Stereotypes and Social Judgment: The Consequences of Group Variability

Alan J. Lambert
Washington University

The present investigation examined the effects of group variability on judgments of single group members. Male and female participants formed impressions of a group of 50 men or women on the basis of their performance on a test of perceptual-motor skills. The variability of group performance varied across conditions. Participants then made speeded typicality judgments and ability ratings of several “new” group members whose performance varied in its discrepancy from the group. Compared with participants in the high variability condition, participants in the low variability condition were (a) more likely to judge discrepant group members as atypical and (b) faster to assess their atypicality. This latter effect decreased the probability that participants in the low variability condition used the group as a basis for judging atypical group members.

Stereotypes have long been regarded as a set of generalizations about the attributes of the average (prototypical) group member (e.g., Allport, 1954; Katz & Braly, 1933). More recent research has indicated, however, that perceivers’ mental representations of social and nonsocial categories are likely to contain information about not only the central tendency of categories but also the variation across their members (e.g., Fried & Holyoak, 1984; Lambert & Wyer, 1990; Linville, Fischer, & Salovey, 1989; Park & Hastie, 1987; Park & Rothbart, 1982). 

Recognition that the members of a group are very different from one another might be viewed as diametrically opposed to what it means to hold a stereotype. That is, one is no longer generalizing about the group; distinctions are being made. To this extent, perceivers ought to rely less on a person’s group membership as a basis for judgment if the group is perceived to be heterogeneous than if it is not. This intuitive view of group variability was nicely summarized in a recent article by Judd, Ryan, and Park (1991):

Presumably, if one sees a group as relatively diverse or heterogeneous, one is more reluctant to routinely apply the perceived attributes of the prototypic group member to a newly-encountered individual. As the perceived variability of the group increases, one may be more willing to take into account individuating information rather than prejudging newly encountered individuals simply on the basis of group membership. (p. 366)

Judd et al. (1991) are certainly not alone in making this point. Indeed, the notion that heterogeneous categories should have relatively weak or undiagnostic implications for social judgment has been widely assumed among researchers in this area (e.g., Kraus, Ryan, Judd, Hastie, & Park, 1993; Linville et al., 1989; Park & Hastie, 1987; Park & Judd, 1990; Pettigrew, 1981).

Although this assumption regarding the moderating effects of group variability on social judgment seems self-evident, surprisingly few studies have examined directly the effects of group variability on judgments of single group members. Of those few studies that have, at least one (Lambert & Wyer, 1990) generated results that are at odds with this intuitive view of group variability. Specifically, Lambert and Wyer found that participants were more, not less, likely to directly apply their beliefs about the group to an atypical group member if the group was heterogeneous than if it was homogeneous. Conceptually similar findings were recently obtained by Lord, Desforges, Ramsey, Trezza, and Lepper (1991).

Although such findings are provocative, the processes underlying these seemingly paradoxical effects are not entirely clear. Moreover, existing models of category-based inferences (e.g., Brewer, 1988; Fiske & Neuberg, 1990) are silent on the role of group variability in social judgment. The overriding goal of the present research, therefore, was to develop a general framework that specifies the process by which group variability can influence perceivers’ reactions to single group members. I use this conceptualization not only to account for the present set of findings but also to provide a framework for integrating the results of several previous investigations that bear directly or indirectly on the role of group variability in social judgment (e.g., Lambert & Wyer, 1990; Lord et al., 1991; Manis, Nelson, & Shedler, 1988; Park & Hastie, 1987).

Acquisition and Representation of Information About Categories and Individuals

Suppose participants are asked to form an impression of a group on the basis of information about the test performance...
of its members. Several models of social categorization (e.g., Hintzman, 1986; Linville et al. 1989) assume that perceivers will form cognitive representations of the individual group members (exemplars) as information about them is received. In addition, perceivers may also abstract (i.e., summarize) various properties of the group, such as central tendency, group variability, and the frequency with which certain attributes occur in the group (e.g., Barsalou, 1990; Kraus et al., 1993; Park & Hastie, 1987).

At present there is considerable disagreement as to the extent to which (and how) exemplar versus abstracted information is represented in memory and the boundary conditions under which such representations are likely to be formed (e.g., Judd & Park, 1988; Linville & Fischer, 1993; Messick & Mackie, 1989; E. R. Smith & Zarate, 1992). Nevertheless, a number of theorists (e.g., Park & Hastie, 1987) have argued that both abstract-based and exemplar-based models of categories are likely to be correct. That is, under some conditions perceivers can form representations of categories that are largely in the form of exemplars, but under other conditions, abstract-based representations may be more likely. These possibilities were considered in the present study.

Consequences of Group Variability for Social Judgment

In this study, I hypothesized that group variability can have at least three general effects on the way that perceivers process information about a single group member. Specifically, group variability could influence (a) whether group members are judged as typical or atypical; (b) the ease with which typicality is assessed; and (c) the likelihood that perceivers apply their beliefs about the group as a whole to any given group member. In the discussion to follow, I consider the theoretical and empirical issues that are germane to each of these three possibilities. In the course of this discussion, I also consider the possible role that ingroup versus outgroup membership might play in these processes.

Effects of Group Variability on Perceptions of Typicality

Although there appears to be no single criterion for assessing the typicality of category members (Barsalou, 1985; Malt & Smith, 1982), several investigations have indicated that typicality is often based on the frequency with which the person’s attributes occur in the group (Barsalou, 1985, 1987; Fried & Holyoak, 1984). According to this view, for example, the extent to which a Graduate Record Examination verbal score of 650 would be perceived as typical of college students would be based on an assessment of how often similar scores occur among members of this group.

Some straightforward implications emerge from this perspective. Suppose participants are given information about a group whose performance on a standardized test is either heterogeneous or homogeneous. A representation of these two distributions is shown in Figure 1, along with (a) the percentage with which each score occurs in the group and (b) the positions of three individual scores (x, y, z). Although the standard deviations of these two distributions are obviously different, the mean and median values, as well as the range, are identical.

As seen in Figure 1, performances in the tails of the distribution (i.e., y) occur more often when the group is highly variable than when it is not (10% vs. 2%). To this extent, these performances should be more likely to be judged as typical in the former case than in the latter. On the other hand, note that performance level x actually occurs somewhat less often when the group is heterogeneous than when it is not (12% vs. 20%), suggesting that such persons should be judged as slightly less typical if the group is heterogeneous than if it is not.1

Although performance level z falls outside the range of both distributions, group variability could still exert some effect on judged typicality here as well. This is especially likely given that perceivers may be somewhat unsure about the precise distribution of attributes among group members. In particular, participants who recall that a particular score occurs with a given frequency might infer that similar scores occur with about the same frequency. When the group is heterogeneous, scores that fall just outside the range of the distribution are, nevertheless, very similar to other scores that do occur relatively frequently. However, this is less true when the group is homogeneous. Therefore, such outlier scores might be judged as relatively more typical (or alternatively, less atypical) if the group is heterogeneous than if it is not. Although they did not directly assess perceived typicality, Park and Hastie (1987) reported results that are consistent with these suppositions. In that study, individuals whose behaviors were discrepant from the group were more likely to be classified as group members if the group was heterogeneous, regardless of whether the target's attributes were within or just outside the range of attributes presented about other group members. Conceptually similar findings were reported by Fried and Holyoak (1984) using nonsocial categories.

Effects of Group Variability on the Ease of Assessing Typicality

Typicality may be more difficult to assess under some conditions than under others. In particular, typicality should be relatively easy to assess when the relevant attribute of the group member occurs either very frequently or very infrequently. In these cases, conclusions that the person is either a typical or an atypical group member can be made quickly and with a minimum of cognitive effort. On the other hand, the typicality of a person whose attributes occur with only moderate frequency is less clear and, thus, should take more time to assess.2

1 Although the predictions to follow are applicable for unimodal distributions (the kind of distribution under investigation in the current research), this may not necessarily be the case for other (e.g., bimodal) distributions. The generalizability of this (as well as other) aspects of the methodology will be discussed in more detail later in this article.

