when a salient categorical (i.e., social grouping) variable is correlated with the continuous (i.e., psycho-behavioral) variable being judged (Tajfel, 1969; Wyer, Sadler, & Judd, 2002).

If people perceived no gender differences at all, the categorization hypothesis would have to be false. If, however, people do perceive group differences, they have an opportunity to base their judgments of trait typicality on these differences. By extension, one might conclude that the larger these perceived differences are, the more strongly they are associated with judgments of trait typicality. The meta-contrast principle of self-categorization theory suggests as much (Oakes, Haslam, & Turner, 1994). In contrast, the earlier study on national stereotypes yielded evidence suggesting that the inverse was true: categorization effects—to the degree that they occurred at all—were associated with *smaller* perceived group differences (Krueger, 1996). In the present study, we sought to replicate this finding empirically, and to explicate it analytically.

The third reason for the potential viability of the categorization hypothesis has to do with the way judgments of trait typicality are elicited. In earlier studies that tested the attribution hypothesis and the categorization hypothesis, participants made typicality judgments in an absolute format (e.g., 'how typical is the trait of sensitivity of women?'). Instructions did not refer to relevant comparison groups. No such reference is necessary from the perspective of the attribution hypothesis, which assumes that people automatically look up their own prevalence estimates for the trait, and proceed to make correspondingly high or low typicality ratings. Typicality ratings may, as Schneider (2004, p. 50) suspected, 'simply be another way of asking what percentage of a group has a particular feature'. According to the categorization hypothesis, however, reference to a comparison group may be critical for people to compute differential prevalence estimates. Such computations may be stimulated by instructions that explicitly call for comparative typicality

ratings (e.g., 'how typical is the trait of sensitivity of women compared with men?'). This idea fits with self-categorization theory, which assumes that stereotyping increases inasmuch as intergroup comparisons are salient.

Explicit instructions may be necessary if comparative thinking is more effortful and resource-dependent than associative thinking (Dawes, 2001). People may be perfectly able to make comparative judgments, but they may not feel compelled to do so when presented with absolute rating formats. Their reliance on non-comparative associations may be a satisfactory response in light of conversational norms that require experimental instructions to be informative but not overly detailed (Grice, 1975; Schwarz, 1998).

The main goal of the present research was to revisit the relative strength of the attribution and the categorization hypotheses by using both absolute and comparative response formats. If the strength of the attribution hypothesis in past research reflected the operation of associative and automatic reasoning, evidence for this hypothesis should also be strong when instructions call for comparative typicality ratings. If, however, comparative thinking can be stimulated by proper instructions, the categorization hypothesis should be supported under the appropriate response format.

Comparative ratings have been used in research on gender stereotypes (Hall & Carter, 1999; Spence & Buckner, 2000), and one particularly interesting study employed both formats. After finding that some perceptions of gender differences were weak, or even reversed, when ratings were absolute, Diekman and Eagly (2000) suggested that absolute ratings induce participants to hold each gender to a different standard. An agentic trait, such as assertiveness, may seem equally characteristic of women and men even when women exhibit fewer assertive behaviors. Comparative ratings might overcome such effects of 'shifting standards' (Biernat & Manis, 1994). Consistent with this possibility, perceptions of gender differences reappeared in Diekman and Eagly's fourth study when participants judged gender differences with regard to the target attributes (2000).

Hypotheses

The first hypothesis was that the evidence for the attribution hypothesis would be stronger than the evidence for the categorization hypothesis when typicality ratings are made in an absolute format. Evidence for the categorization hypothesis would emerge when typicality ratings are made in a comparative format. Specifically, an increase in the categorization effect should occur because participants more strongly base their typicality ratings on their percentage estimates for the opposite gender, rather than reducing their reliance on their percentage estimates for the target gender.

The second hypothesis was that participants would strongly differentiate between the two genders. We expected the correlation between the percentage estimates made for the two genders to be low or even negative. The critical question regarding intergroup differentiation was how it might be related to the strength of the categorization effect. Recall that, according to one view, the perception of larger group differences should be associated with a greater use of these perceived differences in the construction of trait typicality ratings. According to empirical precedent, however, the opposite may be true (Krueger, 1996). We addressed this issue both analytically and empirically: analytically, we decomposed the mathematical formula for a difference-score correlation (Cohen & Cohen, 1983) and asked how changes in the individual elements of this formula would affect the categorization effect as a whole. Empirically, we asked the same question using multiple regression analyses to examine the role of individual differences in the constituents of the difference-score correlation.

