Title

On-line and memory-based processes in group variability judgments

Authors
Mackie, DM
Sherman, JW
Worth, LT

Publication Date
2024-01-20

Peer reviewed

Social Cognition, Vol. 11, No. 1, 1993, pp. 44-69

ON-LINE AND MEMORY-BASED PROCESSES IN GROUP VARIABILITY JUDGMENTS

DIANE M. MACKIE, JEFFREY W. SHERMAN, AND LEILA T. WORTH
University of California, Santa Barbara

Are group variability judgments made in an on-line or memory-based fashion? In a first experiment addressing this question, subjects made judgments of a group's variability, a judgment intended to be on-line (liking), or a judgment intended to be memory-based (religiousness). Variability judgments were made more slowly than on-line judgments and at the same speed as memory-based judgments. Independently of this effect, in-group variability was judged more rapidly than out-group variability. In addition to replicating these results, a second experiment using a minimal group paradigm demonstrated that the amount of similarity information recalled predicted both the latency and extremity of variability judgments. We discuss the implications of our conclusion that variability judgments are predominantly memory-based for models of group variability judgments.

The term Out-Group Homogeneity Effect (OHE) refers to perceivers' tendency to see their own groups as relatively more heterogeneous than groups to which they do not belong (Quattrone & Jones, 1980). The perception of out-groups as homogeneous has been shown to contribute to both the development and maintenance of out-group stereotyping and discrimination (see Linville, Salovey & Fisher, 1986; Park, Judd, & Ryan, 1991, for reviews). Such a role warrants extensive research attention, and the articles in this special issue are testament to social psychology's sustained interest in understanding the antecedents and consequences of the OHE. Unfortunately, its causal underpinnings are still not well understood.

Perhaps the greatest promise of progress in this regard comes from

Experiment 2 was completed as part of the third author's doctoral dissertation at the University of California, Santa Barbara. The support of NIMH grant No. MH430141 to the first author and of a National Science Foundation Graduate Fellowship to Jeffrey W. Sherman is gratefully acknowledged. We thank Holly Schroth for her assistance in conducting this research.

Correspondence concerning this article should be addressed to Diane M. Mackie, Department of Psychology, University of California, Santa Barbara, CA 93106-9660.

recent attempts to specify the information on which in-group and out-group variability judgments are based (see Messick & Mackie, 1989, for a review). In the research we describe in this article, we used such recently developed models to derive predictions about the processes by which group variability judgments are made. We begin by briefly reviewing four different models that have been proposed to explain the OHE. We then draw out the processing implications of these models, particularly with regard to the issue of whether judgments of in-group or out-group variability are made in an on-line or in a memory-based fashion (Hastie & Park, 1986). The findings of two experiments that bear on this issue are then described. Finally, we discuss the implications of our results for the various mechanisms proposed as explanations for the OHE.

MODELS OF GROUP VARIABILITY JUDGMENTS UNDERLYING THE OHE

THE EXEMPLAR FAMILIARITY MODEL

One of the best specified models of how variability judgments are made has been proposed by Linville and her colleagues (Linville, Fischer, & Salovey, 1989; Linville, et al., 1986). The model, which we will call the exemplar familiarity model (EFM), is loosely based on exemplar-based models of category representation (Hintzman, 1986, 1988). According to the EFM, judgments regarding group variability are based on a set of instances or exemplars that are retrieved from long-term memory when the judgment is called for. In addition to individual category members serving as exemplars, multiple feature sets that have been previously abstracted about certain members of a group—typically called subtypes—may also contribute to variability judgments (Linville et al., 1986). Judgments made about the group as a whole that derive from previous experience or social learning may also be stored in long-term memory and used as exemplars in subsequent judgments. Although the EFM acknowledges the usage of such group-level information in later judgments, information based upon abstraction from many cases has no special role and is given no greater weight in subsequent group-level judgments compared to information about a single or individual exemplar.

The EFM is a strictly memory-based model. Group-level variability judgments (indeed, any group judgment) are calculated and subsequently stored *only* when a group-level judgment is explicitly requested. The sole basis for group variability judgments in this model is the set of exemplars that are retrieved when the judgment is requested.

Within the parameters of this model, Linville and her colleagues propose that differences in perceived in-group and out-group variability reflected in the OHE are the result of differential familiarity with the in-group and out-group. As familiarity with a group increases, the number and variety of encountered exemplars also increase. Because more exemplars are retrieved when judging the in-group than the out-group, the in-group will be judged to have more variability. Thus, from the perspective of the EFM, the combination of greater familiarity with the in-group and a strictly retrieval-based judgment process accounts for the OHE.

What evidence supports the EFM of variability judgments? Computer simulations of the retrieval mechanisms posited by this model (Linville et al., 1989) have demonstrated that greater familiarity with a group (as operationalized by exposure to individual exemplars) in combination with a retrieval-based judgment process is *sufficient* to create the OHE. In addition, Linville and her colleagues have demonstrated increased variability judgments as familiarity with groups increases over time and greater differentiation among in-group than out-group members (Linville et al., 1986). However, there is little evidence that the sheer number of group members known is related to increased perceived variability in a straightforward way (Jones, Wood, & Quattrone, 1981; Linville et al., 1986). In fact, OHEs have been found in conditions in which more exemplar information was retrieved about the out-group than about the in-group (Judd & Park, 1988).

THE DUAL PREDICTOR MODEL

A second model of variability judgments, the dual predictor model (DPM; Kashima & Kashima, this issue), is adapted from Tversky's (1977) model of similarity judgments. According to Tversky's model, the similarity of two exemplars is dually determined by the number of features the two exemplars share and the number of distinct features each exemplar possesses uniquely. As the number of shared features increases, similarity increases; as the number of distinct features increases, similarity decreases.

Kashima and Kashima suggest that group variability judgments are essentially judgments of the similarity among group members. Moreover, they argue that such judgments are made in an explicitly retrieval-based fashion. When a group variability judgment is called for, available group members are retrieved and their similarities and differences are compared. Based on the overall numbers of similarities and differences that occur in the exemplars retrieved, an overall estimate of group variability is formed. Although the DPM is a retrieval-based model, the relation between the number of exemplars retrieved and perceived group variability is not monotonic. From the perspective of the DPM, variability depends not so much on the raw number of exemplars retrieved, but on the similarities and differences present in the retrieved exemplars.