2 These assertions are generally derived from a model of typicality by E. E. Smith, Shoben, and Rips (1974). However, Smith et al. were primarily concerned with classification judgments (i.e., decisions whether a stimulus belongs to a particular category or not) rather than whether the stimulus is a typical or atypical category member. Although these judgments are obviously related, predictions for the relative speed to make them are actually quite different. For example, Smith et al.'s formulation predicts that participants should take relatively long to decide if a penguin is a bird, because its attributes occur relatively frequently among other members of that category. However, assuming that participants already know that a penguin belongs to this category, they...
Two predictions follow directly from this reasoning. First, the typicality of scores that occur near the mode of the distribution, or outside of it, should be easier to assess than scores that fall between these two extremes. This general effect should occur regardless of perceived group variability. Independent of this effect (and more important for present concerns), typicality should generally be easier to assess if the group is homogeneous than if it is heterogeneous. In the homogeneous condition, note that there is a relatively narrow range of performances that are very common (i.e., obviously typical), but all other levels of performances are very rare (i.e., obviously atypical). In contrast, there is considerably more grey area in which the typicality of the performance is less obvious when the group is heterogeneous. Indeed, as Park and Hastie’s results (1987) imply, attributes that fall completely outside the range of other known group members are not so clearly atypical when the group is heterogeneous than when it is not. Furthermore, to the extent that outgroups are perceived as more homogeneous than ingroups (e.g., Park & Rothbart, 1982), the typicality of outgroup members might be more easily assessed than that of ingroup members.

However, a quite different set of considerations suggests that typicality of outgroup members might be easier to assess for reasons that have nothing to do with differences in perceived group variability per se. These considerations surround differences in the way that perceivers cognitively organize informa-

---

Figure 1. Relative frequency of performance in heterogeneous (top) and homogeneous (bottom) groups for performance levels x, y, and z.
tion about outgroups versus ingroups. A recent study by Os- 

trom, Carpenter, Sedikides, and Li (1993; see also Park & 

Rothbart, 1982) suggested that information about ingroups is organized in terms of person categories, whereas information about outgroups is organized in terms of the attributes that characterize the group as a whole (but see Mullen & Cooper, 1989). Such effects could be driven, at least in part, by the tendency for perceivers to focus on category-based attributes (e.g., gender or race) that distinguish the self from others and vice versa (McGuire & McGuire, 1987; Park & Rothbart, 1982; Taj- 

fel, 1981). That is, because the category membership of persons belonging to the outgroup is (by definition) different from one's own, perceivers may focus on and cognitively organize (i.e., abstract) information at this category level to a greater extent than for ingroups.

In any event, these considerations suggest that the ease of as- 

sessing typicality might differ for ingroups versus outgroups. As noted earlier, the frequency with which attributes occur in the group could, in principle, be abstracted at the category level along with other properties of the group (cf. Barsalou, 1990). If this is so, typicality could be easier to assess if frequency-based information has already been abstracted than if it has not. For instance, suppose that frequency information has not already been abstracted in memory. If this is so, the perceivers might engage in a memory search for any exemplar that has this attrib- 

ute, compute the frequency with which it occurs, and then use this value as a basis for determining typicality. On the other hand, when the relative frequency of this attribute is already represented in memory, this judgment should presumably take less time. If this line of reasoning is valid, then perceivers might be faster to assess the typicality of outgroup members than in- 

group members. Moreover, this should be true even after control- 

ling for differences in perceived variability.

**Summary**

Group variability could thus influence whether a person is judged as a typical or atypical group member and, along with ingroup versus outgroup membership, could also affect the ease with which this judgment is made. These two classes of effects could, in turn, produce yet a third effect, namely, to influence the manner in which perceivers apply their beliefs about the group to judgments of single group members. This possibility is explored below.

**Effects of Group Variability on the Use of Category- 

Based Information as a Basis for Judgment**

One motivation for conducting this investigation was to try to understand the seemingly paradoxical results obtained by Lambert and Wyer (1990), which were noted at the beginning of this article. In particular, why would participants be more likely to apply their beliefs about the group to an individual group member if the group were heterogeneous than if it were homogeneous? On a priori grounds, there seem to be two possi- 

ble answers to this question.

The first explanation is rather simple. Previous models of ste- 

tereotyping (e.g., Fiske & Neuberg, 1990) assume that perceived typicality acts as an informational cue or "gatekeeper" such that if the person is typical, inferences about the target are based on perceivers' more general beliefs about the group (i.e., an assimilative effect). If the target is considered to be atypical, however, perceivers may be more likely to base their judgments on more specific (e.g., behavioral) criteria than on the group as a whole. Under these conditions, perceivers could also use the original category as a judgmental anchor, producing a contrast effect (e.g., Manis et al., 1988). As discussed earlier, a group member whose attributes are either moderately or highly dis- crepant from the group should be more likely to be judged as typical if the group is heterogeneous than if it is homogeneous.

If this is so, and if typicality acts in the manner specified by Fiske and Neuberg (1990), then perceivers' beliefs about the group should be more likely to have a positive (i.e., assimilative) effect on moderate and highly discrepant group members if the group is homogeneous than if it is heterogeneous.

The second explanation is somewhat more complex and rests on a few underlying assumptions. When perceivers can easily assess typicality, it seems likely that typicality plays the moder- 

ator role assigned to it by Fiske and Neuberg (1990). Suppose, however, that typicality is not easy to assess. If this is so, perceiv- 

ers may be relatively unlikely to rely on typicality as a basis for judgment. There are two (somewhat related) reasons why this might be so. In light of the "cognitive miser" view of social judg- 

ment (Fiske & Taylor, 1984), perceivers who find it relatively effortful to assess an attribute (e.g., typicality) may be less likely to rely on this attribute than perceivers who can assess it more easily. It could also be the case that perceivers who find it diffi- 

cult to assess typicality may simply regard this attribute as a relatively unimportant or meaningless feature and, for this rea- 

son, be unlikely to actually use it.

In any event, if perceivers do not use typicality as a basis for judgment, how might their beliefs about the group influence judgments of individual targets? One possibility is that perceiv- 

ers might independently assess the implications of the target's behavior and those of the group to which he or she belongs and add or average these implications into a final judgment (Anderson, 1981), producing a kind of assimilation effect. Note that this process does not necessarily involve a consideration of the target's discrepancy from the group at all. That is, if the person is known to be a member of the group, perceivers might simply apply their beliefs about that category (i.e., the perceived central tendency of the group) to judgments of that person, regard- 

less of his or her similarity to the other members of the group.

According to this line of reasoning, therefore, ease of assess- 

ing typicality can influence which of two different processes will guide use of the group as a basis for judgment. The first, typi- 

cality-based process, occurs when typicality is easy to assess.
In this process, beliefs about the group should have a positive (assimilative) effect on typical group members but no effect (or a contrast effect) on atypical group members. On the other hand, what I refer to in this article as an independent-attribute process (cf. Anderson, 1981) occurs when perceivers are not able to easily assess typicality. In this case, perceivers’ beliefs about the group should have a positive, assimilative effect regardless of the targets’ discrepancy from the group. The distinct pattern of effects that should arise from these two processes is depicted in Figure 2.

Suppose that the implications arising from Figure 2 are correct. If this is so, and (for reasons outlined earlier) typicality is easier to assess when the group is homogeneous, a prediction follows directly. Whereas the typicality-based process shown in the top panel of Figure 2 should be likely to arise when the group is homogeneous, the independent-attribute process should be more likely to occur when the group is heterogeneous. It is important to emphasize that it is not group variability per se that most directly determines the two different processes shown in Figure 2 but rather the ease with which perceivers can assess typicality.