Method

Participants
Brown University undergraduates participated ($N = 195$, mean age of 19.5 years). Some received course credit, whereas others were recruited from around the campus as volunteers. The data of 11 participants were discarded for being either incomplete or without variance (thus

precluding the computation of correlation coefficients). The final sample consisted of 109 women and 75 men.

Procedure

Participants were presented with a list of 12 personality-descriptive terms from the short version of Bem's Sex Role Inventory (BSRI: Bem, 1981). The list of 12 traits was compiled using the results of a reliability analysis, which was performed on the full list of 10 feminine, 10 masculine, and 10 gender-neutral traits (Krueger et al., 2003). Using the data set from that previous study ($N = 86$), we found that among the feminine traits, the four highest corrected item-total correlations were obtained for the adjectives sensitive (.78), soothing (.84), tender (.76), and warm (.81). Of the masculine traits, the adjectives assertive (.69), dominant (.68), risk-taking (.59), and taking a stand (.65) were the most reliable, and of the gender-neutral traits, the adjectives helpful (.69), likeable (.54), reliable (.60), and sincere (.65) were the most reliable.

Participants generated three responses for each trait item: they estimated the percentage of women who could be described by the trait, the percentage of men who could be described by the trait, and how typical the trait was of one of the gender groups. About half the participants made the typicality ratings in an absolute format, whereas the other half made these ratings in a comparative format, rating 'women, compared with men', or rating 'men, compared with women'. Typicality ratings could range from 1 (not typical at all) to 9 (very typical). The order of the three sets of ratings was varied across participants. After completing the questionnaires, participants were debriefed.

Results

Typicality at the mean level

Before addressing the hypotheses regarding the size of the attribution and categorization effects, we examined average typicality ratings for the two target groups. We needed to establish that ratings reflected the expected pattern of gender-typedness, and to ask whether the response format moderated this basic pattern. To meet these goals, we submitted each of the three sets of traits to a multivariate analysis of variance (MANOVA), in which the format of the scale (absolute vs. comparative) and the gender of the target group were the independent variables, and the ratings of the four traits belonging to a given gender type were the dependent variables. As expected, feminine traits were rated as more typical of women ($M = 6.46$) than of men ($M = 3.99$) ($F(4, 183) = 50.80$, $p < .001$), whereas masculine traits were rated as less typical of women ($M = 4.57$) than of men ($M = 6.83$) ($F(4, 183) = 50.78$, $p < .001$). Gender-neutral traits were rated as being somewhat more typical of women ($M = 5.96$) than of men ($M = 5.12$) ($F(4, 183) = 12.26$, $p < .001$). Univariate analyses replicated these findings for each of the 12 trait items. No other effects were significant.

These preliminary null findings suggest that the use of a comparative response format by itself did not increase perceived gender differences. By the lights of self-categorization theory, one could have expected that a comparative format increases the salience of social categorization and thereby the perceptual differentiation between groups. The lack of an effect of rating format on typicality ratings was further supported by the findings that average typicality ratings obtained with the two formats were highly correlated ($r = .99$ and $.97$, respectively, for the female and male target groups).

Attribution versus categorization

Our first hypothesis was that the attribution effect would be larger than the categorization effect when typicality ratings were absolute. This difference should be reduced or reversed when these ratings were comparative. To repeat, the unique attribution effect was captured by the average idiographic correlation between trait prevalence estimates for the target group and trait typicality ratings for that group, with the differences between prevalence estimates for the two groups being controlled.[1] The unique categorization effect was captured by the correlation between the differences in prevalence

405

estimates and trait typicality ratings, with the prevalence estimates for the target group being controlled.

In a 2 (format: absolute vs. comparative) × 2 (participant gender) × 2 (target gender) × 2 (measure: attribution vs. categorization) mixed-model ANOVA with repeated measures on the last variable, the effect of format was statistically significant ($F(1, 180) = 4.55$, $p = .034$), and so was the predicted interaction between the type of format and the type of measure ($F(1, 180) = 5.50$, $p = .02$) (all other Fs < 1). Consistent with our first hypothesis, the attribution effect was more than twice as large as the categorization effect when ratings were absolute (see Figure 1, left panel) ($F(1, 88) = 11.93$, $p = .001$, d = .38). When compared against the null hypothesis of zero, the categorization effect remained significant, although it was small ($t(90) = 4.77$, $p < .001$).

The expected re-emergence of the categorization effect was also observed: as shown in the right panel of Figure 1, the categorization effect was as strong as the attribution effect when participants made comparative typicality ratings. Simple effects analyses showed that, compared with the effects obtained with absolute

typicality ratings, the categorization effect was stronger ($F(1, 180) = 3.82$, $p = .052$, d = .29), and the attribution effect was weaker ($F(1, 180) = 5.74$, $p = .018$, d = .35).