According to the DPM, the OHE results from differential utilization of similarity and difference information in judgments of in-groups and out-groups. This might occur for four different reasons: 1) there is equal similarity and difference information pertaining to in-group and out-group members but in-group difference information is more salient than out-group difference information; 2) there is equal information but out-group similarity information is more salient than in-group similarity information; 3) in-group members possess more difference information than out-group members; or 4) out-group members possess more similarity information than in-group members.

The DPM was recently tested by Kashima and Kashima (this issue). As predicted, increased similarity information decreased perceived group variability and increased difference information increased perceived variability independently. In addition, these effects were stronger when salience was manipulated, but primarily because of the impact of difference information. Increases in salient similarity information did not decrease variability perceptions. Although these studies provide support for some aspects of the DPM, they do not speak to whether variability judgments were made on-line or in a memory-based fashion.

THE ABSTRACTION PLUS EXEMPLAR MODEL

A third model of variability judgments, proposed by Judd and Park (1988; Park & Judd, 1990; Park et al., 1991), is loosely based on Fried and Holyoak's (1984) category density model of category representation. In contrast to pure retrieval-based models such as the EFM and the DPM, what we term the abstraction-plus-exemplar model (AEM) suggests that both on-line abstraction and exemplar-retrieval processes contribute to variability judgments. The primary source of variability judgments about groups is variability information abstracted on-line (Park & Hastie, 1987). As new information is obtained, the variability of the group along different dimensions is spontaneously and continually updated. In addition to storing these constantly updated group-level judgments, however, particular exemplars that are encountered may also be stored. When a variability judgment about a group is requested, the most recently updated summary estimate of group variability can be retrieved and utilized to make this judgment; it is not necessary to re-compute a variability estimate by retrieving specific instances from long-term memory. However, when a variability judgment is re-quested, subjects may also retrieve specific judgment-relevant exem-plars in addition to the stored summary judgment. When this occurs, these instances may be used to update the summary judgment. Thus, both abstract summaries and specific instances are available for usage in making variability judgments, although the former plays a more central role.

According to the AEM, the OHE results from different kinds of in-group and out-group information being encoded and/or retrieved. In-group judgments are more likely to rely on both abstracted group-level and specific exemplar-level information, whereas out-group judgments are more likely to rely on abstracted group-level information alone. The greater use of specific exemplars in making judgments about the in-group compared to the out-group is responsi-ble for the OHE. In fact, even if there is *equal* information about the in-group and out-group, the OHE may still occur because group-level and exemplar information is *used* differently in in-group and out-group judgments.

Park and her colleagues (1991) suggest several reasons why exem-plar information should play a bigger role in in-group as compared to out-group judgments. First, there may be greater accuracy motivation for in-group judgments than out-group judgments, so that more exemplars are encoded and more exemplars are used to "check" or update group-level summary judgments in the former than the latter case. Second, different levels of experience or familiarity with the in-group may mean that more in-group than out-group exemplars are available. Third, differential use of the self as an exemplar of the in-group can influence in-group but not out-group judgments. Overall, then, the AEM predicts that variability judgments can be made in a completely on-line fashion. If subjects do retrieve exemplars in addi-tion to the abstracted group-level judgment, the AEM suggests that out-group judgments will be made in a relatively on-line fashion (relying on the previously-abstracted, group-level information alone), whereas in-group variability judgments will be comparatively mem-ory-based (with more exemplars available for, and used in, retrieval and judgment).

Evidence in support of the AEM's contention that variability judg-ments are formed on-line comes largely from failures to find relations between retrieved exemplars and variability judgments under condi-

tions predicted by retrieval-based models (Judd & Park, 1988; Park & Hastie, 1987). More specific support for the AEM comes from findings that the variance of a set of retrieved exemplars predict variability judgments for the in-group, but not for the out-group (Park & Judd, 1990). This suggests, as the AEM predicts, that in-group variability judgments rely more on retrieved instances than out-group judgments. However, when the effects of the retrieved exemplars were removed, the OHE remained intact, suggesting that the OHE is not dependent on this difference. In addition, there has been little direct evidence that using the self as an exemplar completely explains the OHE (Park & Judd, 1990).

THE FREQUENCY DISTRIBUTION MODEL

A new model proposed by Park and Judd and their associates (Kraus, Ryan, Judd, Hastie, & Park, this issue), the frequency distribution model (FDM), offers yet another account of the processes underlying variability judgments and the OHE. According to the FDM, people often spontaneously create mental frequency distributions that sum-marize the number of group members with particular attributes at different levels of various dimensions. For example, subjects may store the number of high, moderate, and low intelligence behaviors per-formed by group members, or the number of smart, stupid, or averagely bright individuals in the group. These frequency distribu-tions are not themselves variability estimates. Rather, when a variabil-ity judgment is required, subjects retrieve the distributions and base their variability judgments largely on the number of levels—labeled subtypes in the model—used to discriminate among group members along an attribute continuum. Thus, a group represented by five different levels (or subtypes) of intelligence will be judged as more variable than a group represented by three different categories of intelligence. While the formation of frequency distributions does occur on-line, the formation of variability judgments is memory-based, relying on retrieval of the distributions, levels, and numbers associated with levels.

According to the FDM, the OHE occurs because in-group members are spontaneously classified along a greater number of subtypes on any dimension. This presumably occurs because, for a number of reasons, people are more interested in the differentiating and individu-ating details of in-group members than of out-group members (Park & Rothbart, 1982). Thus, in-group members may be classified into five different levels of intelligence, and out-group members may be classified into only three levels of intelligence. This would result in higher judgments of in-group variability than out-group variability.

Kraus et al. (this issue) report some initial supportive evidence for their model. Analysis of protocols produced when subjects "thought aloud" while receiving information about a group's SAT scores suggested that subjects spontaneously kept count of the number of scores that fell into certain numerical ranges (e.g., 400–500, 500–600). Similarly, it appeared that subjects spontaneously classified and counted behaviors as reflecting high, medium, and low values on relevant trait dimensions. In support of the FDM, subjects' variability judgments were affected by the number of levels subjects used to classify group members. The more levels or subtypes used, the higher the variability judgments made. In a second study, subjects instructed to generate subtypes produced more in-group subtypes than out-group subtypes, a finding consistent with the FDM.