Two aspects of the investigation by Lambert and Wyer (1990) become important to note at this juncture. When participants in that study were explicitly asked (at the end of the experiment) to judge the typicality of the target, perceptions of typicality did not significantly vary as a function of group variability. Thus, whereas these results appear to be inconsistent with the “typicality moderator” hypothesis, they are more compatible with the “cognitive effort” hypothesis. However, because Lambert and Wyer did not measure typicality reaction times, this hypothesis remains untested. Second, all participants in that study judged an outgroup member. Thus, although that study demonstrated the effects of group variability on judgments of an atypical outgroup member, the effects of variability on judgments of ingroup members remain to be examined. This point becomes important in the context of the results to be reported.45

The Present Research

This study investigated the preceding theoretical ideas by asking male and female participants to form an impression of a group of 50 persons on the basis of information provided about their performance on a test of perceptual-motor ability. The variability of group performance (high vs. low) was varied across condition and was identical to the two distributions displayed in Figure 1. In addition, the gender of the group (all men or all women) was also varied. The similarity of the participants’ gender to that of the group members provided the basis for designating the group as an ingroup or an outgroup. After forming an impression of the group, participants were asked to form typicality- and ability-related judgments of several individuals whom participants had not seen before but who ostensibly belong to the same group as the 50 persons described earlier. The discrepancy of these new targets from the group was either low, moderate, or high and corresponded to the three levels of target discrepancy as depicted in Figure 1. (A matched set of targets whose performance was below the mean was also presented.) Latency of response to queries about the targets’ typicality was assumed to provide a measure of the ease with which participants could assess typicality. The relation between participants’ beliefs about the group and judgments of the individual targets was assessed by examining the correlation between estimates of the groups’ perceptual ability and judgments of the targets with respect to this dimension (hereinafter, “group–target correlation”).

Summary of Predictions

Typicality Judgments

Hypothesis 1 (typicality judgment hypothesis). On the basis of the distribution of group performances in the high- and low-variability distributions (see Figure 1), participants should judge both moderate- and high-discrepancy targets as more typical (but low-discrepancy targets as slightly less typical) in the former than in the latter condition. This overall pattern should be reflected in a Group variability × Target discrepancy interaction.

Typicality Reaction Times

Hypothesis 2 (frequency hypothesis). Typicality should generally be easier to judge when the target’s performance occurs very frequently (or very infrequently) than if this is not the case. Two predictions follow directly from this premise: (a) Independent of group variability, typicality response times (RTs) should be faster for low- and high-discrepancy targets than for moderately discrepant targets. This should be reflected in a target discrepancy main effect. (b) Typicality RTs for all three targets should be faster when the group is homogeneous than when it is not. This should be reflected in a group variability main effect.

---

4 It should be noted that in Lambert and Wyer’s (1990) study, the attributes of the target (i.e., an unambiguously dishonest priest) were very highly discrepant from the group. It seemed conceivable that group variability might be more likely to influence the judged typicality of atypical targets whose attributes were relatively less extreme. As I show shortly, this was in fact the case.

5 Although a second study reported by Lambert and Wyer (1990) did focus on ingroups versus outgroups, perceived variability was (unlike in the first study) confounded with ingroup versus outgroup status.
Hypothesis 3 (cognitive organization hypothesis). Independent of group variability, participants should be faster to judge the typicality of outgroup members than ingroup members, and this should be true for all three targets. This should be reflected in an ingroup versus outgroup status main effect.

Effects of Participants' Beliefs About the Group on Judgments of Individual Targets

Hypothesis 4 (typicality moderator hypothesis). To the extent that moderate- and high-discrepancy targets are more likely to be judged as typical if the group is heterogeneous (cf. Hypothesis 1), participants' beliefs about the group should be more likely to have a positive (assimilative) effect on these persons if the group is heterogeneous than if it is not. This should be reflected in higher group-target correlations for high and moderate targets when the group is heterogeneous than when it is not. However, these effects should completely disappear once differences in judged typicality are partialed out.

Hypothesis 5 (cognitive effort hypothesis). The dual effects of group variability (Hypothesis 2b) and ingroup versus outgroup status (Hypothesis 3) should make typicality easiest to assess in the low-variability outgroup condition and hardest to assess in the high-variability ingroup condition. If this is so, then the former condition should be particularly likely to show evidence of typicality-based processing (positive group-target correlations for low targets but zero (or negative) correlations for moderate and high targets). However, the latter condition should be particularly likely to show evidence of independent-attribute processing (positive group-target correlations, regardless of target discrepancy).6

One additional prediction should be noted. As discussed earlier, Hypothesis 3 is based, in part, on the assumption that perceivers tend to focus on categorical attributes that distinguish themselves from others and vice versa (e.g., McGuire & McGuire, 1987). In the present study, therefore, participants were expected focus more attention on the gender of persons who belonged to the outgroup. In addition, characteristics that depart from cultural "default values" (i.e., attributes that are ascribed to other persons in the absence of individuating information) might attract more attention than attributes that do not depart from such values (Goffman, 1963). To the extent that "male" represents one such default value (E. R. Smith & Zarate, 1992), all participants also were expected direct more of their attention to the gender of the female group than to that of the male group.

Method

Participants and Design

A total of 334 University of Illinois undergraduates (158 men and 176 women) participated in partial fulfillment of a course requirement. The design consisted of four between-subject factors: (a) target group gender, (b) participant gender, (c) variability of the group (high vs. low), and (d) judgment order (i.e., typicality judgments preceded vs. followed by ability judgments of the individual new targets). The discrepancy of the individual target performances from the group (high, moderate, and low) as well as the favorableness of their performance (above vs. below group mean) was varied within participants.7

Group Judgment Phase

Initial instructions. Participants were informed initially that the general purpose of the experiment was to understand how individuals form perceptions of other groups of people. Participants were then told that they were to form an impression of a group of 50 persons on the basis of their performance on a test of perceptual-motor ability. Participants were told very little about the exact nature of the group and its members other than it was a "campus group" that had been contacted as part of an earlier investigation.

Presentation of group information. Information about the performance of the members of the target group was presented on a series of 50 cards. Each card provided information about the name and test score performance of one group member (e.g., "Joe: 9"). Depending on the condition to which participants were assigned, all 50 of these names were either unambiguously male or unambiguously female. The mean (and modal) test score performance of the 50 persons was identical (i.e., 8) across experimental condition, as was the range of these scores (minimum = 4 and maximum = 12). However, the variability of the scores was varied across condition. The distribution of scores in the high-variability (SD = 2.46) and low-variability (SD = 1.55) conditions was identical to the two distributions shown in Figure 1.

Dependent variables. After receiving information about the group, perceptions of group variability were measured by asking participants to generate a frequency distribution pertaining to the perceptual-motor abilities of the group. Participants were presented with a scale ranging from 0 (very low perceptual-motor abilities) to 10 (very high perceptual-motor abilities) and were asked to estimate the number of group members that they would assign to each of the 11 scale categories. Participants also judged how similar the 50 persons were to one another with respect to two dimensions that were related to the group's performance on the test (i.e., perceptual-motor abilities, intelligence) as well as two stereotypically masculine (assertive, opinionated) and two stereotypically feminine (sensitive, dependent) dimensions. These judgments were reported along a scale ranging from 0 (not at all similar) to 10 (extremely similar). After judging the variability of the group, participants made direct ratings of the group with respect to these same dimensions along a scale ranging from 0 (extremely low) to 10 (extremely high).

Individual Judgment Phase

After judging the group, participants were then informed that "we are also interested in the process by which people form impressions of sin-

6 One possible qualification of Hypothesis 5 should be noted. Suppose that participants are explicitly asked to infer the typicality of the targets before making descriptive judgments of them. Then, all participants might use typicality as a basis for these latter judgments, regardless of the ease with which that attribute could be computed. This possibility was explored in the present study by varying whether typicality judgments preceded or followed descriptive judgments of the target.

7 In addition to these factors, approximately half of the participants were given the test of perceptual-motor ability (and given false feedback regarding their performance on it) before receiving information about the group, whereas the other half of the participants did not. Analyses revealed that the feedback manipulation had no theoretically significant effects other than a general tendency for participants to use their own performance as a standard of comparison for judging the performance of the group and its individual members. More important, none of the effects reported in this article were significantly contingent on this variable. In addition, presentation time for the information about each of the group members was also varied (2 s vs. 4 s), but this factor had no effect whatsoever. The present results are thus collapsed over these two variables.
ingle individuals as well as groups.” Participants then read the following instructions:

In a few moments, we will be presenting you with some new information about several separate individuals. They are members of the same campus group to which the fifty persons belonged, but you have not seen information about these specific individuals before.

Nature of target information. All participants were presented with the first name and test performance information of six individuals on a computer screen. All targets were of the same gender as the 50 members of the group that participants had considered earlier and included two highly discrepant scores lying outside the range of scores achieved by members of the specific target group (3, 13), two moderately discrepant scores lying within the range of group scores (4, 12), and two typical scores centered around the mean group performance (7, 9). The presentation order of these six scores was counterbalanced. The scores compose a 3 x 2 design including the discrepancy of these scores from the mean (high, moderate, and low) and their favorableness (above or below the mean).