According to a strict version of our hypothesis, there should have been only an increase in the categorization effect, but not a decrease in the attribution effect. To explore the reasons for the equivalence of the two effects in the comparative condition, we examined the zero-order correlations between prevalence estimates and typicality ratings. We designated the correlation between estimates for the target gender and typicality ratings for the target gender $r(P_t,T)$ and the correlation between estimates for the opposite gender and typicality ratings for the target gender $r(P_o,T)$. A 2 (format) × 2 (typicality ratings for women vs. men) × 2 (percentage estimates for women vs. men) mixed-model ANOVA with repeated measures on the last variable yielded a significant three-way interaction ($F(1, 180) = 4.65$, $p = .032$).[2] Examination of this interaction suggested that instructions to make comparative rather than absolute typicality ratings produced a tendency to associate prevalence estimates for the
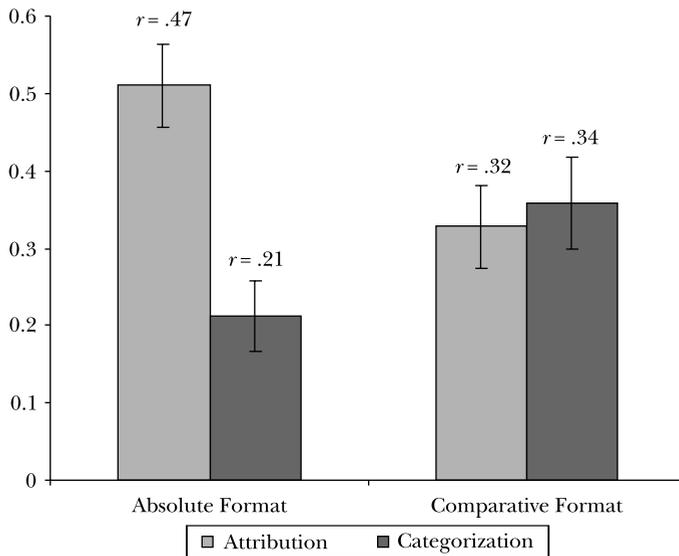

Figure 1. Trait typicality as predicted by prevalence estimates and difference scores: Z-scored partial correlations (with standard error bars).

opposite gender more strongly—and more negatively—with typicality ratings for the target gender (M = −.45 vs. −.39), (F < 1, d =.10), and to associate estimates for the target gender less strongly with these typicality ratings (M = .71 vs. 78), (F(1, 180) = 3.63, p = .058, d = .28). This pattern is weak evidence for the categorization hypothesis, which assumes that only the first would be significant.

Perception of group differences

Our second hypothesis was that participants would perceive strong gender differences. The correlations between prevalence estimates for the two genders, $r(P_t, P_o)$, were indeed negative and of virtually the same size as the effect observed in previous research (Krueger et al., 2003). The lack of any difference between the conditions using a comparative (M = −.30) and an absolute rating format (M = −.32) was consistent with our preliminary analyses of typicality ratings, which had suggested that a change in the rating format per se does not strengthen perceptions of group differences.

The critical question was whether differences in the size of the accentuation effect would predict the strength of the categorization effect. Consider first the analytical approach to this question. Recall that the simple attribution effect (i.e., prior to partialing the categorization effect) is the correlation between prevalence estimates for the target gender and typicality ratings, $r(P_t, T)$. This zero-order correlation should be the only predictor of the attribution effect. In contrast, the simple categorization effect is the correlation between typicality ratings and the differences between estimates for the target gender and estimates for the opposite gender. A difference-score correlation can be recovered from the three underlying zero-order correlations and the three variances (Cohen & Cohen, 1983, p. 416). In the present case,

$$r(P_t - P_o, T) = \frac{s(P_t)r(P_t, T) - s(P_o)r(P_o, T)}{\sqrt{s^2(P_t) + s^2(P_o) - 2s(P_t)s(P_o)r(P_t, P_o)}}$$

This formula allows one to appreciate how the categorization effect would change if any of the constituent correlations were to change. First, confirming the idea that the simple categorization effect is naturally confounded with the simple attribution effect, it is evident that the difference-score correlation increases as $r(P_t, T)$ becomes more positive. Second, confirming the idea that the categorization effect captures the contrast between target group and opposite group, the difference-score correlation increases as $r(P_o, T)$ becomes more negative. Third, the analytical approach confirms an earlier claim that larger perceived group differences are associated with *smaller* categorization effects. The difference-score correlation becomes larger as $r(P_t, P_o)$ becomes *less* negative.