ON-LINE AND MEMORY-BASED PROCESSES IN GROUP VARIABILITY JUDGMENTS

The models proposed as explanations for the OHE differ on a number of dimensions, and findings that definitively support or refute particular models have been elusive. Nor do we claim to seek such evidence here. However, there is one dimension on which the proposed models differ markedly that invites a relatively straightforward empirical test: The models differ in the degree to which they assume variability judgments are computed in an on-line or memory-based fashion. Thus, evidence as to whether variability judgments are made as exemplars of the group are encountered, or whether group-level variability estimates depend on later exemplar retrieval and calculation, would lend support to some conceptualizations of the OHE and detract from others.

In the research reported here we examined the degree to which variability judgments are made in an on-line or memory-based fashion by utilizing two types of evidence: the speed with which relevant judgments can be made and the relationship between judgment and recall (Hastie & Park, 1986; Lichtenstein & Srull, 1987; Mackie & Asuncion, 1990). Our focus in Experiment 1 was on response latencies. To assess the various models of group variability judgments, we compared the time it took subjects to make in-group and out-group variability judgments with the time it took them to make other judgments assumed to result from either on-line or memory-based processes.

EXPERIMENT 1

If group variability judgments are made on-line, we expected them to be made relatively rapidly, whereas slow latencies would suggest the occurrence of memory-based processing. Both the EFM and the DPM argue for retrieval-based calculation of variability judgments, and thus slow latencies would be supportive of the mechanisms proposed in these models. The FDM also relies on memory-based processes, and thus would predict slower latencies than if variability judgments had been precomputed (the model is unclear about whether retrieving frequency distributions would facilitate variability judgments relative to retrieving "raw" exemplars). On the other hand, the on-line abstraction of group variability information proposed by the AEM argues for relatively rapid response latencies. As an additional refinement, the AEM predicts relatively fast response times for out-group judgments and relatively slow ones for in-group judgments. A similar prediction might be made from the EFM and FDM. To the extent that there are more exemplars or subtypes to retrieve and combine for the in-group, responses to in-group variability judgments should be relatively retarded.

One problem in interpreting latencies, of course, is to know what constitutes a fast or a slow response time. Although we intended to compare differences in in-group and out-group judgments, there was no prior research available to guide our interpretations of what could be termed a fast or slow (and by implication an on-line or memory-based) variability judgment in any absolute sense. Our strategy in the first experiment was therefore to compare judgments of variability to judgments we believed to be made in either an on-line or a memory-based fashion. We selected liking judgments as an on-line comparison. Evaluative judgments of liking are frequently made spontaneously as information about a group or individual is encountered (Hastie & Park, 1986). Thus, we expected subjects to make judgments of how much they liked the group as information was encountered. As a memory-based comparison judgment, we asked some subjects how religious the group they had read about was. As none of the behavioral items that subjects saw made reference to religious behaviors, and as it was unlikely that subjects would spontaneously judge a group of males or females on this dimension, we anticipated that subjects' only means of responding to the item would be to try to retrieve relevant behavioral items and make a memory-based judgment. By comparing the speed with which subjects made judgments of similarity about the in-group and the out-group with the speed with which they made liking and religiousness judgments about the same groups, we hoped to learn more about whether similarity judgments were made in an on-line or memory-based fashion.

SUBJECTS AND DESIGN

Sixty-seven male and 74 female members of a University of California at Santa Barbara (UCSB) introductory psychology course received partial course credit for their participation. Subjects were randomly assigned to read information about either male or female targets and to make judgments about the group's similarity, likeability, or religious proclivities.

PROCEDURE

Subjects were told that they would be reading about behaviors performed by two different groups—one of males and one of females—and would be asked to answer some questions about each of the groups after seeing the relevant items. Subjects were instructed in the use of the 9-point rating scale, so that they would be able to answer those questions as quickly and accurately as possible. When subjects understood their task, they pressed a key to begin presentation of the items.

Manipulation of Target Group. Half of the subjects saw behavioral items about females first, thinking they would then see items about males. The rest of the subjects saw items about males first, thinking they would later see items about females. In fact, subjects made judgments about only one group before the experiment ended. Presentation of the behavioral items was prefaced by the label, FEMALE GROUP or MALE GROUP, as appropriate.

Presentation of Behavioral Items. Subjects saw 28 different sentences, each of which described an activity engaged in by a single group member (these items constituted the high variability set of behaviors used by Park & Hastie, 1987). Half the items described activities related to sociability and half described activities that pertained to intelligence. The items were presented in random order, each remaining on the computer screen for 5 seconds. Each item was prefaced by a common male or female first name, as appropriate for the target condition.

Measurement of Liking, Similarity, and Religious Judgments. As soon as the last item had been presented, subjects responded to the dependent measures. Approximately one-third of the subjects were asked "How likeable are the members of this group?" Subjects responded by choosing a number between 1 and 9, where 1 was labeled "not likeable" and 9 was labeled "very likeable." A second group of subjects were asked "How religious are the members of this group?" and responded on a 9-point scale where 1 was labeled "not religious" and 9 was labeled "very religious." The final third of the subjects were asked "How similar are the members of this group to one another?" Subjects responded by choosing a number on a 9-point scale labeled "not similar" at the low end and "very similar" at the high end. Responses and response latencies were recorded automatically. Subjects were debriefed and thanked.

RESULTS AND DISCUSSION

Judgments of Liking, Similarity, and Religiousness. Subjects' responses to the first question they were asked after seeing the behavioral items were analyzed in a 2 (male and female Subject Gender) x 2(male or female Target Gender) x 3 (similarity or likability or religious judgment) analysis of variance (ANOVA). Not surprisingly, subjects responded differently to the three different questions, $F(2,129) = 60.64$, $p < .0001$. Of more interest, this effect was qualified by a marginally significant three-way interaction involving Subject Gender and Target Gender, $F(2,129) = 2.43$, $p < .09$. Separate analyses were performed to see if in-group and out-group differences in similarity, liking, and religiousness were obtained.

As expected, analysis of the similarity judgments indicated that male and female subjects made different judgments about how similar the members of male and female target groups were to one another, $F(1, 44) = 4.11$, $p < .05$. Males saw female targets as more similar to one another ($M = 5.00$) than they did male targets ($M = 4.00$), whereas females saw male targets as more similar ($M = 4.58$) than they did female targets ($M = 3.75$). These responses thus reflected a typical out-group homogeneity effect—members of groups to which the subjects did not belong were seen as more similar than members of groups to which they did belong.

Analysis of ratings of likeability revealed no effects: When males and females judged the likeability of their own and other groups, there was no sign of in-group bias. Males rated both groups as more religious ($M = 4.28$) than did females ($M = 3.28$), $F(1,42) = 4.04$, $p < .05$; otherwise, there were no intergroup effects for judgments of how religious the groups were.

