

Running Head: IN-GROUP AND OUT-GROUP HOMOGENEITY

This self-archived version is provided for scholarly purposes only. The correct reference for this article is as follows:

Rubin, M., & Badea, C. (2007). Why do people perceive in-group homogeneity on in-group traits and out-group homogeneity on out-group traits? *Personality and Social Psychology Bulletin*, 33, 31-42. doi: [10.1177/0146167206293190](https://doi.org/10.1177/0146167206293190)

We are grateful to Sandrine Arp for her assistance with data collection. We are also grateful to Rob Goldstone, Dan Navarro, and Vladimir Sloutsky for their advice regarding Tversky's (1977) diagnosticity hypothesis. The research reported in this article was presented at the 14th General Meeting of the European Association of Experimental Social Psychology, Würzburg, Germany. Correspondence concerning this article should be addressed to Mark Rubin at the School of Psychology, University of Newcastle, Callaghan, NSW 2308, Australia. Tel: +61 (0)2 4921 6706. Fax: +61 (0)2 4921 6980. E-mail: mark.rubin@newcastle.edu.au

Word count: 9,322

Abstract

People tend to perceive in-group homogeneity on in-group stereotypical traits and out-group homogeneity on out-group stereotypical traits (e.g., Kelly, C. A., 1989; Simon, B., 1992a; Simon, B., & Pettigrew, T. F., 1990). If it is assumed that people use homogeneity ratings to indicate the extent to which groups possess traits, then this *stereotype effect* may be interpreted as an expression of perceived trait possession (i.e., in-groups possess in-group stereotypical traits and out-groups possess out-group stereotypical traits). If it is further assumed that research participants abide by the conversational norm of appropriate quantity (e.g., Bless, H., Strack, F., & Schwarz, N., 1993), then this stereotype effect should be significantly reduced following prior expressions of perceived trait possession. A literature review and two minimal group experiments ($Ns = 75$ & 104) supported this prediction. This evidence is discussed in relation to the out-group homogeneity effect and self-categorization theory.

Key words: in-group homogeneity, out-group homogeneity, intragroup variability, stereotype, social identity, minimal group.

Why do People Perceive In-Group Homogeneity on In-Group Traits and Out-Group Homogeneity on Out-Group Traits?

The out-group homogeneity (OH) effect is the tendency to perceive members of one's own social group (in-group members) as being more variable than members of groups to which one does not belong (out-group members). The OH effect is a small but robust effect (Mullen & Hu, 1989) that has been identified across a wide variety of different social groups, including national groups (S. A. Haslam, Oakes, Turner, & McGarty, 1995), religious groups (Hewstone, Islam, & Judd, 1993), political groups (Kelly, 1989), age groups (Linville, Fischer, & Salovey, 1989, Study 1), and even artificial laboratory-based groups (e.g., Rubin, Hewstone, & Voci, 2001, Study 1; for reviews, see Devos, Comby, & Deschamps, 1996; Linville, 1998; Mullen & Hu, 1989; Ostrom & Sedikides, 1992; Voci, 2000).

The OH effect is a theoretically important effect. Differences in the perceived variability of social groups can cause differences in the extent of stereotyping (Lambert, 1995; Lambert & Wyer, 1990; Park & Hastie, 1987) as well as differences in the confidence ascribed to stereotype-based judgments (Ryan, Judd, & Park, 1996; Ryan, Robinson, & Hausmann, 2001). Hence, a better understanding of the OH effect should lead to a better understanding of stereotyping in general.

One approach to understanding the OH effect has been to identify the limits of its occurrence. As Simon (1992b) put it, the OH effect is by no means a "universal law" (p. 4). Researchers have identified a long list of moderators that reduce, nullify, or even reverse the OH effect, including group status (e.g., Boldry & Kashy, 1999), group size (e.g., Simon & Brown, 1987), group power (e.g., Guinote, Judd, & Brauer, 2002), group formation (Brown & Wootton-Millward, 1993), in-group threat (Rothgerber, 1997), in-group identification (e.g., De Cremer, 2001), group membership salience (e.g., Lee & Ottati, 1995), inter- vs intra-group rating context (e.g., S. A. Haslam et al., 1995, 1996), and familiarity with out-group members (e.g., Linville, Salovey, & Fischer, 1986).

Adding to this list of moderators, Simon and Pettigrew (1990) proposed that the OH effect varies as a function of the stereotypicality of the traits on which judgments of intragroup variability are made. Specifically, they proposed that the OH effect should occur only on traits that are stereotypical of the out-group, and that a reverse OH effect – an in-group homogeneity (IH) effect – should occur on traits that are stereotypical of the in-group (for similar predictions, see Brewer, 1993). There is some evidence supporting the existence of this *stereotype effect* (Brown & Wootton Millward, 1993; Castano & Yzerbyt, 1998; Kelly 1989; Pickett & Brewer, 2001; Simon, 1992a; Simon & Pettigrew, 1990).

Previous Explanations of the Stereotype Effect

Simon (1992a, 1992b) offered two explanations for the stereotype effect. The first explanation is based on Tversky's (1977) diagnosticity principle, and the second explanation is based on Tajfel and Turner's (1979) social identity theory.

According to Tversky's (1977) diagnosticity principle, stimuli appear more similar to one another when they share attributes that are diagnostic of intercategory distinctions. This diagnosticity effect should increase perceived intragroup similarity when ratings are made on stereotypical traits because group members share stereotypical traits and stereotypical traits are diagnostic of intergroup distinctions.

According to the social identity explanation, because group members are "assumed to strive for a positive social identity...[they] should therefore emphasize relative ingroup homogeneity on typical ingroup attributes" (Simon, 1992b, p. 13). However, there are reasons to question whether this social identity explanation can provide a complete account of the stereotype effect. The perception of IH on in-group traits should only contribute to a *positive* social identity if the traits are perceived to be *positive*. Perceiving IH on in-group *negative* traits should result in a *negative* social identity and should therefore be avoided.

Likewise, the need for a positive social identity should motivate the perception of OH only on out-group *negative* traits, not on out-group *positive* traits (for discussions, see Kelly, 1989, p. 248; Voci, 2001, pp. 185-186). To be clear, we accept that trait valence acts as a significant moderator of perceived group homogeneity (e.g., S. A. Haslam et al., 1995, 1996; Simon, Glässner Bayerl, & Stratenwerth, 1991; Simon, 1992b, p. 25). However, because the stereotype effect occurs *independent* of trait valence, it cannot be explained in terms of a search for *positive* social identity.