Applying the analytical approach to the *unique* attribution and categorization effects, we find that the attribution effect increases as $r(P_t, T)$ becomes larger and as $r(P_o, T)$ becomes less negative, while accentuation, $r(P_t, P_o)$, has little effect.[3] The categorization effect (formula not shown) increases as $r(P_o, T)$ becomes more negative and as $r(P_t, P_o)$ becomes less negative. Here, changes in $r(P_t, T)$ have little effect. In short, both stereotyping effects respond to changes in two of the three underlying zero-order correlations (see Table 1). Note that the correlation between prevalence estimates for the opposite gender and typicality ratings is the only one that affects both effects, and that it does so in opposite ways. From this, it follows that the two partial correlations are inversely related. As the unique categorization effect increases, the unique attribution effect decreases.

The analytical patterns, which were obtained under the *ceteris paribus* assumption, can serve as baseline hypotheses for the exploration of empirical data. We performed regression analyses across participants, which allowed us to detect how associations specific to the target gender, $r(P_t, T)$, associations specific to the opposite gender, $r(P_o, T)$, and intergroup accentuation effects, $r(P_t, P_o)$, contribute to the attribution effect and to the categorization effect in stereotype judgments. Because the findings were similar regardless of the gender of the participants, the gender of the target group, and the instructional set, we considered only analyses across all participants. First we regressed

Table 1. Predicted and empirical relationships between zero-order correlations and the two unique effects on typicality ratings

Correlation	Attribution		Categorization	
	Prediction	β Weight	Prediction	β Weight
$r(P_t,T)$	+	.65	0	.01
$r(P_o,T)$	+	.62	−	−.88
$r(P_t,P_o)$	0	0	+	.53

the unique attribution effect on the three zero-order correlations. As expected, the size of the unique attribution effect increased as $r(P_t,T)$ became more positive ($\beta = .65$, $p < .001$), and as $r(P_o,T)$ became less negative ($\beta = .62$, $p < .001$). The degree of accentuation, $r(P_t, P_o)$, did not play a role ($\beta = 0$). We then regressed the unique categorization effect on the same predictor variables. Again, as expected, the categorization effect increased as $r(P_o,T)$ became more negative ($\beta = −.88$, $p < .001$), and as $r(P_t, P_o)$ became less negative ($\beta = .53$, $p < .001$). Changes in $r(P_t,T)$ played no role ($\beta = .01$). These findings are also summarized in Table 1.

The final analytic prediction was that the strength of the attribution effect would be negatively related to the strength of the categorization effect. The data showed that this was the case: $r(182) = −.49$, $p < .001$. In sum, the analytical derivations provided a good fit for the empirical data.

Robustness of results

Regarding the robustness of the findings, four points are worth noting. First, the difference scores can be recovered from ratios: $a − b = a(−\frac{b}{a})$, and vice versa: $\frac{a}{b} = 1 + \frac{a-b}{b}$ Computer simulations show that these transformations have little impact on correlational analyses (Krueger et al., 2003). No differences emerged in a past study using both methods (Krueger, 1996), nor did a re-analysis of the present data reveal any. Second, the correlational analyses were idiographic (i.e., within participants and across items) rather than nomothetic (within items and across participants). Theoretically (Kenny & Winquist, 2001) and empirically (Robbins &

Krueger, 2005), these two approaches are equivalent. A re-analysis of the present data confirmed this. Third, the unique effects of attribution and categorization were assessed by partial correlations. Partial regression weights should—and did—yield similar results. We retained the partial correlation approach to preserve comparability of the present data with past research. Fourth, treating participants as the units of analysis and computing all pairwise correlations across traits, we found that typicality ratings were more reliable in the comparative format ($M = .59$) than in the absolute format ($M = .48$), and that difference scores were more reliable ($M = .59$) than their constituent prevalence estimates ($M = .46$). These findings ($sem < .02$) contravene the idea that the relatively modest size of the categorization effect resulted from the unreliability of its measurement.

Discussion

Judd, Park, Yzerbyt, Gordijn, and Muller (2005, p. 678) asserted that 'all beliefs about social groups are comparative'. Such strong claims regarding the power of the categorization effect may have to be reconsidered. The present study confirmed previous findings which suggested that when people rate how typical a trait is of a target group, their judgments can be modeled by the attribution hypothesis. Probabilistic estimates of trait prevalence in the target group predict typicality judgments well, whereas estimates for a comparison group are virtually irrelevant. We speculated that the failure of the categorization hypothesis could stem, in part, from typicality judgments being cast as absolute rather than comparative judgments. When the

instructions for typicality judgments matched the categorization hypothesis by explicitly calling attention to the comparison group, the categorization hypothesis contributed as much to the modeling of stereotype judgments as the attribution hypothesis did.