Dependent variables. Three dependent variables were of primary interest: (a) judgments of each target’s typicality, (b) the speed with which these judgments were made, and (c) descriptive (ability) judgments of the individual targets. Half of the participants in each condition completed trials in which they made typicality judgments of each of the six targets. Following these trials, participants were again presented with information about the same six targets (in the same order as in the initial block) and made descriptive judgments of each person. For the remaining participants, the order of making the two types of judgments was reversed.

To acquaint participants with the general nature of the speeded typicality task, participants first performed a practice task in which they judged the typicality of the height of four basketball players. Following this practice task, participants were told that “we are interested in your judgments of the typicality of performance of several individuals on the perceptual-motor task.” Participants were informed further that if they thought that “the person’s score is generally typical with how the members of the group perform,” they should press the “b” key (labeled typical), but if they thought that the person’s score is generally atypical, they should press the “m” key (labeled atypical). In the descriptive (ability) judgment task, participants were asked the following questions about each person: (a) “How would you judge this person’s perceptual-motor skills?” (b) “How would you judge this person’s overall level of cognitive abilities?” and (c) “How would you rate this person’s overall intelligence?” To answer each question, participants typed in a number along a scale from 0 (very low) to 10 (very high).

Salience of Gender

At the end of the experimental session, participants estimated how much “weight or importance” they attached to the fact that the group was all of the same gender when trying to form an impression of these persons along a scale from 0 (no weight/very low importance) to 10 (high weight/very high importance). Participants also were asked to think back to when you received information (on the deck of cards) about the 50 group members during the first part of the experiment. When, if at all, did you first notice that all the members of the group were all of the same sex?

Participants responded to this question by checking one of the following response alternatives: (a) almost immediately; (b) by the time I saw the first 10 names or so, (c) by the end of the deck, (d) while I was forming judgments of the group, (e) during the computer task, and (f) I only realized that they were all of the same sex at this moment (now). These responses were coded from 0 to 5, with higher numbers indicating earlier attention to and awareness of the group’s gender.

Scoring

To increase the reliability of measurement, the individual items pertaining to each of the relevant variables in this study were converted to z scores and averaged to provide a single index of the variable in question. Decisions to include the specific items for each composite were based on a priori theoretical considerations as well as the pattern of interitem correlations.

Perceptions of performance variability were based on an average of the z scores pertaining to (a) the standard deviation computed from the frequency distribution task and (b) similarity judgments with respect to judgments of the group’s perceptual-motor ability and intelligence. Perception of overall trait variability was based on an average of the similarity ratings with respect to the traits assertive, opinionated, sensitive, and dependent. Each participant’s perceptions of the mean ability of the group was based on an average of the z scores pertaining to (a) the mean of the frequency distributions that participants generated, (b) their rating of the group’s perceptual-motor ability, and (c) ratings of its intelligence. Perception of the group’s masculinity and femininity was based on an average of participants’ ratings of the group as assertive and opinionated, and sensitive and dependent, respectively. Because the two questions pertaining to participants’ perceived attention to the gender of the target group were highly correlated with one another, they were averaged to provide a composite measure of participants’ attention to and the importance of gender. Finally, participants’ overall assessment of the perceptual-motor ability of each of the individual targets was based on an average of all three of the ratings made of each person, namely, perceptual-motor skills, overall cognitive ability, and intelligence.

Results

Manipulation Checks

Judgments of group variability/central tendency. Judgments of variability were submitted to a 2 (group variability) X 2 (subject gender) X 2 (target gender) analysis of variance (ANOVA). Collapsed over all other variables, group members’ perceptual-motor ability was judged to be more variable in the high-variability condition (M = .42) than in the low-variability condition (M = -.43), F(1, 326) = 126.19, p < .001. Thus, the manipulation of actual group variability was successful. Similar results were obtained for perceptions of the target group with respect to their more general personality traits (Ms = .21 vs. -.22), F(1, 326) = 16.54, p < .001. No other main effects or interactions were significant (all ps > .20). In particular, the ingroup was not perceived as reliably more variable than the outgroup, and this was true both with respect to perceptual-motor ability as well as the more general personality traits. These null results are consistent with those of previous researchers who have obtained either relatively small outgroup homogeneity effects (Park & Judd, 1990) or no evidence at all for this effect (e.g., Linville et al., 1989) in the domain of gender. Perceptions of the average ability of the group was unaffected by manipulated variability (Ms = -.06 vs .06 for the high- and low-variability conditions, respectively, p > .20). Nor did these judgments vary as a function of participant gender, target gender, or these two variables in combination. Thus, the variability of the group was successfully manipulated without affecting
participants' perceptions of the group's average ability. This fact becomes critical in the context of subsequent analyses. Additional analyses revealed that the male group was rated higher with respect to stereotypically masculine traits, whereas the reverse was true for stereotypically feminine traits (both ps < .05). Although not surprising, these results confirm that participants were, on average, sensitive to the gender of the target group. As noted below, however, there were substantial differences in participants' awareness of this factor.

Gender salience  As anticipated, participants judged the gender of the outgroup to be more salient than that of the ingroup, (Ms = .12 vs. -.13), F(1, 326) = 5.21, p < .02. In addition, participants rated the gender of the female group (M = .15) as more salient than that of the male group (M = -.15), F(1, 326) = 9.76, p < .002. Additional analyses revealed that this latter effect was somewhat more pronounced in the high-variability condition (Ms = .19 vs. -.11) than in the low-variability condition (M = -.07 vs. -.23), although the interaction implied by this pattern was not significant (p > .20). Overall, these findings are consistent with the idea that perceivers will tend to notice those features that distinguish themselves from others (McGuire & McGuire, 1987; Park & Rothbart, 1982) or that depart from cultural defaults (Goffman, 1963).

Judgments of the individual targets. Not surprisingly, judgments of the targets' perceptual-motor ability were strongly determined by their specific test performance. Table 1 shows the pattern of means for the six targets, broken down by favorableness of performance and discrepancy from the overall group performance. Judgments of ability were obviously higher for targets whose scores were above the average performance of the group than for targets whose scores were not (p < .0001). Second, within each favorableness level, judgments increased in extremity as the performance scores became more discrepant from the group average (both ps < .0001). This increase in extremity was more pronounced when the scores were below the mean than when they were above. This effect, which is reflected by an interaction between favorableness of performance and its discrepancy from the group mean, F(2, 325) = 421.50, p < .0001, is consistent with previous evidence showing that negative information has a greater impact on judgments than does positive information (e.g., Birnbaum, 1974; Skowronski & Carlston, 1989). This finding does not qualify any of the results reported below.

Typicality judgments

Typicality judgments of each of the six targets were initially submitted to a loglinear analysis in which all possible contingencies of these judgments on group variability, participant gender, target gender, and judgment order were examined for each target separately. Aside from some four-way interactions that appeared to be devoid of theoretical interest, the only significant determinant of typicality to arise from these analyses was, as predicted, group variability. The contingency of typicality judgments on group variability is shown in the top panel of Table 2. To facilitate interpretation of these data, the percentage with which each level of performance actually occurred in the group is displayed as well.

According to Hypothesis 1, participants should have judged moderately and highly discrepant targets as more typical if the group was heterogeneous than if it was homogeneous, and this pattern should have been slightly reversed for the low-discrepancy targets. The pattern of categorical typicality judgments, displayed in Table 2, offers partial support for these predictions. As predicted, a greater proportion of participants judged the moderately discrepant targets as typical in the high-variability condition than in the low-variability condition, and this was true regardless of whether the performance of those targets was unfavorable (.20 vs. .08), χ²(1, N = 334) = 9.14, p < .01, or favorable (.40 vs. .21), χ²(1, N = 334) = 14.38, p < .01. In addition, a higher proportion of participants judged the highly discrepant and favorable performances as typical when the group was heterogeneous than when it was homogeneous, χ²(1, N = 334) = 5.56, p < .05, but no such effect occurred when the performance was unfavorable. Group variability had no discernable effect on the low-discrepancy targets.