A Stereotype Expression Explanation of the Stereotype Effect

To add to these two accounts, we tested a stereotype expression explanation of the stereotype effect that is based on the assumption that research participants adopt a “homogeneity equals possession” heuristic when making ratings of intragroup variability. We assumed that one of the reasons that research participants might indicate that a group is relatively homogeneous with respect to a particular trait is to indicate that the group *possesses* that trait (see also S. A. Haslam et al., 1995, p. 154). According to this “homogeneity equals possession” heuristic, the perception that one group possesses a trait to a greater extent than another group can be conveyed by rating the members of the first group as being more homogeneous than the members of the second group.

Based on the assumption of a “homogeneity equals possession” heuristic, we interpreted the stereotype effect as an act of stereotype expression. If research participants presume that “homogeneity equals possession”, then the stereotype effect may be explained as an expression of perceived differences in the extent to which groups possess stereotypical and counterstereotypical traits. Specifically, participants may rate the in-group as more homogeneous on in-group traits in order to express their perception that the in-group possesses these stereotypical traits. Likewise, participants may rate the out-group as more homogeneous on out-group traits in order to express their perception that the out-group possesses these stereotypical traits.

It is instructive to compare and contrast the social identity, diagnosticity, and stereotype expression explanations. The social identity and diagnosticity explanations treat the stereotype effect as a *biased* perception of *group homogeneity* caused by *cognitive and/or motivational factors*. In contrast, our stereotype expression explanation treats the stereotype effect as an *accurate* representation of socially shared perceptions of *trait possession*. Also note that, like the diagnosticity explanation but unlike the social identity explanation, our stereotype expression explanation does not depend on perceivers’ group membership: Both target group members *and* nonaffiliated observers should indicate greater group homogeneity on stereotypical traits than on counterstereotypical traits.

It is also important to note that the social identity, diagnosticity, and stereotype expression explanations are not mutually exclusive. In other words, it is possible that trait diagnosticity, the need for a positive social identity, and stereotype expression all contribute to the stereotype effect, at least in some circumstances. However, given that the social identity explanation cannot easily account for the stereotype effect on either in-group negative or out-group positive traits, we suggest that social identity concerns are unlikely to play a large role in producing the stereotype effect on these traits. In contrast, our stereotype expression explanation can account for the stereotype effect regardless of trait valence.

The Relationship Between Group Homogeneity and Trait Possession

Our stereotype expression explanation is based on the assumption that participants rate groups as more homogeneous on stereotypical traits than on counterstereotypical traits in order to indicate that groups possess stereotypical traits to a greater extent. Below, we discuss three reasons why people might equate group homogeneity with trait possession in this way.

First, homogeneous groups are more likely to be perceived as coherent groups rather than as incidental aggregates of individuals (for a discussion, see Hamilton, Sherman, &

Lickel, 1998). This perception of group *entitativity* (Campbell, 1958) is usually accompanied by the perception that the group possesses an underlying essence (e.g., N. Haslam, Rothschild, & Ernst, 2000; for a review, see Yzerbyt, Corneille, & Estrada, 2001). The perception of a group essence would seem to be necessary in order for people to attribute traits to a group as a holistic entity, rather than to individuals within a group (Plaks, Levy, Dweck, & Stroessner, 2004, p. 141). Hence, increasing a group's perceived homogeneity increases the perception of a group essence to which traits may be attributed. Consistent with this assumption, Levy, Stroessner, and Dweck (1998, Experiment 3) found that people who were more disposed towards characterising groups in terms of traits were also more likely to perceive greater similarity between group members. In addition, Bastian and N. Haslam (2006) found that individual differences in essentialist beliefs correlated positively with stereotype endorsement in relation to a variety of social categories.

Second, perceiving greater homogeneity within the in-group and/or the out-group on a particular trait enhances the perceived reliability of the distinction between the two groups with respect to that trait (Doosje, Spears, & Koomen, 1995; Ford & Stangor, 1992). In statistical terms, the perception of relatively small within-group variances increases the significance of the difference between group means (Doosje, Ellemers, & Spears, 1999; Ford & Stangor, 1992). Hence, increasing a group's perceived homogeneity makes intergroup differences in trait possession appear more reliable.

Third, Park and Hastie (1987) found that perceived homogeneity influences the extent to which people generalize traits from one member of a group to other members of that group. The more homogeneous a group, the greater the generalization of stereotypical traits. Hence, increasing a group's perceived homogeneity on a particular trait makes individual group members appear more integrated in possessing that trait (see also Ford & Stangor, 1992; Ryan, Judd et al., 1996; Ryan, Robinson, et al., 2001).

In summary, there is evidence that increasing the perceived homogeneity of a group with respect to a particular trait is liable to (a) reify the group's essence in relation to the trait, (b) increase the perceived reliability with which the group's possesses the trait, and (c) increase the generalizability of the trait to the group's members. In short, people perceive homogeneous groups to be more real, more reliable, and more integrated in terms of their trait possession.

We assumed that groups that are perceived to be real, reliable, and integrated in terms of their trait possession would be perceived as possessing traits to a greater extent than groups that are perceived to be less real, reliable, and integrated in their trait possession. Consequently, there should be a positive relationship between a group's perceived homogeneity on a given trait and the extent to which that group is perceived to possess the trait. This positive relationship between group homogeneity and trait possession opens up the possibility of a stereotype expression explanation for the stereotype effect: People show IH on in-group traits and OH on out-group traits in order to express their perception that the in-group possesses in-group traits and the out-group possesses out-group traits.

Experimental Rationale: A Conversational Approach

We adopted a conversational approach (e.g., Grice, 1975; Hilton, 1995) in order to test our stereotype expression explanation. This approach is based on Grice's (1975) proposal that people tend to use a set of norms to infer meaning during their conversations with others. Specifically, if communicative partners are co-operative with one another, then they should only provide one another with an appropriate *quantity* of information that is *true*, *relevant*, and *unambiguous*.

As Hilton (1995) observed "all experiments and surveys are forms of social interaction between the experimenter and participant, which invariably involves communication through ordinary language" (p. 249). Following this interpretation, several

researchers have used Grice's (1975) conversational principles to understand how participants interpret research situations (e.g., Bless, Strack, & Schwarz, 1993; Hilton, 1995; Schwarz, 1996; for a review, see Schwarz, 1999). Of particular relevance here, Bless et al. (1993) have considered the influence that the norm of providing appropriate *quantities* of information has on responses in multi-item questionnaires.