Taken together, these findings suggest that stereotype judgments are both parsimonious and sensitive to the logic of conversation. Being more laborious than associative thinking, comparative thinking is not engaged unless it is specifically elicited. Surprisingly, participants achieved the equal strength of the attribution and the categorization effect not exclusively through an increase in the categorization effect, but partly through a reduction in the attribution effect. Instead of merely giving greater weight to their prevalence estimates for the opposite gender, they also gave less weight to their estimates for the target gender.

The present findings fit well into a larger emerging theme. The validity of comparative self-judgments, for example 'how happy am I compared with the average person?', was long taken for granted until componential analyses showed that people heavily rely on absolute self-judgments and virtually ignore their own absolute judgments of the average person (Chambers & Windschitl, 2004; Moore & Small, 2008). This associationist pattern arises in part from egocentric weighting of self-referent information, and in part from attention being focused on the self. Our findings show that a focalist bias can emerge when self-judgments play no role.

Research on implicit prejudice also tends to conceptualize and measure biases against certain social groups in comparative terms. The popular implicit association test (Greenwald, McGhee, & Schwartz, 1998) yields an index of bias that is a difference between two differences (Krueger, 2008). For example, reaction times elicited from White participants show how rapidly they associate White stimuli with positive stimuli, White stimuli with negative stimuli, Black stimuli with positive stimuli, and Black stimuli with negative stimuli. Blanton, Jaccard, Gonzales, and

Christie (2006) found that only the last type of association predicted participants' scores on a self-report measure of prejudice.

The other major result of this study was that although participants perceived women and men differently, the degree to which they did so was negatively related to the strength of the categorization effect. In other words, a stronger accentuation effect was associated with a weaker categorization effect. This finding may be intuitively surprising and at variance with the meta-contrast principle of self-categorization theory. Yet this conflict is easily resolved. Recall that our analysis was focused on the unique effects of the three zero-order correlations. This strategy of exploring unique effects paralleled our approach to the analysis of unconfounded attribution and categorization effects. Statistically, however, the three correlations systematically constrain one another. If $r(P_t,T)$ and $r(P_o,T)$ are respectively positive and negative, it is likely that $r(P_t,P_o)$ is negative. Without controlling $r(P_t,T)$ and $r(P_o,T)$, it will thus appear that stronger inter-group accentuation effects are associated with stronger categorization effects.

Rationality

We now consider a theoretical and a methodological implication of the present research. The theoretical implication is concerned with the rationality of social judgment. In its pure form, the categorization hypothesis exemplifies the ideal of rational judgment, which is the construction of a belief system that is free from internal contradictions (Dawes, 2001). Such coherence can be attained if typicality ratings for a target group are equally related, though with different signs, to prevalence estimates for the target group and to prevalence estimates for the comparison group. If this standard is achieved, the difference-score representing the categorization effect is maximized, and judgments of typicality amount to judgments of diagnosticity: they allow the categorization of a person into a social group on the basis of a known attribute. Statistically, such coherence means that prevalence estimates for both groups tap into the

409

same underlying construct (i.e., diagnosticity), which legitimates the use of difference scores (Blanton et al., 2006).

As the correlation between prevalence estimates for the opposite gender and typicality ratings for the target gender becomes less extreme, the difference-score correlation becomes smaller. Statistically, the difference score becomes confounded because it integrates variables that measure different constructs (Johns, 1981). Ultimately, the unique categorization effect disappears. If, for example, the correlation between prevalence estimates for the target gender and typicality ratings is .5 and the correlation between prevalence estimates for the opposite gender and typicality is 0, the difference-score correlation is still positive, but the unique categorization effect is nil.

The empirical fragility of the categorization effect marks a difficulty inherent in comparative ratings. Unlike absolute ratings, which can be made on the basis of simple associations, comparative ratings require the kind of effortful cogitation that is characteristic of rational judgment. The distinction between associative processes underlying the attribution effect and reflective, comparative processes underlying the categorization effect evokes distinctions made by two-systems theories of social judgment (Strack & Deutsch, 2004).[4] In many areas of social life, the associative systems yield judgments of reasonable accuracy without consuming precious mental resources (Gigerenzer, 2006; Willis & Todorov, 2006). When social stereotyping is understood as such an ordinary form of heuristic reasoning, its adaptive benefits can be documented (Macrae, Milne, & Bodenhausen, 1994).