Although these data provide only partial support for Hypothesis 1, it should be noted that although dichotomous measures were necessary for the reliable measurement of the RT data, simple yes versus no measures of typicality obviously provide a far less sensitive measure of perceived typicality than continuous measures. Therefore, to conduct a more sensitive test of Hypothesis 1, I asked a separate group of participants (N = 96) to judge the typicality of the six targets using a scale that ranged from 0 (not at all typical) to 10 (extremely typical). All other aspects of the procedure prior to this judgment task were identical to the main study. However, because the loglinear analyses indicated that ingroup versus outgroup status had no effects whatsoever on judged typicality, this factor was not manipulated in this study (all participants judged outgroup members). These continuous judgments are also displayed in Table 2.

These continuous typicality judgments were submitted to a Group Variability X Target Discrepancy X Target Favorableness repeated measures ANOVA. These analyses yielded strong support for Hypothesis 1. As predicted, a significant Group Variability X Target Discrepancy interaction emerged from these analyses, F(2, 93) = 7.88, p < .001, such that both the moderate- and high-discrepancy targets were judged as significantly more typical when the group was heterogeneous than

### Table 1

<table>
<thead>
<tr>
<th>Favorableness of performance</th>
<th>Target discrepancy</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>Moderate</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>Above the mean</td>
<td>8.86</td>
<td>8.42</td>
<td>6.92</td>
<td></td>
</tr>
<tr>
<td>(13)</td>
<td>(12)</td>
<td>(9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below the mean</td>
<td>2.39</td>
<td>3.25</td>
<td>6.08</td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>(4)</td>
<td>(7)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Scale ranges from 0 (very low ability) to 10 (very high ability). Numbers in parentheses refer to target's score on the test of perceptual-motor ability and ranged from 14 (highest possible) to 0 (lowest possible).
Table 2
Categorical and Continuous Typicality Judgments as a Function of Group Variability, Performance Discrepancy, and Performance Favorableness

<table>
<thead>
<tr>
<th>Target discrepancy and favorableness of performance</th>
<th>High</th>
<th>Moderate</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Favorable</td>
<td>Unfavorable</td>
<td>Favorable</td>
</tr>
<tr>
<td>Actual occurrence of score in group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High variability</td>
<td>0%</td>
<td>0%</td>
<td>10%</td>
</tr>
<tr>
<td>Low variability</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
</tr>
</tbody>
</table>

Type of judgment

Categorical (yes-no; n = 334)
High variability .17 .04 .40 .20 .95 .96
Low variability .08 .04 .21 .08 .94 .98
Difference .09 .00 .19 .12 .01 -.02

Continuous (0-10; n = 96)
High variability 2.36 2.09 3.68 3.96 7.38 8.06
Low variability 1.51 1.55 2.94 2.49 7.84 8.24
Difference .85 .54 .74 1.47 -.46 -.18

when it was not. This trend (as predicted) was slightly reversed for the low-discrepancy targets. Internal analyses confirmed that moderate-discrepancy targets were judged as more typical when the group was highly variable than when it was not (3.82 vs. 2.72), and this was true regardless of whether the performance was favorable (3.68 vs. 2.94) or unfavorable (3.96 vs. 2.49). F(1, 94) = 14.82, p < .001. Similarly, high-discrepancy targets were judged as significantly more typical when the group was highly variable than when it was not (2.23 vs. 1.53), and this was true for both favorable (2.36 vs. 1.51) and unfavorable (2.09 vs. 1.55) performances, F(1, 94) = 6.33, p < .01. In neither case were the effects of group variability contingent on target favorableness (both ps > .10). In contrast to the moderate- and high-discrepancy targets, the low-discrepancy targets were judged nonsignificantly less typical when group variability was high than when it was low (7.72 vs. 8.04).8

Typicality Response Times

Typicality RTs were initially evaluated on the basis of a six-way (Group Variability × Participant Gender × Target Gender × Judgment Order × Target Discrepancy × Target Favorableness) repeated analyses ANOVA in which all but the final two variables were between-subject variables. Although a number of significant effects emerged from these analyses, the effects of primary theoretical interest involved group variability, ingroup versus outgroup membership, and the discrepancy of the target performance from the group mean. RTs (in milliseconds) are shown as a function of these three variables in Table 3.9 (Additional effects arising from these analyses are considered shortly.)

According to Hypothesis 2 (frequency hypothesis), typicality RTs should have been faster when (a) the target’s performance was highly, or not at all, discrepant from that of other group members and (b) when the group was homogeneous. Both of these predictions were strongly confirmed. First, participants took less time to judge the typicality of both low-discrepancy (M = 1,086 ms) and high-discrepancy (M = 1,145 ms) targets than to judge moderate-discrepancy targets (M = 1,332 ms), F(2, 317) = 21.89, p < .001. Second, they took less time to judge the target’s typicality when the group was low in variability (M = 1,131 ms) than when it was high (M = 1,243 ms), F(1, 318) = 5.07, p < .03. Overall, these results support the proposition that typicality is generally easier to assess when the target’s attribute—or one similar to it—occurs either very frequently or very infrequently.

8 The extent to which participants judged favorable performances as more typical than unfavorable performances (or vice versa) was contingent on both target discrepancy and group variability. When the group was variable, favorable performances were judged as less typical for both the moderate- and low-discrepancy targets, whereas this pattern was reversed for the high-discrepancy targets. When the group was low in variability, however, favorable performances were judged as more typical than unfavorable performances for moderate-discrepancy targets, less typical for low-discrepancy targets, and equaled typical for high-discrepancy targets. This highly complex pattern was reflected in a Group Variability × Target Discrepancy × Target Favorableness interaction, F(2, 93) = 4.21, p < .02. This interaction was not expected and is very difficult to interpret. In any case, it does not qualify the implications of the analyses just presented.

9 Extremely slow or extremely fast response times were assumed to indicate that participants were either not concentrating on the experimental task or were responding so quickly that they did not have time to adequately consider the question being asked. Therefore, a conservative criterion was adapted such that if the response latencies were either less than 200 ms or more than 2 standard deviations from the mean for that particular target, the value was replaced with the participant’s mean response time to questions about the remaining targets. Such extreme response times composed less than 10% of the typicality judgments made in the experiment. A separate set of analyses performed on a logarithmic transformation of the raw latency scores produced a pattern of results very similar to that presented below.
CONSEQUENCES OF GROUP VARIABILITY

Table 3
Typicality Reaction Times (in Milliseconds) as a Function of Group Variability, Ingroup/Outgroup Membership, and Discrepancy of Target Performance From Group Mean

<table>
<thead>
<tr>
<th>Target group and ingroup vs. outgroup membership</th>
<th>Group variability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
</tr>
<tr>
<td>High discrepancy</td>
<td></td>
</tr>
<tr>
<td>Ingroup</td>
<td>1,191</td>
</tr>
<tr>
<td>Outgroup</td>
<td>1,195</td>
</tr>
<tr>
<td>Moderate discrepancy</td>
<td></td>
</tr>
<tr>
<td>Ingroup</td>
<td>1,424</td>
</tr>
<tr>
<td>Outgroup</td>
<td>1,419</td>
</tr>
<tr>
<td>Low discrepancy</td>
<td></td>
</tr>
<tr>
<td>Ingroup</td>
<td>1,185</td>
</tr>
<tr>
<td>Outgroup</td>
<td>1,038</td>
</tr>
<tr>
<td>All targets (average)</td>
<td></td>
</tr>
<tr>
<td>Ingroup</td>
<td>1,266</td>
</tr>
<tr>
<td>Outgroup</td>
<td>1,217</td>
</tr>
</tbody>
</table>

According to Hypothesis 3 (cognitive organization hypothesis), participants should have been faster to judge the typicality of outgroup members than ingroup members. This prediction was also strongly confirmed. Specifically, typicality RTs were faster when the target belonged to the outgroup ($M = 1,127$ ms) than to the ingroup ($M = 1,246$ ms), $F(1, 318) = 4.86, p < .03$. There appeared to be some variation in this difference as a function of both group variability and discrepancy of target performance (see Table 3). However, the effect of ingroup versus outgroup membership on RTs was not contingent on any other factor (all $p$s > .10). Independent of targets' status as ingroup or outgroup members, the typicality of female targets was judged faster than that of male targets in the high-variability ($Ms = 1,193$ vs. $1,300$ ms) but not the low-variability ($Ms = 1,155$ vs. $1,104$ ms) condition. $F(1, 318) = 4.53, p < .03$, for the relevant two-way interaction.