According to Bless et al. (1993), research participants are liable to assume that the researcher is a co-operative communicator and that, consequently, the researcher's questions are designed to elicit suitable quantities of information at each stage of the research. Furthermore, co-operative participants are obliged to respond to these questions with suitable quantities of information. In particular, co-operative participants should not provide the researcher with information that they have already provided in response to previous questions. In other words, co-operative participants are obliged to provide *new* information to the researcher on each occasion that they are asked the same or similar question. Clark and Haviland (1977) labelled this conversational obligation the *given-new contract* (for evidence, see Strack, Martin, & Schwarz, 1988). To illustrate, imagine that a researcher asks her participants to describe their "weekend activities" and then, during a later part of the research, to describe their "weekly activities". In this scenario, participants are likely to interpret "weekly" as meaning activities that are carried out from Monday to Friday, rather than from Monday to Sunday, because they will assume that the researcher is a co-operative communicator who is soliciting new and nonredundant information at each stage of her research. Consequently, participants are liable to exclude information about their weekend activities when responding to the "weekly activities" item in order to conform to the conversational norm of quantity.

We assumed the operation of a given-new contract in our research in order to test our stereotype expression explanation. If the stereotype effect is attributable to stereotype expression, then it should be nullified if participants engage in an alternative form of stereotype expression prior to making variability ratings. In other words, the stereotype effect should not occur if participants have already provided the researcher with information about the extent to which the in-group possesses in-group traits and the out-group possesses out-group traits. To be clear, we are proposing that stereotype expression causes the stereotype effect and that the given-new contract is a moderator of stereotype expression: If participants have already expressed a stereotype once to the researcher on one measure (trait ratings), then they would contravene the conversational norm of quantity if they were to express this stereotype again on a second measure (variability ratings).

Literature Review

We began our research by conducting a review of the studies in which researchers had measured perceived in-group and out-group variability on in-group and out-group traits.¹ Consistent with our stereotype expression explanation, we only expected the stereotype effect to occur in studies in which measures of perceived intragroup variability were not preceded by measures that allowed the expression of trait stereotypicality. The results of our literature review are displayed in Table 1.

----- Insert Table 1 here -----

Looking at Table 1, it can be seen that the stereotype effect was only evident in the seven studies in which measures of perceived intragroup variability were not preceded by measures of perceived trait possession (Brown & Wootton-Millward, 1993; Castano & Yzerbyt, 1998, Studies 1 & 2; Kelly 1989; Pickett & Brewer, 2001; Simon, 1992a; Simon & Pettigrew, 1990). No significant stereotype effect was found in any of the seven studies in which measures of perceived intragroup variability were preceded by measures of perceived trait possession (Hopkins & Cable, 2001; Judd, Park, Ryan, Brauer & Kraus, 1995, Studies 1 & 2; Rubin et al., 2001, Studies 1, 2, & 3; Simon & Hamilton, 1994, Study 1). This pattern of

evidence is consistent with our view that the stereotype effect is a form of stereotype expression.

Interestingly, our literature review also suggested that the stereotype effect is moderated by the *type* of variability measure that researchers have used. The stereotype effect has been found most often using the similarity and distribution measures (SD & P_d). Only two of the eight studies that have used the range measure has found evidence of the stereotype effect (Brown & Wootton-Millward, 1993; Kelly, 1989). The relative scarcity of the stereotype effect on the range measure suggests that there may be something about this measure that precludes the effect. It is useful to consider operational details of the range measure in order to investigate this discrepancy.

In typical operationalizations of the range measure, participants are asked to indicate where the average group member falls on a particular trait – the *mean rating* – before they indicate where the two most extreme group members fall on that trait – the *range rating*. Consequently, the range measure usually provides participants with the opportunity for expressing trait possession via mean ratings before they make their judgments of intragroup variability. Consistent with our stereotype expression explanation, this order of events may preclude the stereotype effect: The given-new contract obliges participants to provide new information in their range ratings compared with the information that they have already provided in their mean ratings.

Based on the above analysis, we made the additional hypothesis that the stereotype effect would be found on the range measure provided that there was no prior opportunity for stereotype expression, including the presence of mean ratings. Consistent with this hypothesis, the only studies to have found the stereotype effect using the range measure did not precede range ratings with mean ratings (Brown & Wootton-Millward, 1993; Kelly, 1989). In order to test our hypothesis, we used several different measures of perceived intragroup variability in our research, including the range measure, and we included variability measure as a factor in our analyses.

Overview of the Research

We predicted that the stereotype effect would be nullified if stereotype expression preceded variability judgments. In order to test this hypothesis, we conducted two experiments based on the *minimal group paradigm* (e.g., Tajfel, Billig, Bundy, & Flament, 1971). In these experiments, we randomly assigned participants to two new and unfamiliar social groups on the basis of bogus feedback about their performance on an alleged test of personality. This procedure allowed us to exclude objective differences in intragroup variability from our analyses (for discussions, see Brown & Wootton-Millward, 1993; Guinote et al., 2002; Rubin, Hewstone, Crisp, Voci, & Richards, 2004).

In our first experiment, we manipulated the opportunity for stereotype expression in a relatively subtle manner by including or excluding mean ratings prior to range and similarity measures. In our second experiment, we manipulated the opportunity for stereotype expression more directly by including or excluding a separate measure of perceived trait typicality prior to range, similarity, and distribution measures.

Experiment 1

Method

Participants and design. We used a 2 (mean ratings: absent/present) x 2 (variability measure: range/similarity) x 2 (trait stereotypicality: in-group/out-group) x 2 (trait valence: positive/negative) x 2 (target group: in-group/out-group) mixed model design with repeated measures on the last three factors. We recruited 75 French first year psychology undergraduate students (7 men, 68 women), aged 18 to 44 years old ($M = 21.23$), from a French university. Eighteen to twenty participants were randomly assigned to each of the four experimental conditions. A female experimenter conducted the experiment in five sessions,

each containing between 15 and 20 participants.