Item selection

The methodological implication of this work is concerned with the types of group characteristics chosen for the study of stereotyping. In many areas of social judgment research, trait-descriptive adjectives are the common grist for the data-analytic mill. Usually, trait adjectives are selected for their ability to capture individual and group differences. Consider three different types of trait: one type consists of features that all (or no) members of all groups share, such as

being mortal; another type consists of features that all members of one group share, but no member of the other, such as having two X chromosomes; yet another consists of features that some members of one group share, but all members or no member of the other share, such as having a caesarian section.

Traits of the first type (e.g., mortality) do not involve variation across individuals or groups, and thus yield no insights into processes of attribution or categorization. Traits of the second type (e.g., having two X chromosomes) perfectly confound attribution with categorization, and thus offer no statistical leverage for competitive hypothesis testing. Traits of the third type are more intriguing. According to the attribution hypothesis, having had a caesarian is not typical of women because most have had none. According to the categorization hypothesis, having had a caesarian is highly typical of women because it yields a positive difference score. If only items of this type were presented, or if statistical analyses were nomothetic (i.e., for a single item), the attribution and categorization effects would be perfectly confounded. All difference scores would be equal to the prevalence estimates for the target group, and hypothesis tests would be reduced to investigators' intuitive judgments of whether mean typicality rating for such items can be considered high or low. The constraints imposed by certain types of judgment item suggest that it is reasonable to use personality-descriptive terms for stereotype assessment. Such items permit both individual and group differences across the entire percentage scale.

Conclusion

The idea that one core function of social stereotypes is to simplify social perception has enjoyed broad acceptance since the days of Lippmann (1922) and Allport (1954). The present research suggests that people rely on simple associations between trait prevalence and trait typicality as a default when constructing mental images of groups. When prompted, however, they can also call upon perceived group differences to construct more complex images. We believe that the empirical strength of the attribution effect

410

may not only be a matter of mental parsimony and resource conservation. As we noted at the outset, all trait judgments involve implicit comparisons. Inasmuch as social stereotypes are not mere conglomerations of unrelated attributes, prevalence estimates for a set of traits within the same group tend to cohere into a theme or *Gestalt* (Rehder & Hastie, 2001; Wittenbrink, Gist, & Hilton, 1997). As people perceive inter-item correlations within a group, judgments about any trait serve as cues for other traits. Here lies a final advantage of parsimony: it is easier to infer that trait Y has a high prevalence in the group if the prevalence of trait X is also high, than it is to infer that trait Y positively differentiates group A from group B if this is what trait X does.

Notes

1. Correlations involving difference scores and a female target gender were multiplied with –1 so that they could be aggregated with the corresponding correlations involving the male target gender.
2. Participant gender was not included in this analysis.
3. The partial correlation representing the unique attribution effect is

$$r(P_t, T.P_t - P_o) = \frac{r(P_t, T) - r(P_t, P_t - P_o)r(T, P_t - P_o)}{\sqrt{(1 - r[P_t, P_t - P_o]^2)(1 - r[T, P_t - p_o]^2)}},$$

where the term $r(P_t, P_t - P_o)$ can be expressed as

$$\frac{s_o - s_t r(P_t, P_o)}{\sqrt{s_t^2 + s_o^2 - 2s(P_t)s(P_o)r(p_t, P_o)}} \text{ (McNemar, 1969,}$$

p. 177), and the term $r(T, P_t - P_o)$ can be expressed by the standard formula for difference score correlations.
4. Eiser (2003) recently presented a connectionist model and computer simulations to argue that inter-group comparisons and accentuation can arise from a set of associations being made unconsciously and in parallel—that is, without the burden of added effort.

Acknowledgments

References

Allen, B. P. (1995). Gender stereotypes are not accurate: A replication of Martin (1987) using diagnostic vs. self-report and behavioral criteria. *Sex Roles, 32,* 583–600.

Allport, G. W. (1954). *The nature of prejudice.* Garden City, NY: Doubleday/Anchor.

Anderson, J. R., & Schooler, L. J. (1991). Reflections on the environment in memory. *Psychological Science, 2,* 396–408.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review, 88*, 354–364.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, *66*, 5–20.

Blanton, H., Jaccard, J., Gonzales, P. M., & Christie, C. (2006). Decoding the implicit association test: Implications for criterion prediction. *Journal of Experimental Social Psychology, 42,* 192–212.