Note that these effects of ingroup versus outgroup membership and target gender exactly parallel the effects of these variables on the salience of the group's gender reported earlier in this article. Specifically, recall that participants judged the gender of the group as more salient if its members (a) belonged to the outgroup than if they belonged to the ingroup and (b) were women than if they were men, especially in the high-variability condition. The reason for the parallel contingency of both the salience and RT data on group variability and target gender is somewhat unclear. In any event, the data clearly show that the conditions under which the gender of the target group was salient were also the conditions under which typicality RTs were relatively fast. The correlation between salience ratings and RTs was in the direction implied by this relation, although its magnitude was surprisingly trivial ($r = -.06$). I return to this pattern of findings later in this article.

One additional aspect of these data should be noted. As seen in Table 3, the dual effects of group variability and ingroup versus outgroup status led to RTs that, for all three classes of targets, were clearly faster in the low-variability outgroup condition than in the other three conditions. This means that participants in that particular experimental condition were able to most easily assess whether the individual targets were typical or atypical. This aspect of the data becomes important in the context of the group–target correlations, to be reported presently.

Supplementary analyses. Overall, RTs were similar regardless of whether targets were judged as typical ($M = 1,148$ ms) or atypical ($M = 1,185$ ms; $p > .15$), and none of the effects reported above were contingent on typicality choice (all $Fs < 1.0$). Thus, the effects of group variability and ingroup versus outgroup status on RTs was not contingent on whether targets were actually judged as typical or atypical. RTs were slower for favorable target performances ($M = 1,273$ ms) than for unfavorable performances ($M = 1,102$ ms), $F(1, 318) = 31.53, p < .001$. This effect is probably attributable to the tendency, noted earlier, for participants to view unfavorable performances as more extreme than favorable performances and, thus, more easily assessed as atypical. This difference was considerably more pronounced for the moderate-discrepancy targets ($Ms = 1,486$ vs. $1,178$ ms) than for either the low-discrepancy ($Ms = 1,120$ vs. $1,051$ ms) or the high-discrepancy targets ($Ms = 1,212$ vs. $1,078$ ms), $F(2, 317) = 6.93, p < .001$. These findings do not compromise any of the conclusions drawn earlier.

Effects of Group-Based Beliefs on Judgments of Individual Targets

A final issue concerned the effect of participants' beliefs about the group on their judgments of single group members. Operationally, the relation between judgments of the group and the individual targets was assessed by examining the correlation between estimates of the group's general perceptual-motor ability and judgments of the individual targets along this same dimension (i.e., group–target correlation). The main goal in examining these data was to both (a) replicate previous findings obtained by Lambert and Wyer (1990) in the domain of natural stereotypes and (b) gain greater insight into the processes underlying these seemingly paradoxical effects.

In an initial examination of these matters, average group-target correlations were computed for each of the three classes of low-, moderate-, and high-discrepancy targets across each of the four combinations of group variability and ingroup versus outgroup status. The pattern of these correlations is displayed in Table 4. Consider first the pattern of results for the outgroup. The first thing to note about the pattern arising here is that it is consistent with the basic finding obtained by Lambert

---

10 In addition to these effects, two 3-way (Group Variability $\times$ Participant Gender $\times$ Target Favorableness and Group Variability $\times$ Ingroup vs. Outgroup Status $\times$ Target Favorableness) interactions emerged from these analyses. These interactions were unexpected and difficult to interpret. Neither of these effects qualifies any of the conclusions drawn above.

11 When target performances were favorable and either moderately or highly discrepant from the group, ability judgments were very negatively skewed ($\sim$-2.4 and $\sim$-2.7, respectively), whereas the distributions for the other targets, as well as participants' beliefs about the group, were normally distributed (all $skews < 11$). Before these group–target correlations were calculated, therefore, a power transformation was applied to these two targets to adjust for this difference.
and Wyer (1990). As noted earlier, Lambert and Wyer found that when participants were forming impressions of outgroup members, participants were more likely to use the group as a basis for judging an atypical target if the group was heterogeneous than if it was homogeneous. A similar pattern arose here, although the magnitude of the effect is much less pronounced than that obtained in Lambert and Wyer’s investigation. As seen in the top half of Table 4, group–target correlations for the moderate-outgroup targets were positive and significant when the group was heterogeneous but were close to zero when the group was homogeneous (rs = .21 vs. .07). A similar difference arose for the high-discrepancy target (rs = .15 vs. -.03).

As for the underlying processes at work here, the overall pattern of results appears to be at least somewhat more consistent with Hypothesis 5 than Hypothesis 4. According to Hypothesis 4, any effects of group variability on group–target correlations should be completely eliminated when judged typicality is partialed out. However, the pattern of group–target correlations after judged typicality was partialed (shown in the bottom half of Table 4) is virtually identical to the zero-order correlations. On the other hand, recall that typicality RTs in the low-variability outgroup condition were consistently faster than in the other three conditions (see Table 3). According to Hypothesis 5, therefore, the pattern of group–target correlations in this condition should resemble the pattern of typicality-based processing, whereas group–target correlations in the other three conditions should resemble independent-attribute processing (cf. Figure 2). In fact, the overall pattern of correlations matches reasonably well with this prediction.

A more direct way of examining the moderating role of ease of assessing typicality is to examine the pattern of group–target correlations as a function of whether typicality RTs were relatively fast or slow, collapsing over experimental variables. Separate median splits were thus performed on RTs to each of the individual targets, collapsing over group variability and ingroup versus outgroup status. This separated participants into groups of fast and slow RTs for each target and thus divided participants according to the relative ease or difficulty with which they assessed the typicality of that individual. Following these median splits, average group–target correlations were computed (after judged typicality was partialed out) for each of the three classes of target discrepancy. The results of these analyses were consistent with the implications of Hypothesis 5, although the differences in correlations were again relatively small. When participants were relatively fast in assessing typicality, the correlation for the high-discrepancy targets was positive and significant for the low-discrepancy targets but was consistently smaller for the moderate- and high-discrepancy targets (rs = .26, .10, and .10, respectively). When participants found assessment of typicality more difficult, however, group–target correlations were actually identical for the low- and moderate-discrepancy targets (both rs = .28), but they were somewhat lower for the high-discrepancy target (r = .19).12

Although the overall pattern of results is consistent with the (much stronger) findings reported by Lambert and Wyer (1990) in the natural stereotype domain, the magnitude of these correlations (as well as the differences between them) were relatively small. In isolation, therefore, these data do not permit any definitive conclusions to be reached regarding the support for Hypothesis 5 (or the lack of support for Hypothesis 4). Nevertheless, when considered along with the pattern of results reported by Lambert and Wyer, the available evidence provides reasonable support for the idea that the effect of group variability on perceivers’ reactions to discrepant group members is moderated by the ease of assessing typicality rather than by the effects of judged typicality per se. As for the small effects obtained here, this may have been due, in part, to the relative strength of artificial versus natural stereotypes, an issue which I consider in more detail below.

Discussion

Group variability was predicted to have at least three main consequences for how perceivers form judgments of single
group members. The present results generated support for several distinct hypotheses that bear on these consequences and yielded numerous insights into the process by which perceivers judge and react toward single group members. The present results may be summarized as follows:

**Judged typicality.** The probability that a given group member was judged as typical was determined not only by his or her actual characteristics per se but also by the variability of the group. Persons whose attributes were moderately or extremely discrepant from the group were more likely to be judged as typical if the group was heterogeneous than if it was not.

**Ease of assessing typicality.** Three factors had independent effects on the ease and speed of assessing typicality. Typicality RTs were faster (a) if the target’s performance was close to the group mode or far removed from it, (b) if the group was homogeneous than if it was heterogeneous, and (c) if targets belonged to the outgroup.