Procedure. We provided all instructions and stimuli in the French language. We told participants that we were investigating how people view people who overestimate and underestimate the number of physical stimuli that they perceive. We provided participants with false background information about the distinction between overestimators and underestimators (adapted from Brown, Collins, & Schmidt, 1988, p. 446). Specifically, we told participants that research had established that people tend to either consistently overestimate or underestimate the number of physical stimuli that they perceive, that approximately half of the population are underestimators and half are overestimators, and that the tendency to over- or under-estimate is related to a number of personality traits. We cited fictitious research that purported to show that overestimators tend to be ambitious and imaginative but also arrogant and clumsy and that underestimators tend to be modest and responsible but also pessimistic and reserved. Previous research has established that these traits are generally perceived to be either socially desirable (ambitious, imaginative, modest, responsible) or socially undesirable (arrogant, clumsy, pessimistic, reserved; Anderson, 1968; Bochner & van Zyl, 1985; Hampson, Goldberg, & John, 1987). In addition, Rubin et al. (2001, Study 1) found that six of these eight traits are perceived to be stereotypical of overestimators (ambitious, imaginative, arrogant) or underestimators (modest, pessimistic, reserved).²

Following the background information, we asked participants to take part in a dot estimation task that was allegedly used to distinguish between overestimators and underestimators (e.g., Brown et al., 1988). Participants were shown five overhead transparencies, each for a period of three seconds. Each slide showed between 50 and 100 dots. Participants estimated the number of dots that appeared on each slide and wrote their estimates on a response sheet.

After the dot estimation task, the experimenter categorized each participant as an overestimator or underestimator, ostensibly on the basis of their responses during the dot estimation task, but actually on a random basis. After group assignment, we gave participants a questionnaire in which they were informed that they would be rating both overestimators and underestimators during the research. This information was provided in order to ensure that participants made their group ratings in an intergroup context.³

Participants began the questionnaire by completing an 8-item measure of in-group identification. This measure was based on items from the identity self-esteem subscale of Luhtanen and Crocker's (1992) Collective Self-Esteem Scale and Brown et al.'s (1986) measure of group identification.

Participants then made mean ratings and/or variability ratings. We used the same type of scale for all group ratings in order to exclude scale type as a confounding variable. This scale consisted of a 95mm line anchored *not at all* at the left end and *extremely* at the right end.⁴

Participants in the mean ratings present/range measure condition received a questionnaire in which they were presented with the eight traits arranged in a single random order, with each trait followed by a 95mm scale. Participants placed a cross along the 95mm scale in order to indicate how much of each trait they imagined most people in a target group possessed (0 = *Not at all*, 95 = *Extremely*). Immediately after completing these mean ratings for a target group, participants returned to the same eight scales and provided range ratings for that group. These range ratings involved participants making two slash marks on either side of their original crosses in order to indicate where they imagined the two most extreme target group members fell for each trait (e.g., Park & Judd, 1990, Study 1). Participants in the mean ratings absent/range measure condition performed a similar task except that they made their range ratings without providing mean ratings beforehand.

Participants in the mean ratings present/similarity measure condition performed a similar task to those in the mean ratings present/range measure condition except that they only made one slash mark on each scale after making their mean ratings. This single slash mark indicated how similar target group members were perceived to be to one another on each trait (0 = *Not at all*, 95 = *Extremely*). Participants in the mean ratings absent/similarity measure condition made their similarity ratings without providing mean ratings beforehand.

Participants made their group ratings on separate pages of the questionnaire for each target group. All participants judged the overestimator group first, followed by the underestimator group. Given that half of our participants belonged to the overestimator group and half belonged to the underestimator group, this rating order meant that the order of in-group and out-group ratings was counterbalanced across participants.

Finally, in order to check the perceived valence of the eight traits, we asked participants to indicate the extent to which they considered each trait to be a desirable or undesirable characteristic of their personality. Participants made their ratings on a 5-point Likert-type scale (1 = *Extremely undesirable*, 5 = *Extremely desirable*).

Results and Discussion

Checks on perceived trait stereotypicality and valence. We performed a series of 2 (participants' group: overestimators/underestimators) x 2 (variability measure: range/similarity) x 2 (target group: overestimators/underestimators) mixed-model analyses of variance (ANOVAs) with repeated measures on the last factor on the mean ratings data for each of the eight traits in order to establish whether participants in the mean ratings present conditions perceived each trait to be stereotypical of the appropriate target group. The expected main effect of target group was significant on all of the traits ($ps < .01$) apart from "clumsy" ($p = .73$). Consequently, we excluded data from this trait from our subsequent analyses. The main effect of target group was not involved in any significant interactions with variability measure ($ps > .06$), but it was qualified by participants' group for all of the traits apart from "imaginative" ($ps < .05$). In all of the cases apart from "responsible", participants' group interacted with target group in an ordinal manner such that the main effect of target group remained significant at each level of participants' group ($ps < .01$). In the case of "responsible", the simple main effect of target group was not significant among participants in the overestimator group ($p = .11$). Consequently, we excluded the data from this trait from our subsequent analyses.

Note that our check on perceived trait stereotypicality also served as a check on the effectiveness of our manipulation of stereotype expression. The results indicated that participants used their mean ratings to express to the researcher their perception that the in-group possessed in-group traits but not out-group traits and that the out-group possessed out-group traits but not in-group traits.

We performed a one-sample t test on the valence check data for each trait in order to determine whether the mean value was significantly different from the scale's neutral midpoint of 3. All traits were significantly different from neutral and in the predicted directions ($ps < .01$ apart from "reserved", where $p = .05$).

Creating OH-IH indices. For the range measure, we measured the distance between the left and right slash marks to the nearest millimetre for each trait. Larger distances between these two ratings indicated greater perceived intragroup dispersion. For the similarity measure, we reverse scored ratings so that larger scores indicated less perceived intragroup similarity. Because our hypotheses were specifically related to OH and IH effects, we performed our analyses on indices representing these effects. We created these OH-IH indices by subtracting average out-group ratings from average in-group ratings (e.g., Rubin et al., 2001). Scores on these indices could range from +95 to -95, with positive values indicating OH and negative values indicating IH. We found three outliers on the OH-IH

indices (< -3.07 or > 3.32 SDs). These three participants were excluded from all subsequent analyses.

Testing our stereotype expression explanation. We performed a 2 (mean ratings: absent/present) $\times$ 2 (variability measure: range/similarity) $\times$ 2 (trait stereotypicality: in-group/out-group) $\times$ 2 (trait valence: positive/negative) mixed model ANOVA on the OH-IH indices. We found a significant main effect of trait stereotypicality, $F(1, 68) = 9.22, p < .01$, a significant two-way interaction between mean ratings and trait stereotypicality, $F(1, 68) = 6.53, p = .01$, and a significant two-way interaction between variability measure and trait stereotypicality, $F(1, 68) = 15.87, p < .01$. However, all of these effects were qualified by a significant three-way interaction between mean ratings, variability measure, and trait stereotypicality, $F(1, 68) = 8.01, p < .01$. In order to investigate this three-way interaction, we performed 2 (mean ratings: absent/present) $\times$ 2 (trait stereotypicality: in-group/out-group) mixed model ANOVAs on the data from each of our two variability measures: range and similarity.