Latency of Liking, Similarity, and Religious Judgments. Our main focus of interest in this experiment was the speed with which subjects made the various judgments about the in-group and the out-group. To compare responses to the liking, religiousness, and similarity judgments, the speed with which male and female subjects made ratings about the male and female group's likeability, religiousness, and similarity were analyzed. The first result of interest was a main effect for the type of judgment made, $F(2,129) = 8.74$, $p < .0003$. Post-hoc Bonferroni comparisons indicated that likeability judgments were made more quickly ($M = 6.94$ seconds) than either similarity judgments ($M = 9.13$ seconds, $t(94) = 4.45$, $p < .001$) or religiousness ratings ($M = 9.00$ seconds, $t(92) = 3.42$, $p < .002$). Similarity and religiousness ratings did not differ. The relative speed with which liking judgments were made, and the relative slowness with which religiousness judgments were made, was consistent with our expectation that the former would be made in an on-line fashion, whereas the latter would be made in a memory-based fashion. These results also suggested that the similarity judgment was more like the putative memory-based judgment than like the putative on-line judgment.

Males and females also made judgments about males and females at different speeds, $F(1,129) = 4.07$, $p < .05$. In general, females made faster judgments about females ($M = 7.59$ seconds) than about males ($M = 9.24$ seconds), whereas males made slightly faster judgments about males ($M = 8.17$ seconds) than about females ($M = 8.41$ seconds). This tendency to make faster judgments about the in-group was qualified, however, by a significant three-way interaction involving Subject Gender, Target Gender, and Judgment, $F(2,129) = 3.06$, $p < .05$. Separate analyses were performed to better understand these effects.

Analysis of the speed with which liking judgments were made revealed no significant effects. Liking judgments were made relatively quickly overall, and the speed with which judgments about the in-group and the out-group were made did not differ.

Analysis of the speed with which similarity judgments were made revealed an interaction between Subject Gender and Target Gender, $F(1,44) = 3.98$, $p < .05$. Results appear in the left panel of Figure 1. Females made faster judgments about females ($M = 7.94$ seconds) than they did about males ($M = 10.39$), whereas males made faster judgments about males ($M = 8.73$ seconds) than they did about females ($M = 9.58$ seconds). Similarity judgments were made relatively slowly overall, but were made more quickly for the in-group than for the out-group.

A similar but weaker pattern emerged in the speed with which religious judgments were made, $F(1,42) = 3.42$, $p < .07$. Again, females made judgments more quickly about female targets ($M = 7.61$ seconds) than about male targets ($M = 10.56$ seconds), whereas males made

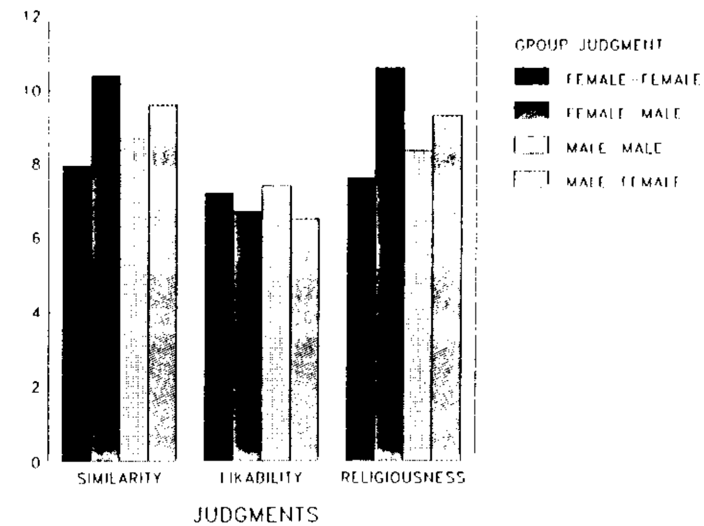

FIGURE 1
Mean latencies for in-group (female-female and male-male) and out-group (female-male and male-female) similarity, liking, and religiousness judgments, Experiment 1.

faster judgments about male targets ($M = 8.35$ seconds) than about female targets ($M = 9.27$ seconds). The pattern of similarity judgments was identical to the pattern of religious judgments: Both judgments were made relatively slowly, and both judgments were made more quickly for the in-group than for the out-group (see Figure 1).

The results of the first experiment suggested that variability judgments are not abstracted on-line, but are memory-based. Several aspects of our data converge on this conclusion. First, it took subjects longer to make variability judgments than to make liking judgments about the groups. Of course, in making this comparison we followed previous research in assuming that liking judgments were made on-line. If this assumption is true, it is hard to see why variability judgments take longer than evaluative judgments unless one concludes that the former are made in a memory-based fashion. However, even if the assumption about liking judgments is false, it is hard to see why it would take less time to calculate an evaluative than a variability judgment (if both judgments were memory-based), or why it would take longer to retrieve a previously calculated variability judgment than it would to retrieve relevant exemplars and calculate a liking judgment (if liking judgments were memory-based and variability judgments on-line). The most parsimonious explanation is thus that variability judgments are made in a memory-based fashion, or at least have more memory-based components.

Second, this conclusion is strengthened by the finding that judgments of variability took about the same amount of time as did judgments of the group's religious proclivities. Given little evidence that college students spontaneously abstract religious information as they learn about groups, and given that none of the information presented pertained explicitly to religious behaviors, it seems likely that these judgments were made in a memory-based fashion. If variability judgments were in fact made on-line, it is hard to imagine why retrieving a previously calculated variability judgment would take as long as reviewing retrieved behaviors and calculating a group-level judgment about religion.

Thus, these results suggest that variability judgments were made in more of a memory-based than on-line fashion. To this extent, they are more consistent with the retrieval-based mechanisms of the EFM, DPM, and FDM than with the AEM's proposal that variability judgments are made on-line.

Despite the fact that variability judgments about both in-group and out-group were made relatively slowly, in-group variability judgments were made significantly more quickly than out-group judgments. This suggests possible processing differences in making in-group and out-group variability judgments, but not of the type suggested in the AEM. This model would have predicted slightly longer processing times for in-group judgments, which according to the model are more likely to be checked and updated by exemplar retrieval. Nor would faster in-group judgments be consistent with predictions from the EFM or FDM: the greater number of instances or subtypes available for retrieval from the in-group should, if anything, have increased rather than decreased judgment latency. However the fact that in-group variability judgments were made more quickly than out-group judgments is compatible with the idea that in-group exemplars are more easily retrieved. This notion is explicit in the AEM and implicit in the EFM; no doubt it would be easily integrated into the FDM.