**Reliance on the group as a basis for judging individual group members.** The ease with which participants could assess typicality was, in turn, predictive of whether they relied on the group as a basis for judging persons whose attributes were discrepant from those of the group. Relatively fast typicality RTs suggest that participants were able to easily discern that the moderate- and high-discrepancy targets were indeed atypical group members but that the low-discrepancy target was typical. Hence, these participants tended to use the group as a basis for judging only the latter, but not the former, class of targets. This kind of process was most evident in the low-variability outgroup condition, when typicality RTs were fastest. On the other hand, relatively slow typicality RTs suggest that participants found the typicality of the targets to be a relatively unimportant or meaningless attribute, and, hence, these participants tended to use their beliefs about the group as a basis for judgment independent of the target’s discrepancy from the group.

Although the overall magnitude of the group-target correlations (as well as the differences between them) was relatively small, several kinds of evidence in combination suggest that the kinds of effects described above were driven by the ease of assessing typicality rather than judged typicality per se. For one thing, the overall pattern of correlations across experimental condition showed a reasonably close parallel between the ease of assessing typicality and the extent to which participants used the group as a basis for judging moderately and highly discrepant group members (see Table 4). Similar implications arose from the median split analyses. On the other hand, partialing out judged typicality had no effect whatsoever on the overall pattern of correlations. These conclusions are also consistent with the fact that, as noted earlier in this article, judged typicality played no role whatever in the findings reported by Lambert and Wyer (1990).13

**Factors Influencing the Ease of Assessing Typicality**

The present results suggest that group variability had an indirect, rather than direct, effect on the process by which perceivers formed impressions of single group members. That is, this effect was most directly driven by the ease with which participants could assess typicality. However, group variability was by no means the only factor that influenced the ease of assessing typicality. Specifically, typicality was easier to judge if the targets belonged to the outgroup rather than the ingroup and if they were female rather than male (at least under conditions of high group variability). Theoretically, such effects were the result of differences in the way that participants organized information about the target group.

The current investigation provides no direct evidence that the cognitive organization of information about the target group varied across experimental condition. Nevertheless, several studies have shown that perceivers are more likely to store information about the group as a whole when forming impressions of outgroups than of ingroups (Ostrom et al., 1993; Ostrom & Sedikides, 1992; Park & Rothbart, 1982). Judgments of typicality are presumably easier to assess if perceivers have already stored this kind of summary information about the group than if they have not. To this extent, the fact that participants were significantly faster to judge the typicality of outgroup members than ingroup members is consistent with the implications of this literature. Considerably less research has examined the cognitive organization of information about men versus women. However, the same factor that presumably leads perceivers to store information about outgroups at the category level (i.e., greater salience of their category membership) is characteristic of women as well. The close parallel between the salience of the (shared) gender of the group and typicality RTs is consistent with this line of reasoning. On the other hand, although a factor common to both outgroups and women other than the one noted above cannot be entirely dismissed, its nature is not at all clear.

A somewhat related interpretation is that when the gender of the group was particularly salient, participants could have used this attribute as an organizing category around which they could store information about the group (cf. Rothbart, Fulero, Jensen, Howard, & Birrell, 1978). Although all participants were told that the target persons belonged to the same campus group, very little was actually said about the nature of this group, and thus its nature was likely quite vague. On the other hand, gender is a personally relevant dimension along which to categorize a group of individuals and thus could have facilitated participants’ attempts to organize information about the group. Although the pattern of gender salience precisely paralleled RTs, one weakness in this line of reasoning is that the correlation between gender salience and response times was trivial. This null finding is most likely attributable to the fact that the current study used a relatively crude, off-line estimate of perceived (but not actual) attention to the target group gender. The mea-

---

13 One might still wonder at this point why, then, group variability did not influence the judged typicality of the target in Lambert and Wyer’s (1990) study even though it apparently influenced the extent to which participants used the person’s typicality as a basis for judgment. As noted earlier, however, the target’s attributes in that experiment were very atypical. In combination with the present results, these findings suggest that there may be somewhat different thresholds for the two different classes of effects of group variability discussed in this article. That is, even though a target’s behaviors may be too extreme for group variability to exert any effect on typicality judgments per se, it apparently is still possible for this factor to affect the underlying importance of or participants’ sensitivity to the target’s discrepancy from the group.
sure used here evidently was sufficiently sensitive to reveal mean differences in participants' attention to, for example, the outgroup versus the ingroup. However, it evidently was not sufficiently sensitive to be used as a measure of individual differences in attention, which is how this measure was used in the correlational analyses.

On the other hand, some recent models (e.g., Hintzman, 1986; Kahneman & Miller, 1986) apparently would not predict ease of assessing typicality to vary as a function of whether perceivers have summarized or organized information about the group as a whole (see footnote 3). Nevertheless, it is somewhat unclear how these models would explain the fact that participants in this experiment were, in fact, faster to judge the typicality of outgroup members than ingroup members in ways that did not involve differences in the cognitive organization of these two kinds of groups. In any event, it is important to emphasize that differences in the way that people organize information about categories, and the role of attention in these representations, is currently a matter of considerable theoretical debate (e.g., Kraus et al., 1993; Linville & Fischer, 1993; Ostrom & Sedikides, 1992). Thus, further research involving direct manipulation of attention as well as measures that more directly tap the underlying representation of social categories (e.g., Ostrom et al., 1993) are obviously needed to more firmly establish the validity of the interpretation offered above.

General Discussion

This article began with the highly intuitive and widely accepted idea that beliefs associated with heterogeneous categories ought to have weaker effects on social judgment compared to homogeneous categories. By now, I hope to have convinced the reader that this view is not necessarily correct. Nevertheless, conclusions regarding the effects of group variability on social judgment may, of course, depend on the specific category under consideration by the experimenter (cf. Wyer & Srull, 1989, for a relevant discussion). Issues of generalizability are particularly relevant when, as in the present investigation, researchers rely on artificial (laboratory-induced) groups. That is, the precision and control gained by this methodology may potentially lead to the creation of processes that have little relation to the mechanisms that occur outside of the laboratory. To this extent, the value of the present conceptualization lies not only in its ability to account for the results reported in this article but also in the extent to which it can provide a theoretical framework for the results obtained in other settings (and with other types of categories) as well. In fact, the present conceptualization is able to account for several other findings obtained by previous researchers across a wide spectrum of natural stereotypes and experimental methodologies. The relation of this conceptualization to this earlier work is considered below.

Implications for Previous Models of Category-Based Inferences

Previous research and theory suggests that a group member's "fit" or similarity to the group as a whole determines the kind of effect that group stereotypes can have on judgments of that person (Brewer, 1988; Fiske & Neuberg, 1990). For example, previous research has often demonstrated that when a group member's attributes are not consistent with the group, participants' beliefs associated with the category (e.g., the group stereotype) tend to have either null (or contrast) effects on judgments of that person (e.g., Fiske, Neuberg, Beattie, & Milberg, 1987; Lord, Lepper, & Mackie, 1984).

How can the present results (in which participants tended to use the group as a basis for judging atypical group members under certain conditions) be reconciled with these earlier studies and theoretical models? This apparent contradiction is, in fact, easily resolved. The majority of studies in the stereotyping literature have investigated participants' beliefs about outgroups rather than ingroups. Furthermore, the members of these outgroups typically were persons with whom participants were likely to have had relatively little contact and thus were likely to be perceived as homogeneous. This includes, for example, the voluminous research on (White) participants' impressions of racial and ethnic minorities. As the data from the present study suggest, homogeneous outgroups are precisely the kinds of categories in which the discrepancy of the group member's attributes from the group stereotype should play the strongest role. Therefore, rather than competing with previous stereotyping models, the present conceptualization has established some boundary conditions of when these models may, and may not, apply. As implied by previous models, group stereotypes may indeed have very little direct effect on persons whose attributes are atypical of the group. However, this may only be true when perceivers form judgments of homogeneous outgroups.