On the range measure, we found no significant effects ($ps > .23$). Hence, contrary to our predictions, we failed to find the stereotype effect on the range measure in the mean ratings absent condition. This condition replicated the range procedure used by Brown and Wootton-Millward (1993) and Kelly (1989) in the only two studies to find the stereotype effect using a range measure. We concluded that, although these studies demonstrate that it is possible to find the stereotype effect on the range measure, the range measure may not be the most sensitive measure of this particular effect.

On the similarity measure, we found a main effect of trait stereotypicality, $F(1, 33) = 13.96, p < .01$. This main effect was qualified by a two-way interaction between mean ratings and trait stereotypicality, $F(1, 33) = 8.22, p < .01$. The mean values and simple effects for this two-way interaction are presented in Table 2.

----- Insert Table 2 here -----

Looking at Table 2, it can be seen that, in the mean ratings absent condition, significant IH occurred on in-group traits ($M = -2.18$), $t(15) = -4.51, p < .01$, significant OH occurred on out-group traits ($M = 1.94$), $t(15) = 3.16, p < .01$, and there was a significant difference between these two effects, $F(1, 33) = 20.08, p < .01$, indicating a significant stereotype effect. In contrast, in the mean ratings present condition, there were no significant IH or OH effects on either in-group or out-group traits ($ps > .48$), and the difference between means was not significant ($p = .52$). This pattern of results is consistent with our stereotype expression explanation of the stereotype effect: Participants who expressed trait stereotypicality via their mean ratings felt obligated not to provide the same information in the form of a stereotype effect during their variability ratings.

Experiment 2

In Experiment 2, we aimed to increase the validity of our findings by using a different manipulation of stereotype expression and an additional measure of perceived intragroup variability. We elaborate on each of these changes in turn.

First, it is possible that the moderating effects of stereotype expression that we found in Experiment 1 were limited to the particular methodology that we used to manipulate this variable. In order to demonstrate the generalizability of prior stereotype expression as a moderator of the stereotype effect, we used an alternative manipulation of stereotype expression in Experiment 2. Specifically, instead of asking participants to provide mean ratings about the *possession* of traits (*not at all, extremely*), we asked them to indicate how *typical* traits were of the target groups (*very typical of underestimators, very typical of overestimators*). This new procedure provided participants with a more specific and direct method for expressing trait stereotypicality to the researcher than the method that we used in Experiment 1.

Second, the stereotype effect has frequently been found using distribution measures (Castano & Yzerbyt, 1998, Studies 1 & 2; Simon, 1992a; Simon & Pettigrew, 1990). Hence, we included a distribution measure in Experiment 2 in order to increase the comparability of our research with this previous research.

Method

Participants and design. We recruited 104 French first year psychology undergraduate students (11 men, 93 women), aged 19 to 57 years old ($M = 22.01$), from a French university. We used a 2 (ratings order: variability first/trait typicality first) $\times$ 3 (variability measure: range/similarity/distribution) $\times$ 2 (trait stereotypicality: in-group/out-group) $\times$ 2 (trait valence: positive/negative) $\times$ 2 (target group: in-group/out-group) mixed model design with repeated measures on the last three factors. Sixteen to nineteen participants were randomly assigned to each of the six experimental conditions. A female experimenter conducted the experiment in six sessions, each containing between 15 and 20 participants.

Procedure. The procedure was the same as that for the mean ratings absent conditions of Experiment 1 apart from the following changes: Participants indicated how typical each of the eight traits was of underestimators or overestimators on a 5-point scale (1 = *Extremely typical of underestimators*, 5 = *Extremely typical of overestimators*). Participants in the variability first conditions made these ratings immediately after their variability ratings. Participants in the trait typicality first conditions made these ratings immediately before their variability ratings.

Participants in the distribution measure conditions were asked to imagine 100 target group members and then, for each trait, to indicate how many group members fell into each of seven boxes anchored *not at all* and *extremely* (e.g., Linville et al., 1989). Hence, participants could indicate how many target group members did not possess the trait by writing a number from 0 to 100 inside the first box, how many possessed a very small amount of the trait by writing a number from 0 to 100 inside the second box, and so on. For each trait, participants were asked to write numbers inside all seven boxes and to make sure that their ratings summed to 100.

In light of the convincing evidence from Experiment 1, we did not include the measure of trait valence in this second experiment.

Results and Discussion

Check on perceived trait stereotypicality. We performed one sample t tests on the data from the perceived trait typicality measure, using the midpoint of the response scale (3) as the reference value. As in Experiment 1, participants rated all traits as being significantly different from the scale midpoint ($ps < .01$), apart from “clumsy” ($p = .08$). In order to check that perceptions of trait stereotypicality generalised across participants assigned to the overestimator and underestimator groups, we conducted the same set of one sample t tests within each subsample. These tests revealed that participants in the overestimator group did not perceive the traits “clumsy” and “responsible” as being significantly typical of the relevant target groups ($ps > .26$). Consequently, as in Experiment 1, we also excluded data from the trait “responsible” from our subsequent analyses.

In order to check that participants in the trait typicality first condition expressed trait stereotypicality, we restricted our previous tests to this subsample. We found that participants in the trait typicality first condition rated all traits as being significantly different from the scale midpoint ($ps < .03$), although participants in the overestimator group did not perceive traits “clumsy” and “responsible” as being significantly typical of the relevant target groups ($ps > .20$).

Creating OH-IH indices. We converted the distribution, range, and similarity data to z -scores in order to provide a standard metric between the distribution indices and the range

and similarity indices. We followed Linville et al.'s (1986, pp. 184-186) formulae in order to compute the probability of differentiation (P_d) and standard deviation (SD) from the perceived distribution measure. P_d is defined by the formula $P_d = 1 - \sum_{i=1}^n P_i^2$, where P_i denotes the proportion of the group's members that are described by the level (scale value) i of the trait in question. SD is defined by the formula $SD = \sqrt{\sum_i P_i (X_i - M)^2}$. A higher value of P_d and SD indicated a higher perceived variability.