EXPERIMENT 2

If in-group and out-group variability judgments were made in a memory-based fashion, we would expect to find some relation between the exemplars recalled and the variability judgments made (Park & Hastie, 1987). In Experiment 2 we wished to assess the relation between

judgment and recall, as well as attempt to replicate our findings with regard to variability response times. Because our response time data supported both the EFM and the DPM as viable memory-based models of the OHE, we attempted to distinguish them by assessing the relation between variability judgments and both the sheer amount of information retrieved (relevant to the EFM) and the different types of information retrieved (relevant to the DPM). Specifically, we examined the correspondence between variability judgments and the retrieval of information about in-group and out-group similarities and differences.

The second experiment also allowed us to examine variability judgments in a minimal group situation. The first experiment utilized gender groups, about which subjects no doubt had preconceptions. Although we presented equal information about in-groups and out-groups, and although we asked subjects to make judgments only about the group members they saw, it is nevertheless possible that differences in the speed with which in-group and out-group variability judgments were made were due to greater knowledge about, familiarity with, or accessibility of in-group exemplars. In the second experiment, therefore, we used a minimal group paradigm to investigate the OHE. By allowing us to control the amount and nature of in-group and out-group information presented, the minimal group paradigm allowed us a more controlled comparison of predictions from the various models.

SUBJECTS AND DESIGN

One hundred and forty UCSB undergraduates (49 males and 91 females) participated in the experiment for partial fulfillment of an introductory psychology course requirement. Subjects, all of whom were strangers to one another, participated in the experimental sessions in groups of 8 to 10.

PROCEDURE

Group Membership Manipulation. Subjects' performances on an initial embedded figures task provided the basis for ostensible categorization into groups. While their performance on the initial task was supposedly scored, subjects completed a Subject Profile Sheet on which they indicated, among other things, their favorite TV show, their intended major, their favorite sport, a creative behavior they had performed, a sociable behavior, and a responsible behavior. Subjects thus believed that this kind of information had been collected about each of the people in the experimental session.

Subjects were then told that performance on the embedded figures task allowed most people to be classified as "Grounds" (whose perception is strongly influenced by backgrounds rather than figures) and "Figures" (whose perception is more influenced by figures than by backgrounds). In every session four people were told that they were Grounds and four were told that they were Figures, although assignment was in fact random. If more than eight subjects were present, the additional one or two were told that they were unclassifiable.

Presentation of In-group and Out-group Information. Subjects were then told that the purpose of the experiment was to see if the perceptual differences between Grounds and Figures affected perceptions in more subjective domains, such as judgment making and impression formation. After a 10-minute filler task, Figures and Grounds participated in an impression formation task. They were seated at visually isolated computer terminals and given some general instructions about how to use the computer keyboard to respond to questions.

Subjects were told they would be seeing some of the information that the other members of their own group had volunteered on the Subject Profile Sheets, as well as some information that some of the members of the other group had volunteered. Subjects were asked to form an impression of their in-group and of the other group based on this information.[1]

Each subject saw information about his or her three other in-group members (labeled with actual names and sexes and group classification) and about three out-group members (again labeled with actual names, gender, and group label). Each person was described by six stimulus sentences ostensibly from the Subject Profile Sheets, each of which remained on the screen for 7 seconds. In fact, these descriptions were manipulated so that half of the subjects saw in-group members described as having similar preferences and different traits, and the out-group members described as having different preferences and similar traits.[2] For the other half of the subjects, in-group members were described as having different preferences and similar traits, and out-group members were described as having similar preferences and different traits. Half of the subjects saw information about in-group members first, whereas the other half saw out-group information first. Presentation of the 18 sentences about each group was prefaced by the warning "Now you will see information about the GROUND (FIGURE) group."

1. Subjects were also asked to think about the group as a whole or to think about each member of each group as information was presented. This manipulation had no effect on any dependent variable and reported analyses collapse across this variable.

2. Group member descriptions were formed in the following way. First, a total of 36 sentence predicates that were either expressions of preferences within three different domains (TV shows, majors, and sports) or behaviors that exemplified one of three traits (creativity, sociability, and responsibility) were composed. There were six predicates for each preference domain and six predicates for each trait. "Similar" subsets were then composed of preferences or behaviors with high pairwise similarity ratings on pretesting, whereas "different" subsets were composed of preferences or behaviors that had low pairwise pretest ratings (the latter were, in addition, dissimilar to the preferences or behaviors in the similar subset). Items in each subset were equally desirable and equally diagnostic. Using these subsets of stimulus predicates, descriptions of six target people were created. Each target was assigned a predicate from each preference and each trait domain. Three target people were assigned items so that they had similar preferences but differed on the traits, and three target people were assigned different preferences but similar traits.

DEPENDENT MEASURES

Effectiveness of Group Manipulation. To assess the effectiveness of the minimal group procedure in inducing feelings of group membership, subjects were asked how much they thought they would like the members of each group if they got to know them better.

Variability Measures. After reading the information about both groups, subjects considered each group in turn (in the same order as the information had been presented) and responded to the questions: "How similar do you think the members of your group (the other group) are to one another?"; "How similar do you think the members of your group (the other group) are in terms of PERSONALITY?"; and "How similar do you think the members of your group (the other group) are in terms of THE ACTIVITIES THEY LIKE TO DO?" Ratings of the in-group and out-group on these three measures were averaged to form a similarity index for each group (in-group *alpha* = .59; out-group *alpha* = .58). Response latency to these three measures was automatically recorded and also averaged to form a response time index (in-group *alpha* = .54; out-group *alpha* = .71).

Recall. Subjects were given 4 minutes to recall information about the in-group and an additional 4 minutes to recall information about the out-group, in counterbalanced order. After probing for suspicion, the experimenter thanked and thoroughly debriefed the subjects.[3]

3. The data from six subjects (5 males and 1 female) was dropped from all analyses because they expressed some suspicion about the veracity of the group membership manipulation.

RESULTS AND DISCUSSION

Effectiveness of Group Manipulation. In-group bias in group evaluations was taken as evidence that the minimal group manipulation was successful. Subjects indicated that they would prefer to get to know the other in-group members ($M = 6.71$) over the out-group members ($M = 6.42$), $F(1,126) = 4.6$, $p < .04$.