Implications for Previous Investigations of Group Variability

The conclusion that participants are more likely to use their group as a basis for judging atypical targets when the group is heterogeneous than when it is not is consistent with the findings of at least two recent investigations. As noted earlier, the present findings are consistent with previous findings by Lambert and Wyer (1990), who found that participants' beliefs about a particular category (Catholic priests) had a strong assimilative effect on judgments of an unambiguously atypical target when the group was heterogeneous but not when the group was homogeneous. Although they did not directly measure group variability, conceptually similar results were recently obtained by Lord et al. (1991). When participants were either highly familiar with the group (or had a high degree of knowledge about its members), their attitudes about the group had a positive effect on judgments of single group members, regardless of whether they were typical or atypical. In contrast, when participants were not familiar with the group (or were not knowledgeable about its members), the effect of group-based attitudes was highly contingent on typicality. High levels of expertise and familiarity have been cited as conditions under which the members of a category should be perceived as relatively heterogeneous (e.g., Linville et al., 1989). To this extent, the findings reported by Lord et al. (1991) are precisely in line with the present conceptualization.

The relation between the present research and results obtained by Manis et al. (1988) is also worth noting. Briefly,
Manis et al. manipulated whether participants’ stereotype of the level of pathology in a particular hospital was either extreme (e.g., consisting of all pathological stimuli) or more moderate (e.g., consisting of a mixture of both pathological and non-pathological stimuli). Results showed that whereas “new” (midscale) exemplars tended to be contrasted from extreme stereotypes, assimilation effects arose when the stereotype was moderate. Although they did not discuss their results in terms of group variability, the results obtained by Manis et al. are consistent with the finding that participants in this study were more likely to judge a known group member as typical when the group was variable than when it was not. Nevertheless, it should be noted that (unlike in the present study) Manis et al. confounded the manipulation of group variability with at least two other group parameters (extremity of central tendency and range). As Manis et al. themselves noted, this makes it unclear the extent to which their results are driven by the discrepancy of test items from the central tendency of the group or from the similarity (or ranking) of the test items relative to the known exemplars in the group.

The present research not only conceptually replicates and extends the findings obtained by these previous investigations but also clarifies a number of important ambiguities left unresolved by that research. For one thing, the present study (unlike these previous studies) experimentally manipulated the variability of the group while holding all other factors constant (e.g., central tendency and range of the group’s attributes, participants’ familiarity with and liking for the group, perceived complexity). Moreover, the present investigation measured not only whether participants judged the targets as typical but also the ease with which they could make this judgment. This methodology allowed me to clarify the precise mechanism by which group variability influences how perceivers process information about single group members. Indeed, the results of this investigation indicate that group variability plays a less direct role in moderating the use of the group stereotype than I assumed in my previous research (Lambert & Wyer, 1990).

At the same time, the convergence of the present results with those obtained under more natural settings minimizes the possibility that the findings reported here are due to some idiosyncratic features of the particular methodology or stimulus materials used. For example, the perceived (or actual) distribution of group members’ attributes outside of the laboratory may often be different than the kind of unimodal and normal distributions used in this study (Nisbett & Kunda, 1985). Moreover, the information one acquires outside the laboratory about the attributes of category members is usually conveyed piecemeal over a long period of time (rather than in a single judgmental setting). Nevertheless, the fact that similar findings obtain when these artificial constraints are not present makes these liabilities of the present study less serious than they appear at first.

**Some Caveats**

Although the findings that emerged in the present research appear to be fairly generalizable, at least two limitations are worth noting. First, the theoretical considerations at stake here required me to restrict the number of attribute dimensions that participants considered when forming their impressions of the target group and its members. Outside of the laboratory, however, it is far more common to encounter persons or objects whose features have implications for several different, interrelated attributes. The processes that guide perceptions of typicality are likely to be considerably more complicated under these conditions (Rosch & Mervis, 1975).

Second, although one of the obvious benefits of the laboratory-induction procedure used here lies in the degree of precision and control over the characteristics of the target group, one drawback of artificial stereotypes is that they are, not surprisingly, generally not as strong as natural stereotypes. In particular, a recent study by Manis, Paskewitz, and Cotler (1986) showed that assimilative effects produced by laboratory-based stereotypes were much weaker than those produced by natural stereotypes, a finding that could, in part, account for the discrepancy between the size of the group–target correlations reported here and the much stronger effects reported by Lambert and Wyer (1990). A related point is that the kind of information that was provided about the targets in this study (i.e., their performance on a novel test) had relatively few evaluative implications and therefore was relatively unlikely to have influenced participants’ affective reactions to these persons or the group to which they belonged. Indeed, results reported by Lambert and Wyer suggest that participants who have strong affective reactions to homogeneous groups can use the target’s typicality as a positive or negative attribute in its own right.

**Characteristics of Social Categories and Their Impact on Social Judgment**

A central concern among stereotype researchers has been on the variables that might reduce the impact of negative stereotypes on social judgment. As Linville et al. (1989) and others have noted, increasing one’s contact with and knowledge about the group may increase perceptions of variability, that is, lead to the recognition that the members of the group are not all alike. Several authors have suggested that this process might reduce the extent to which perceivers rely on stereotypes as a basis for judgment (Linville et al., 1989; Pettigrew, 1981). An intriguing, and perhaps disturbing, implication of the present findings is that (all else being equal) the process of coming to see the group as heterogeneous could actually increase, not decrease, the extent to which perceivers apply their beliefs about the group to single group members. That is, perceivers who view the group as heterogeneous are likely to pay less attention to whether a known group member does or does not fit the group stereotype, and for this reason, be more likely to simply apply their beliefs about the group (i.e., perceived central tendency) to that person.

Of course, group variability represents only one of the many parameters along which categories can differ, and thus other factors could have important effects on the extent to which perceivers use their beliefs about the group as a basis for judgment. For example, as perceivers become more knowledgeable about the group, they may come to view its members in more cognitively complex terms (i.e., higher in cross-dimensional variability). Research by Linville (1982) showed that affective reactions to persons belonging to complex categories are less extreme than reactions to persons belonging to categories with
relatively simple cognitive organization. Increasing knowledge about the group could also influence the extremity of the perceived central tendency of the group which, in turn, could influence the extent to which perceivers use the target's group membership as a basis for judgment. For example, the average man is perceived as more aggressive than the average woman, but this difference is not as great as the perceived aggressiveness of construction workers versus housewives. To this extent, a person's gender is likely to have a weaker effect on perceptions of that individual's aggressiveness than varying his or her membership in these two occupational categories (Krueger & Rothbart, 1988).

More generally, a rapidly growing number of studies have demonstrated that stereotyping effects are likely to depend on a wide variety of factors, such as the processing capacity at the time of judgment (e.g., Bodenhausen & Wyer, 1985; Gilbert & Hixon, 1991), the type and amount of information available about the group member (Hilton & Fein, 1989; Krueger & Rothbart, 1988; Locksley, Hepburn, & Ortiz, 1982), the goal of the perceiver to be accurate (Neuberg & Fiske, 1987), and mood (Bodenhausen, Kramer, & Susser, 1994). Thus, the variables under consideration in the present project concern only a small subset of the universe of variables that could potentially play a role in the use and disuse of stereotypes in social judgment. Further examination of the role of group variability in the context of these other factors may well be a profitable route of future investigation.

References


Received August 12, 1993
Revision received April 18, 1994
Accepted April 20, 1994 •

MEMBERS OF UNDERREPRESENTED GROUPS: REVIEWERS FOR JOURNAL MANUSCRIPTS WANTED

If you are interested in reviewing manuscripts for APA journals, the APA Publications and Communications Board would like to invite your participation. Manuscript reviewers are vital to the publication process. As a reviewer, you would gain valuable experience in publishing. The P&C Board is particularly interested in encouraging members of underrepresented groups to participate more in this process.

If you are interested in reviewing manuscripts, please write to Leslie Cameron at the address below. Please note the following important points:

- To be selected as a reviewer, you must have published articles in peer-reviewed journals. The experience of publishing provides a reviewer with the basis for preparing a thorough, objective review.
- To select the appropriate reviewers for each manuscript, the editor needs detailed information. Please include with your letter your vita. In your letter, please identify which APA journal you are interested in and describe your area of expertise. Be as specific as possible. For example, "social psychology" is not sufficient—you would need to specify "social cognition" or "attitude change" as well.
- Reviewing a manuscript takes time. If you are selected to review a manuscript, be prepared to invest the necessary time to evaluate the manuscript thoroughly.

Write to Leslie Cameron, Journals Office, American Psychological Association, 750 First Street, NE, Washington, DC 20002-4242.