We created OH-IH indices for all measures by subtracting average out-group ratings from average in-group ratings. We found four outliers on the OH-IH indices (< -3.07 and > 3.06 SDs). These four participants were excluded from all subsequent analyses.

Testing our stereotype expression explanation. We performed a 2 (ratings order: variability first/trait typicality first) $\times$ 3 (variability measure: range/similarity/distribution) $\times$ 2 (trait stereotypicality: in-group/out-group) $\times$ 2 (trait valence: positive/negative) mixed model ANOVA on the OH-IH indices, with the distribution measure based on P_d .⁵ Consistent with Experiment 1, there was a significant two-way interaction between ratings order and trait stereotypicality, $F(1, 94) = 4.41, p = .04$. The mean values and simple effects for this two-way interaction are presented in Table 3.

----- Insert Table 3 here -----

Looking at Table 3, it should be noted that no IH or OH effects were significant ($ps > .07$). Nonetheless, the pattern of means and simple main effects in this two-way interaction were consistent with our stereotype expression explanation: In the variability first condition, the difference between the tendency towards IH on in-group traits ($M = -.12$) and the tendency towards OH on out-group traits ($M = .17$) was approaching significance, $F(1, 94) = 3.34, p = .07$. In contrast, the difference between the two means in the trait typicality first condition was not significant, $F(1, 94) = 1.31, p = .26$. These results are consistent with our prediction that participants in the variability first condition would express the differential possession of in-group and out-group traits using their variability ratings, whereas participants in the trait typicality first condition would not.

General Discussion

Summary

We conducted two experiments in which we manipulated the presence or absence of stereotype expression prior to ratings of intragroup variability. In both experiments, we found that the stereotype effect only occurred in the absence of preceding expressions of trait stereotypicality. The stereotype effect did not occur when participants expressed their perceptions of trait stereotypicality before making their ratings of intragroup variability. Hence, our experimental research complements the results of our literature review by providing evidence that the stereotype effect is moderated by the presence or absence of preceding expressions of trait stereotypicality. Assuming the operation of a given-new contract in research situations (Bless et al., 1993), the finding that the stereotype effect is nullified by preceding acts of stereotype expression suggests that the stereotype effect is itself a form of stereotype expression.

Our predicted effects were qualified by variability measure in Experiment 1, but not in Experiment 2 ($p = .53$). Hence, we cannot draw any firm conclusions about the moderating role of variability measure from our research.

Alternative Explanation

An alternative explanation of our results relates to the confound between stereotype expression and in-group favoritism in our research: It is possible that participants exhibited in-group favoritism during their mean and trait typicality ratings, and that this favoritism increased the perceived positivity of the in-group and its associated social identity, thus

reducing the social identity motive for the stereotype effect during subsequent variability ratings. However, this explanation is only tenable if it is assumed that the stereotype effect addresses the need for a positive social identity. We accept that the stereotype effect can contribute to a positive social identity if in-group traits are perceived to be more positive than out-group traits. We tested this possibility in our research using the trait valence data from Experiment 1. We performed a paired samples *t* test on the average perceived desirability of in-group and out-group traits. In-group traits were not perceived to be significantly more desirable ($M = 2.96$) than out-group traits ($M = 3.02$), $t(71) = -.59$, $p = .56$. Hence, the stereotype effect cannot be conceived as a strategy for increasing the positivity of the in-group and its associated social identity, even after taking into account the overall valence of the particular in-group and out-group traits that we used in our research. Consequently, the nullifications of the stereotype effect that we obtained in our research cannot be attributed to a reduction in the motive for a positive social identity following in-group favouritism.

Implications and Limitations

From a general perspective, our research suggests that people's perceptions of trait possession can influence their ratings of group homogeneity. Researchers who intend to use measures of perceived group homogeneity should consider making provisions for this potential confound when designing and interpreting their research. Having said this, we believe that the influence of the "homogeneity equals possession" heuristic is likely to be restricted to those situations in which there are clear intergroup differences in trait possession. So, for example, our stereotype expression explanation *cannot* be used to account for the general finding of OH because there is no theoretical reason or empirical evidence to suggest that out-groups are perceived to possess traits to a greater extent than in-groups.

Self-Categorization Theory and the Given-New Contract

We conclude by considering the relationship between our research findings and the explanation of group stereotyping that is provided by self-categorization theory (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987). According to self-categorization theory, tasks that require people to compare their in-group with an out-group should increase the salience of the intergroup distinction and associated social identity and, consequently, increase the extent of stereotyping (e.g., S. A. Haslam, Oakes, Reynolds, & Turner, 1999). In our stereotype expression conditions, we asked participants to indicate the extent to which the in-group and an out-group possessed a series of traits before they made judgements of intragroup variability. Hence, according to self-categorization theory, our manipulation of stereotype expression should have increased, rather than decreased, the stereotype effect, by intensifying participants' stereotypical perceptions. How do we account for this apparent discrepancy?

Although we found that participants expressed less, rather than more, group stereotyping following an in-group/out-group rating task, we do not believe that our evidence contradicts self-categorization theory. Instead, we believe that we have identified an important qualification for the self-categorization explanation: Tasks that increase identity salience should only increase subsequent stereotyping when they do not themselves provide an outlet for stereotype expression. Tasks that increase identity salience *and* allow participants to express stereotypes to the researcher should *reduce* stereotype expression on subsequent measures due to the operation of the given-new contract. Future researchers may wish to provide a direct test of this novel hypothesis.