Group Variability Judgments. Analysis of the similarity index revealed that out-group members were seen as more similar to one another ($M = 5.84$) than in-group members ($M = 5.57$), $F(1,126) = 8.68$, $p < .004$. Thus, the OHE was obtained with mere categorization into minimal groups and when equal information about the in-group and out-group was provided. Group members who knew neither the in-group nor out-group members before the experiment, who did not expect to interact with either group in the future, and who received equal information about each group, perceived the out-group to be more homogeneous than the in-group. This result was not affected by any other manipulation, and did not depend upon a competitive inter-group orientation, as was suggested by Judd and Park (1988).

Latency of Variability Judgments. Because our hypotheses focused on whether variability judgments were spontaneously made during presentation of the group information, we examined the time subjects took to respond to the variability questions the first time they were posed.[4] Analysis of the variability judgment latency index revealed that subjects responded more rapidly to variability questions about the in-group ($M = 10.79$) than to questions about the out-group ($M = 13.48$), $F(1,126) = 18.83$, $p < .0001$, replicating the results from Experiment 1. In addition, response times were equal to or slower than the time taken to answer variability questions about the in-group and out-group in the first experiment.[5]

4. Subjects responded to the variability questions first about one group and then about the other. Thus only the first set of questions could be considered to measure whether variability judgments had been spontaneously made. Analysis of latencies to answer both sets of questions revealed that subjects responded more rapidly overall to variability questions about the in-group ($M = 8.91$) than to such questions about the out-group ($M = 10.55$), $F(1,126) = 13.37$, $p < .0004$. Not surprisingly, latencies to answer the variability question when it was asked for the first time were slower ($M = 12.09$) than when the variability question was asked for the second time ($M = 7.31$), $F(1,126) = 215.92$, $p < .0001$. These effects were both qualified by their interaction, $F(1,126) = 10.33$, $p < .002$. In-group responses were faster than out-group responses for the initial inquiry, as noted in the text. When the variability question was asked for the second time in-group variability judgments were slightly but not significantly faster ($M = 7.02$) than judgments about the out-group ($M = 7.62$). These results are consistent with the idea that the second set of judgments was cued by, or relied upon, the first set of judgments.

Recall. Recall was coded as belonging to either the "similar" or "different" subsets of information about each group. Greater recall of information about the out-group ($M = 3.7$) compared to the in-group ($M = 3.4$; $F(1,125) = 5.37$, $p < .02$) and greater recall of similarity ($M = 3.86$) than difference ($M = 3.23$; $F(1,125) = 7.77$, $p < .006$) information was qualified by their marginal interaction, $F(1,125) = 2.56$, $p < .10$. In fact, more information from the similar subsets was recalled about the out-group ($M = 4.12$) than about the in-group ($M = 3.62$), $F(1,125) = 7.62$, $p < .007$. However, this difference was itself qualified by an interaction between order and group, $F(1,125) = 15.10$, $p < .0002$, indicating that it was significant only when out-group information came first and not when in-group information came first. A comparison of recall for the in-group and out-group when they were both described with the same information indicated that in both cases recall for the out-group was greater, although not significantly so. Though qualified by an unexpected interaction, these results suggest that subjects paid more attention to the ways in which out-group members were similar to each other than the ways in which in-group members were similar to each other.

With regard to the difference information, subjects tended to recall more information about the out-group when the in-group came first ($M = 3.12$ and 3.59, for in-group and out-group respectively), and more information about the in-group when the out-group came first ($M = 3.31$ and 2.92, for in-group and out-group respectively), $F(1,120) = 5.14$, $p < .03$.

Relation Between Recall and Perceived Group Homogeneity. Variability judgments about the in-group and the out-group were regressed on the total amount of information retrieved about each group. In-group similarity was marginally predicted from recall of information about the in-group, $F(2,130) = 3.72$, $p < .06$ ($b = .07$), whereas recall of out-group information clearly predicted out-group similarity, $F(2,130) = 15.46$, $p < .001$ ($b = .17$). Although these results seem consistent with the EFM, further analyses indicated that the picture was more complicated. When subjects' variability ratings for the in-group and out-group were separately regressed on their recall of similarity and difference information about the groups, both analyses yielded significant results: for in-group homogeneity, $F(4,128) = 3.12$, $p < .01$, and for out-group homogeneity, $F(4,128) = 5.77$, $p < .003$. As can be seen in Table 1, however, these significant results resulted in both cases from

5. Because of the borderline *alphas*, separate analyses with the three similarity measures and the latency of the three measures as repeated measures were also performed. None of the three variability nor the three latency measures differed from one another.

TABLE 1
Prediction of Perceptions of Group Homogeneity from Recall, Experiment 2

	TYPE OF INFORMATION RECALLED			
	ABOUT THE IN-GROUP		ABOUT THE OUT-GROUP	
	SIMILAR	DIFFERENT	SIMILAR	DIFFERENT
In-group homogeneity	.16*	-.03	.02	.07
Out-group homogeneity	.08	-.05	.23**	.09

Note. Numbers are *b*s.
*$F(4,128) = 9.03, p < .003$
** $F(4,128) = 17.03, p < .0001$

the fact that the amount of similarity information recalled predicted overall perceptions of homogeneity.[6] This relation was present for both in-group and out-group judgments, adding further support to the idea that these variability judgments were made in a memory-based fashion. Difference information, however, had no predictive effect.

Relation Between Recall and Variability Latency. To further bolster the argument that variability judgments were made in a memory-based fashion, the number of items of similarity and difference information recalled about each group was regressed on the time it took subjects to make the appropriate variability judgment. If judgments are memory-based, we might expect that the more items recalled, the longer the judgment would take. In fact, such a relation was present for both judgments. The number of similarity items recalled about the in-group tended to predict the time it took to make in-group variability judgments, $b = .29$, $F(2,130) = 3.34$, $p < .07$. The number of similarity items recalled about the out-group significantly predicted latency to make the out-group variability judgment, $b = .48$, $F(2,130) = 3.87$, $p < .05$. Recall of difference information did not predict variability latency for either the in-group or the out-group.