Brigham, J. C. (1971). Ethnic stereotypes. *Psychological Bulletin*, *76*, 15–38.

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: The role of nonmotivated factors in above-average and comparative-optimism effects. *Psychological Bulletin, 130*, 813–838.

Chiu, C., Hong, Y., Lam, I. C., Fu, J. H., Tong, J. Y., & Lee, V. S. (1998). Stereotyping and self-presentation: Effects of gender stereotype activation. *Group Processes & Intergroup Relations, 1*, 81–96.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Dawes, R. M. (2001). *Everyday irrationality.* Boulder, CO: Westview.

Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin, 26*, 1171–1188.

Diekman, A. B., Eagly, A. H., & Kulesa, P. (2002). Accuracy and bias in stereotypes about the social and political attitudes of women and men. *Journal of Experimental Social Psychology*, *38*, 268–282.

Eagly, A. H, Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men?

An analysis of attitudes, beliefs, and emotions. *Psychology of Women Quarterly, 15,* 203–216.

Eiser, J. R. (1996). Accentuation revisited. In W. P. Robinson (Ed.), *Social groups and identities. Developing the legacy of Henri Tajfel* (pp. 121–142). Oxford: Butterworth-Heinemann.

Eiser, J. R. (2003). The accentuation principle in social judgment: A connectionist reappraisal. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 486–511). New York: Cambridge University Press.

Estes, W. K. (1976). The cognitive side of probability learning. *Psychological Review, 83,* 37–64.

Ford, T. E., & Stangor, C. (1992). The role of diagnosticity in stereotype formation: Perceiving group means and variances. *Journal of Personality and Social Psychology, 63,* 356–367.

Gawronski, B., Bodenhausen, G. V., & Banse, R. (2005). We are, therefore they aren't: Ingroup construal as a standard of comparison for outgroup judgments. *Journal of Experimental Social Psychology, 41,* 515–526.

Gigerenzer, G. (2006). Bounded and rational. In R. J. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 115–133). Oxford: Blackwell.

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory and Cognition, 24,* 608–628.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.)*, Syntax and semantics: 3 speech acts* (pp. 41–58). New York, NY: Academic Press.

Hall, J. A., & Carter, J. D. (1999). Gender-stereotype accuracy as an individual difference. *Journal of Personality and Social Psychology, 77,* 350–359.

Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information. *American Psychologist, 39,* 1372–1388.

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology, 47,* 237–271.

Hyde, J. S. (2005). The gender similarity hypothesis. *American Psychologist, 60,* 581–592.

Johns, G. (1981). Difference score measures of organizational behavior variables: A critique. *Organizational Behavior and Human Performance, 27,* 443–463.

Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review, 100,* 109–128.

Judd, C. M., Park, B., Yzerbyt, V., Gordijn, E. H., & Muller, D. (2005). Attributions of intergroup bias and outgroup homogeneity to ingroup and outroup others. *European Journal of Social Psychology, 35,* 677–704.

Kanahara, S. (2006). A review of the definitions of stereotype and a proposal for a progressional model. *Individual Differences Research, 4,* 306–321.

Kenny, D. A., & Winquist, L. (2001). The measurement of interpersonal sensitivity: Consideration of design, components, and unit of analysis. In F. J. Bernieri & J. A. Hall (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 265–302). Mahwah, NJ: Erlbaum.

Krueger, J. I. (1992). On the overestimation of between-group differences. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 3, pp. 31–56). Chichester, UK: Wiley.

Krueger, J. I. (1996). Probabilistic national stereotypes. *European Journal of Social Psychology, 26,* 961–980.

Krueger, J. I. (2008). The robust beauty of simple associations. In J. I. Krueger (Ed.), *Rationality and social responsibility: Essays in honor of Robyn M. Dawes* (pp. 111–140). New York, NY: Psychology Press.

Krueger, J. I., & Clement, R. W. (1994). Memory-based judgments about multiple categories: A revision and extension of Tajfel's accentuation theory. *Journal of Personality and Social Psychology, 67,* 35–47.

Krueger, J. I., & DiDonato, T. E. (2008). Social categorization and the perception of groups and group differences. *Social and Personality Psychology Compass: Group Processes, 2,* 733–750.

Krueger, J. I., Hasman, J. F., Acevedo, M., & Villano, P. (2003). Perceptions of trait typicality in gender stereotypes: Examining the role of attribution and categorization processes. *Personality and Social Psychology Bulletin, 29,* 108–116.

Kunda, A., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103,* 284–308.