References

- Anderson, N. H. (1968). Likeableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, 9, 272-279.
- Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, 42, 228-235.
- Bless, H., Strack, F., & Schwarz, N. (1993). The informative functions of research procedures: Bias and the logic of conversation. *European Journal of Social Psychology*, 23, 149-165.
- Bochner, S., & Van Zyl, T. (1985). Desirability ratings of 110 personality trait words. *Journal of Social Psychology*, 125, 459-465.
- Boldry, J. G., & Kashy, D. A. (1999). Intergroup perception in naturally occurring groups of differential status: A social relations perspective. *Journal of Personality and Social Psychology*, 77, 1200-1212.
- Brewer, M. B. (1993). Social identity, distinctiveness, and in-group homogeneity. *Social Cognition*, 11, 150-164.
- Brown, J., Collins, R. L., & Schmidt, G. W. (1988). Self-esteem and direct vs. indirect forms of self enhancement. *Journal of Personality and Social Psychology*, 55, 445-453.
- Brown, R., Condor, S., Mathews, A., Wade, G., & Williams, J. (1986). Explaining intergroup differentiation in an industrial organization. *Journal of Occupational Psychology*, 59, 273-286.
- Brown, R., & Wootton-Millward, L. (1993). Perceptions of group homogeneity during group formation and change. *Social Cognition*, 11, 126-149.
- Campbell, D. T. (1958). Common fate, similarity, and other indices of the status of aggregates of persons as social entities. *Behavioral Science*, 3, 14-25.
- Castano, E., & Yzerbyt, V. Y. (1998). The highs and lows of group homogeneity. *Behavioural Processes*, 42, 219-238.
- Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In Freedle, R. O. (Ed.), *Discourse production and comprehension* (pp. 1-40). Norwood, NJ: Ablex.
- De Cremer, D. (2001). Perceptions of group homogeneity as a function of social comparison: The mediating role of group identity. *Current Psychology: Developmental, Learning, Personality, Social*, 20, 138-146.
- Devos, T., Comby, L., & Deschamps, J. (1996). Asymmetries in judgements of ingroup and outgroup variability. *European Review of Social Psychology*, 7, 95-144.
- Doosje, B., Spears, R., & Koomen, W. (1995). When bad isn't all bad: Strategic use of sample information in generalization and stereotyping. *Journal of Personality and Social Psychology*, 69, 642-655.
- Doosje, B., Ellemers, N., & Spears, R. (1999). Commitment and intergroup behavior. In N. Ellemers, R. Spears, & B. Doosje (Eds.), *Social identity: Context, commitment, content*. (pp. 84-106). Oxford: Blackwell.
- 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.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Volume 3, Speech Acts* (pp. 41-58). New York: Academic Press.
- Guinote, A., Judd, C. M., & Brauer, M. (2002). Effects of power on perceived and objective group variability: Evidence that more powerful groups are more variable. *Journal of Personality and Social Psychology*, 82, 708-721.
- Hamilton, D. L., Sherman, S. J., & Lickel, B. (1998). Perceiving social groups: The importance of the entitativity continuum. In C. Sedikides, J. Schopler & C. A. Insko

- (Eds.), *Intergroup cognition and intergroup behaviour* (pp. 47-74). Mahwah, NJ: Lawrence Erlbaum.
- Hampson, S. E., Goldberg, L. R., & John, O. P. (1987). Category-breadth and social-desirability values for 573 personality terms. *European Journal of Personality*, 1, 241-258.
- Haslam, N., Rothschild, L., & Ernst, D. (2000). Essentialist beliefs about social categories. *British Journal of Social Psychology*, 39, 113-127.
- Haslam, S. A., Oakes, P. J., Reynolds, K. J., & Turner, J. C. (1999). Social identity salience and the emergence of stereotype consensus. *Personality and Social Psychology Bulletin*, 25, 809-818.
- Haslam, S. A., Oakes, P. J., Turner, J. C., & McGarty, C. (1995). Social categorization and group homogeneity: Changes in the perceived applicability of stereotype content as a function of comparative context and trait favourableness. *British Journal of Social Psychology*, 34, 139-160.
- Haslam, S. A., Oakes, P. J., Turner, J. C., & McGarty, C. (1996). Social identity, self-categorization, and the perceived homogeneity of ingroups and outgroups: The interaction between social motivation and cognition. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition* (Vol. 3, pp. 182-222). New York: Guilford Press.
- Hewstone, M., Islam, M. R., & Judd, C. M. (1993). Models of crossed categorization and intergroup relations. *Journal of Personality and Social Psychology*, 64, 779-793.
- Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, 118, 248-271.
- Hopkins, N., & Cable, I. (2001). Group variability judgments: Investigating the context-dependence of stereotypicality and dispersal judgments. *British Journal of Social Psychology*, 40, 455-470.
- Jones, E. E., Wood, G. C., & Quattrone, G. A. (1981). Perceived variability of personal characteristics in in-groups and out-groups: The role of knowledge and evaluation. *Personality and Social Psychology Bulletin*, 7, 523-528.
- Kelly, C. A. (1989). Political identity and perceived intragroup homogeneity. *British Journal of Social Psychology*, 28, 239-250.
- Lambert, A. J. (1995). Stereotypes and social judgment: The consequences of group variability. *Journal of Personality and Social Psychology*, 68, 388-403.
- Lambert, A. J., & Wyer, R. S. (1990). Stereotypes and social judgment: The effects of typicality and group heterogeneity. *Journal of Personality and Social Psychology*, 59, 676-691.
- Lee, Y. T., & Ottati, V. (1995). Perceived in-group homogeneity as a function of group membership salience and stereotype threat. *Personality and Social Psychology Bulletin*, 21, 610-619.
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74, 1421-1436.
- Linville, P. W. (1998). The heterogeneity of homogeneity. In J. M. Darley & J. Cooper (Eds.), *Attribution and social interaction: The legacy of Edward E. Jones* (pp. 423-487). Washington, DC: American Psychological Association.
- Linville, P. W., Fischer, G. W., & Salovey, P. (1989). Perceived distributions of 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., Salovey, P., & Fischer, G. W. (1986). Stereotyping and perceived distributions of social characteristics: An application to in-group-out-group