Together, the pattern of recall and regression results provide an explanation for the OHE obtained in this experiment. Retrieved similarity information predicted variability judgments. Thus, to the extent that perceivers retrieved more similarity information for the out-group than for the in-group, the out-group was judged to be less variable than the in-group. The results from the second experiment again conflict with on-line models such as the AEM. Our finding that the OHE can occur with minimal groups—with whom the subjects had equal familiarity—and can be produced for the very reason that more "exemplar-like" information is retrieved about the out-group, is also damaging to the EFM. The fact that retrieved similarity information predicted variability judgments provides partial support for the DPM, although retrieved difference information did not affect perceived variability. Although recall of exemplar information is not strictly relevant to the FDM, the finding that the number of similar items recalled predicted the time taken to make variability judgments is not entirely consistent with the idea that such judgments depend on already summarized frequency distributions.

6. In-group similarity information was a slightly but not significantly better predictor of in-group homogeneity than of out-group homogeneity. Out-group similarity information was a significantly better predictor of out-group homogeneity than of in-group homogeneity, $t(133) = 4.41$, $p < .01$.

GENERAL DISCUSSION

Two findings emerged consistently from the two experiments reported here. First, subjects took a considerable time to report variability judgments, which resembled other memory-based, rather than other on-line, judgments. Second, judgments about in-group variability were made faster than judgments about out-group variability. We discuss the implications of each of these findings in turn.

HOW ARE VARIABILITY JUDGMENTS MADE?

The results of both experiments support the role of retrieval-based mechanisms in variability judgments. Experiment 1 demonstrated that variability judgment latencies were significantly longer than latencies to make liking judgments, which were assumed to be made on-line. In addition, variability latencies closely resembled those produced by subjects making what we believed to be another memory-based judgment: judgments of the group's religious proclivities. Both these findings suggest that group variability judgments are made in a memory-based fashion, rather than being abstracted on-line. Of course, we chose religiousness as a comparison judgment because we felt it highly unlikely that such a judgment would be made on-line; in addition, the difficulty of retrieving or evaluating exemplars on this dimension (none of our stimulus sentences were directly relevant to this trait) would make memory-based judgments particularly slow. Thus, it was particularly instructive that variability judgments were made equally slowly. Our strategy in the first experiment was to assess whether variability judgments were on-line or memory-based by comparing them to "marker" judgments—liking and religiousness. Although variability judgments were more like religious judgments and less like liking judgments, this conclusion would be further bolstered by directly manipulating whether subjects expected to make variability judgments before exposure to group information.

Experiment 2 provided additional evidence that variability judgments are memory-based. First, lengthy variability latencies were also obtained in this experiment, in which the OHE was produced in a minimal group situation. In addition, we found a relation both between retrieved exemplar information and variability judgments and between retrieved exemplars and latencies. This experiment differed from previous research in that we examined the extent to which variability judgments were predicted by retrieved similarity and difference information, rather than by the sheer amount of information retrieved. Retrieved similarity information—but not retrieved difference information—predicted both the latency and extremity of judgments of variability for both the in-group and the out-group, lending further weight to our interpretation of variability judgments as memory-based. Because recall measures always followed collection of variability judgments, it is possible that those similarity judgments may have guided greater recall of similarity information. However, there are two reasons to think this unlikely. First, the number of similarity items recalled predicted how long it took to make the variability judgments. Second, although manipulations of the order in which recall and judgments are assessed would resolve this issue more definitively, such manipulations in the impression formation domain have been found to make little difference (Hastie & Park, 1986).

It is possible that our findings reflect only the extent to which retrieved information contributes to updating (and not formation of) group-level variability judgments that were, in fact, abstracted on-line. After all, retrieved exemplars do not account for all the variance in group judgments. This possibility is made less likely by the finding that variability judgments took significantly longer than on-line liking judgments (unless one also assumes that some on-line judgments are checked and others are not). Thus, even the most cautious interpretation of our results indicates that retrieval mechanisms play a central role in variability judgments. Accordingly, models that suggest that variability judgments themselves are made on-line do not seem consistent with our data.

Our results suggest the importance of the retrieved similarity information in group variability judgments. Our findings provide partial support for the DPM, although difference information—which was equally available—did not influence judgments. It is not clear why difference information did not play a greater role in our experiment. However, the fact that retrieved similarity information predicts variability in judgments suggests that the role of similarity information (rather than all retrieved information) warrants much closer attention by researchers in this domain. Findings that the type of information or exemplar retrieved is more important than the sheer number of exemplars retrieved suggests the importance of looking at differential information encoding, organization, and retrieval strategies as possible mediators of in-group and out-group variability judgments (see other articles in this issue). For example, more similarity information tended to be retrieved about the out-group than about the in-group in our second experiment. When variability judgments are based upon retrieval of similarity information, such an occurrence provides a sufficient explanation for the OHE.

Our interpretation of variability judgments as memory-based is at odds with conclusions reached by Park and Hastie (1987). These researchers concluded that variability judgments were made on-line because they seemed impervious to repetition manipulations designed to influence accessibility (in Experiment 1) and showed primacy rather than recency effects (in Experiment 2). However, it is likely that the relatively small number of exemplars presented in those studies enabled subjects to accurately "keep track" of individual group members and thus discount and eliminate repetitions of particular items. Thus, variability judgments could have been made in a memory-based fashion without repeated items playing any special role (see Rothbart, Fulero, Jensen, Howard, & Birrell, [1978] for a similar interpretation of findings from this paradigm). Given the emerging evidence that subjects pay differential attention to selected types of information, manipulations of accessibility that are not sensitive to this fact do not provide unequivocal evidence for or against memory-based processing.

It is also possible that our results reflect the special nature of our dependent measure, which asked subjects to make a global variability judgment about the group as a whole. Obviously, making global variability judgments about the group (which include or collapse across information about multiple dimensions) might be different from abstracting variability information about a particular dimension on which information is explicitly presented. Park and Judd (1990) have recently demonstrated that different types of measures may produce different out-group homogeneity effects. In particular, they differentiated between measures of variability within a group and the extent to which groups are seen as exemplifying the group stereotype. We are reassured by the fact that our measure of global similarity appears to tap both aspects of variability that contribute to the OHE.

HOW DO IN-GROUP AND OUT-GROUP VARIABILITY JUDGMENTS DIFFER?

In both experiments, we found that perceivers made variability judgments about the in-group more rapidly than they did about the out-group. None of the various models discussed explicitly predict such a difference. In fact, as noted, the EFM, AEM, and FDM would predict the opposite pattern of results. These models argue that a greater number of exemplars or subtypes are retrieved for in-group judgments than for out-group judgments: thus, in-group judgments should take more time. Given that the amount of similarity and difference information about the different groups was equated (especially in Experiment 2), such a difference is not predicted by the DPM either, although the finding could be incorporated by assuming that similarity information was differentially salient. As salience was not objectively manipulated, this ad hoc explanation focuses attention on processing strategies that confer different psychological salience on in-group and out-group information.