Lambert, A. J., Payne, B. K., Jacoby, L. L., Shaffer, L. M., Chasteen, A. L., & Khan, S. R. (2003). Stereotypes as dominant responses: On the 'social facilitation' of prejudice in anticipated

public contexts. *Journal of Personality and Social Psychology, 84,* 277–295.

Lewicki, P., Hill, T., & Czyzewska, M. (1992). Nonconscious acquisition of information. *American Psychologist, 47,* 796–801.

Lippmann, W. (1922). *Public opinion.* New York, NY: Harcourt, Brace.

McCauley, C., & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology, 36,* 929–940.

McNemar, Q. (1969). *Psychological statistics* (4th ed.). New York, NY: Wiley.

Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66,* 37–47.

Mann, J. W. (1967). Inconsistent thinking about group and individual. *Journal of Social Psychology, 71,* 235–245.

Martell, R. F., & DeSmet, A. (2001). A diagnostic-ratio approach to measuring beliefs about the leadership abilities of male and female managers. *Journal of Applied Psychology, 86,* 1223–1231.

Martin, C. L. (1987). A ratio measure of sex stereotyping. *Journal of Personality and Social Psychology, 52,* 489–499.

Moore, D., & Small, D. (2008). When it is rational for the majority to believe that they are better than average. In J. I. Krueger (Ed.), *Rationality and social responsibility: Essays in honor of Robyn M. Dawes* (pp. 141–174). New York, NY: Psychology Press.

Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality.* Oxford: Blackwell.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General, 130,* 323–360.

Robbins, J. M., & Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review, 9,* 32–47.

Rosenkrantz, P., Vogel, S., Bee, H., & Broverman, I. (1968). Sex-role stereotypes and self-concepts in college students. *Journal of Clinical and Consulting Psychology, 32,* 287–295.

Rothbart, M., Davis-Stitt, C., & Hill, J. (1997). Effects of arbitrarily placed category boundaries on similarity judgments. *Journal of Experimental Social Psychology, 33,* 122–145.

Schneider, D. J. (2004). *The psychology of stereotyping.* New York, NY: Guilford.

Schwarz, N. (1998). Communication in standardized research situations: A Gricean perspective. In S. Fussell & R. Kreuz (Eds.), *Social and Cognitive Approaches to Interpersonal Communication* (pp. 39–66). Mahwah, NJ: Erlbaum.

Spence, J. T., & Buckner, C. E. (2000). Instrumental and expressive traits, trait stereotypes, and sexist attitudes. *Psychology of Women Quarterly, 24,* 44–62.

Stamos-Rossnagel, C. (2001). Revealing hidden covariation detection: Evidence for implicit abstraction at study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1276–1288.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247.

Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology, 66,* 21–36.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25,* 149–178.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100-ms exposure to a face. *Psychological Science, 17,* 592–598.

Wittenbrink, B., Gist, P. L., & Hilton, J. L. (1997). Structural properties of stereotypic knowledge and their influences on the construal of social situations. *Journal of Personality and Social Psychology, 72,* 526–543.

Wyer, N. A., Sadler, M. S., & Judd, C. M. (2002). Contrast effects in stereotype formation and change: The role of comparative context. *Journal of Experimental Social Psychology, 38,* 443–458.

Zawadski, B. (1948). Limitations of the scapegoat theory of prejudice. *Journal of Abnormal and Social Psychology, 43,* 127–141.

Biographical notes

JOACHIM I. KRUEGER is professor of psychology at Brown University. He received his PhD at the University of Oregon in 1988. His research is focused on inductive reasoning in social context: specific topics include self-perception and intergroup relations. He edited a volume, with M. Alicke and D. Dunning, on *The self in social*

judgment (Psychology Press, 2005), and a festschrift for Robyn Dawes: *Rationality and social desirability* (Psychology Press, 2008).

JULIE H. HALL earned her PhD in clinical psychology from the University at Buffalo in 2007. She is the author of several book chapters and journal articles. Her research interests include infidelity, forgiveness within romantic relationships, and self-forgiveness.

PAOLA VILLANO is associate professor of social psychology at the University of Bologna, Italy. She earned her PhD here in 1996. Her research interests focus on gender stereotypes, ethnic prejudice, social dominance orientation, and conflict and mediation. Her publications include a book on prejudice and stereotype (*Pregiudizi e stereotipi*, Roma, Carocci, 2003) and social dominance orientation and prejudice in an Italian sample, Psychological Reports, 101 (2007).

MEREDITH C. JONES is a student in the PhD program in child clinical psychology at the University of Denver, where she studies adolescent romantic relationships, attachment, and sexual behavior. She received her BA in psychology and anthropology from Brown University.