- perception. In J. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 165-208). San Diego, CA: Academic Press.
- Luhtanen, R., & Crocker, J. (1992). A collective self esteem scale: Self evaluation of one's social identity. *Personality and Social Psychology Bulletin*, 18, 302-318.
- Mullen, B., & Hu, L. (1989). Perceptions of in-group and out-group variability: A meta-analytic integration. *Basic and Applied Social Psychology*, 10, 233-252.
- Ostrom, T. M., & Sedikides, C. (1992). Out-group homogeneity effects in natural and minimal groups. *Psychological Bulletin*, 112, 536-552.
- Park, B., & Hastie, R. (1987). Perception of variability in category development: Instance-versus abstraction-based stereotypes. *Journal of Personality and Social Psychology*, 53, 621-635.
- Park, B., & Judd, C. M. (1990). Measures and models of perceived group variability. *Journal of Personality and Social Psychology*, 59, 173-191.
- Park, B., & Rothbart, M. (1982) Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, 42, 1051-1068.
- Pickett, C. L., & Brewer, M. B. (2001). Assimilation and differentiation needs as motivational determinants of perceived in-group and out-group homogeneity. *Journal of Experimental Social Psychology*, 37, 341-348.
- Plaks, J. E., Levy, S. R., Dweck, C. S., & Stroessner, S. J. (2004). In the eye of the beholder: Lay theories and the perception of group entitativity, variability, and essence. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 127-146). New York: Psychology Press.
- Rothgerber, H. (1997). External threat as an antecedent to perceptions of in-group and out-group homogeneity. *Journal of Personality and Social Psychology*, 73, 1206-1212.
- Rubin, M., Hewstone, M., & Voci, A. (2001). Stretching the boundaries: Strategic perceptions of intragroup variability. *European Journal of Social Psychology*, 31, 413-429.
- Rubin, M., Hewstone, M., Crisp, R. J., Voci, A., & Richards, Z. (2004). Gender out-group homogeneity: The roles of differential familiarity, gender differences, and group size. In V. Yzerbyt, C. M. Judd & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 203-220). New York: Psychology Press.
- Ryan, C. S., Judd, C. M., & Park, B. (1996). Effects of racial stereotypes on judgments of individuals: The moderating role of perceived group variability. *Journal of Experimental Social Psychology*, 32, 71-103.
- Ryan, C. S., Robinson, D. R., & Hausmann, L. R. (2001). Stereotyping among providers and consumers of public mental health services: The role of perceived group variability. *Behavior Modification*, 25, 406-442.
- Schwarz, N. (1996) *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Mahwah, NJ: Lawrence Erlbaum.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54, 93-105.
- Simon, B. (1992a). Intragroup differentiation in terms of ingroup and outgroup attributes. *European Journal of Social Psychology*, 22, 407-413.
- Simon, B. (1992b). The perception of ingroup and outgroup homogeneity: Reintroducing the intergroup context. *European Review of Social Psychology*, 3, 1-30.
- Simon, B., & Brown, R. (1987). Perceived intragroup homogeneity in minority majority

- contexts. *Journal of Personality and Social Psychology*, 53, 703-711.
- Simon, B., Glässner-Bayerl, B., & Stratenwerth, I. (1991). Stereotyping and self-stereotyping in a natural intergroup context: The case of heterosexual men and homosexual men. *Social Psychology Quarterly*, 54, 252-266.
- Simon, B., & Hamilton, D. L. (1994). Self-stereotyping and social context: The effects of relative in-group size and in-group status. *Journal of Personality and Social Psychology*, 66, 699-711.
- Simon, B., & Pettigrew, T. F. (1990). Social identity and perceived group homogeneity: Evidence for the in-group homogeneity effect. *European Journal of Social Psychology*, 20, 269-286.
- Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology*, 18, 429-442.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1, 149-178.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks-Cole.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 327-352.
- Voci, A. (2000). Perceived group variability and the salience of personal and social identity. *European Review of Social Psychology*, 11, 177-221.
- Yzerbyt, V., Corneille, O., Estrada, C. (2001). The interplay of subjective essentialism and entitativity in the formation of stereotypes. *Personality and Social Psychology Review*, 5, 141-155.

Endnotes

1. We excluded studies that used measures of perceived group stereotypicality from our literature review (e.g., S. A. Haslam et al., 1995; Park & Judd, 1990; Park & Rothbart, 1982) because these measures identify IH and OH effects through an interaction between target group and trait stereotypicality, and they therefore preclude the possibility of demonstrating the moderating effects of trait stereotypicality on IH and OH effects (i.e., the stereotype effect).

2. Rubin et al. (2001) found that the traits “clumsy” and “responsible” were not perceived to be stereotypical of overestimators and underestimators respectively. In contrast to Rubin et al., we explicitly informed participants about the supposed stereotypical relationship between traits and groups as part of the background information that we provided. Given this procedural difference, we hoped that participants would perceive these two traits as being stereotypical of the appropriate target groups.

3. Based on self-categorization theory (Turner et al., 1987), S. A. Haslam et al. (1995, 1996) explained the OH effect as a result of rating the in-group in an intragroup context and the out-group in an intergroup context. This rating situation is most likely to occur when participants rate the in-group before the out-group and they are initially unaware of the intergroup nature of their judgments.

4. We originally intended to use a 100mm scale. However, a printing error resulted in a 95mm scale.

5. When we replaced the P_d measure with the SD measure in our global analysis, we did not find a significant two-way interaction between ratings order and trait stereotypicality ($p = .19$), and no other effects were significant ($ps > .11$).

Table 1
Occurrence of the Stereotype effect in Previous Research

Study	Type of variability measure	Trait possession ratings made before variability ratings?	Stereotype effect present?
Kelly (1989)	Range	No	Yes
Simon and Pettigrew (1990)	P_d	No	Yes
Simon (1992a)	SD and P_d	No	Yes
Brown and Wootton-Millward (1993)	Range	No	Yes
Simon and Hamilton (1994, Study 1)	Similarity	Yes	No
Judd et al. (1995, Study 1)	Range	Yes	No
Judd et al. (1995, Study 2)	Range	Yes	No
Castano and Yzerbyt (1998, Study 1)	SD and P_d	No	Yes
Castano and Yzerbyt (1998, Study 2)	SD and P_d	No	Yes
Rubin et al. (2001, Study 1)	Range	Yes	No
Rubin et al. (2001, Study 2)	Range	Yes	No
Rubin et al. (2001, Study 3)	Range	Yes	No
Pickett and Brewer (2001)	Similarity	No	Yes
Hopkins and Cable (2001)	Range	Yes	No

Note. “Similarity” refers to the similarity measure of perceived intragroup variability (e.g., Park & Rothbart, 1982); “SD” refers to the standard deviation measure and “ P_d ” refers to the probability of distribution measure (Linville et al., 1989); “range” refers to the range measure (e.g., Jones, Wood, & Quattrone, 1981)

Table 2

IH and OH on the Similarity Measure as a Function of Mean Ratings and Trait Stereotypicality

Mean ratings	Trait stereotypicality	
	In-group	Out-group
Absent	-2.18 _a *	1.94 _b *
Present	-.09 _b	.45 _b

Note. Means with different subscripts differ significantly within columns ($ps < .01$) and within rows ($p < .01$). Means with asterisks differ significantly from zero ($p < .01$).

Table 3

IH and OH as a Function of Ratings Order and Trait Stereotypicality

Ratings order	Trait stereotypicality	
	In-group	Out-group
Variability first	-.12 _a	.17 _b
Trait typicality first	.04 _a	-.14 _a

Note. Means with different subscripts are significant within columns ($p = .02$) and approaching significance within rows ($p = .07$). Means are not significantly different from zero ($ps > .07$).