As noted above, these findings could be dealt with by modifying existing models to take into account ease of information retrieval rather than the mere fact of its retrieval. Again, this suggests that models need to pay attention to encoding, storage, or retrieval differences that make particular information about particular groups more likely to be retrieved than other information. For example, one possible explanation of our results is that both in-group and out-group variability judgments are memory-based, but in-group exemplars are more easily and rapidly retrieved than are out-group exemplars.

Of course, faster in-group latencies could also be produced by the operation of other processing mechanisms. First, it could be that in-group judgments are based to a larger degree on thinking about the self as an in-group exemplar, either alone or in addition to other exemplars (Park & Judd, 1990). Such a possibility would be consistent with the finding from Experiment 2 that retrieval of similarity information alone was not as good a predictor of latency or extremity for in-group variability judgments as for out-group judgments. Second, longer response latencies might reflect less certainty about out-group judgments: Out-group variability judgments might be the ones that are checked by retrieval before being reported. In fact, Judd and Park (1988) report one condition in which more out-group than in-group exemplars were retrieved. A third explanation suggests that in-group judgments are formed on-line to a greater extent than are out-group judgments. If subjects are more likely to abstract in-group variability information on-line, then they might not need to retrieve as many

exemplars before being willing to report their judgments, given that they had already established some sort of estimate. Our finding of reduced prediction of in-group judgments by retrieval of similarity information and Judd and Park's finding of increased recall for out-group exemplars is consistent with this explanation. Our experiments were not designed to distinguish among or definitively eliminate any of these possibilities, which await further research.

ON-LINE AND MEMORY-BASED PROCESSES IN THE OHE: CLOSING COMMENTS

Progress on difficult research questions is often made not only by asking the right question, but also by discovering an informative way of assessing the answer to this question. In the research described here, we attempted to use paradigms for assessing the extent to which variability judgments are made in an on-line or memory-based fashion. Like other research that has attacked the issue of mediation of the OHE, our paradigm has strengths and weaknesses. The response latency paradigm has the advantage of providing a straightforward test of whether judgments are made in an on-line or memory-based fashion, although the relative contribution of each of these processes to mixed judgments is difficult to assess. Similarly, evaluating the correspondence between recall and judgment has the difficulties of isolating the relevant information to measure and of providing largely correlational support for hypotheses.

In closing, we would like to suggest a more direct method for identifying the kind of information used in group variability judgments more precisely. This technique extends the priming technique used by Klein and his colleagues (e.g., Klein & Loftus, in press), to examine the extent to which judgments of self and others are dependent on exemplar information. The logic of the methodology is as follows: To the extent that performing one task makes available information that is used in performing a second task, the first task should facilitate the performance of the second task. Extension of this logic to the domain of group variability judgments suggests that if people actually retrieve individual or group-level exemplars when making variability judgments, then having subjects retrieve such exemplars prior to making the variability judgment should decrease the time required to make it, relative to appropriate control conditions. Use of this technique may establish more conclusively the extent to which variability judgments about both real and minimal in-groups and out-groups are based on different kinds of information.

Our research has focused on the extent to which in-group and out-group variability judgments are made on-line or in a memory-based fashion. Although our intent was to help distinguish among different models of the OHE that have been proposed in the literature, the issue of how such variability judgments are made has other important implications for intergroup perception. It has been suggested, for example, that on-line judgments are more difficult to alter than memory-based judgments. Park and Hastie (1987) argue that once an initial estimate has been formed, an anchoring and adjustment process may govern the subsequent integration of new information into the variability estimate. As a result, on-line judgments are underadjusted in response to the variance information in newly encountered exemplars. In contrast, pure retrieval-based models suggest that variability judgments are more responsive to newly encountered exemplars. Thus, establishing the extent to which in-group and out-group variability judgments are on-line and memory-based has implications for our prospect of changing intergroup perceptions, as well as the techniques most likely to do so.

REFERENCES

Fried, L. S., & Holyoak, K. J. (1984). Induction of category distributions: A framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 234–257.

Hastie, R., & Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93,* 258–268.

Hintzman, D. L. (1986). "Schema abstraction" in a multiple-trace memory model. *Psychological Review, 93,* 411–428.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95,* 528–551.

Jones, E.E., Wood, G.C., & Quattrone, G.A. (1981). Perceived variability of personal characteristics of in-groups and out-groups: The role of knowledge and evaluation. *Personality and Social Psychology Bulletin, 7,* 523–528.

Judd, C. M., & Park, B. (1988). Out-group homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology, 54,* 778–788.

Kashima, E. S., & Kashima, Y. (this issue). Perceptions of general variability of social groups. *Social Cognition.*

Klein, S. B., & Loftus, J. (in press). The mental representation of trait and autobiographical knowledge about the self. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 5). Hillsdale, NJ: Erlbaum.

Kraus, S., Ryan, C.S., Judd, C.M., Hastie, R., & Park, B. (this issue). Use of mental frequency distributions to represent variability among members of social categories. *Social Cognition.*

Lichtenstein, M., & Srull, T. K. (1987). Processing objectives as a determinant of the relationship between recall and judgment. *Journal of Experimental Social Psychology, 23,* 93–118.

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

Linville, P. W., 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). New York: Academic.

Mackie, D. M., & Asuncion, A. G. (1990). On-line and memory-based modification of attitudes: Determinants of message recall-attitude change correspondence. *Journal of Personality and Social Psychology, 59,* 5–16.

Messick, D. M., & Mackie, D. M. (1989). Intergroup relations. *Annual Review of Psychology, 40,* 45–81.

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., Judd, C. M., & Ryan, C. S. (1991). Social categorization and the representation of variability information. In M. Hewstone & W. Stroebe (Eds.), *European review of social psychology* (Vol. 2), pp. 211-245. Chichester, England: Wiley.

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.

Quattrone, G. A., & Jones, E. E. (1980). The perception of variability within in-groups and out-groups: Implications for the law of small numbers. *Journal of Personality and Social Psychology, 38,* 141–152.

Rothbart, M., Fulero, S., Jensen, C., Howard, J., & Birrell, B. (1978). From individual to group impressions: Availability heuristics in stereotype formation. *Journal of Experimental Social Psychology, 14,* 237–255.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352.