



Typicality and group variability as dual moderators of category-based inferences

Alan J. Lambert,^{a,*} Alison L. Chasteen,^b B. Keith Payne,^c and Lara Shaffer^d

^a Department of Psychology, 1 Brookings Drive/Box 1125, Washington University, St. Louis, MO 63130-4899, USA

^b University of Toronto, Canada

^c Ohio State University, USA

^d Washington University, USA

Received 15 January 2003; revised 15 February 2003

Abstract

We propose and test two alternative hypotheses bearing on the dual roles of group variability and typicality when people form impressions of single category members. The *latitude of acceptance* hypothesis suggests that a wider range of individual group members are likely to be seen as good-fitting members (i.e., typical) if the group is heterogeneous, thereby increasing the extent to which stereotypical attitudes are used as a basis for responding to these persons. In contrast, the *typicality-functionality* hypothesis suggests that typicality plays different roles depending on group variability. This view suggests that typicality plays the “gatekeeper function” as postulated by Fiske and Neuberg (1990) when the group is homogeneous, but not when it is heterogeneous. Across two studies, stronger support was found for the typicality-functionality hypothesis. Implications for the extant literature on category-based processing are discussed.

© 2004 Elsevier Inc. All rights reserved.

Social psychologists have generated an impressive amount of research on perceived group variability over the past 20 years (see Linville, 1999 for a review). Nevertheless, this literature has mainly focused on the antecedents, rather than the consequences, of group variability. Hence, more is known about the factors that determine whether a group is seen as heterogeneous vs. homogeneous than about the difference that this factor might make in driving behavior and judgment. This state of affairs prompted Linville (1999) to observe that researchers “have made only limited progress toward learning the consequences of variability” (p. 448) and that there is “more speculation than empirical fact” (p. 454) with respect to the effects of this factor on social judgment. The overriding goal of this paper was to shed more light on this surprisingly understudied issue.

Research on the “gatekeeper” function of typicality

Psychologists have long recognized that people’s stereotypic expectations about social categories (e.g., *Blacks*, *elderly people*, *college professors*) can have a powerful effect on people’s impressions of single group members (Allport, 1954; Lippman, 1922). Nevertheless, stereotype-based responses may not be inevitable given that the person has been identified as a member of the category (Kunda & Spencer, 2003). One well-researched moderator of stereotype application is the typicality of the target person, that is, the perceived “goodness of fit” to the category in question. According to Fiske and Neuberg (1990; see also Brewer, 1988; Lord, Lepper, & Mackie, 1984), typicality could be conceived as a kind of “gatekeeper,” moderating the degree to which category-based sentiments guide subsequent reactions to the target. For example, suppose that you judged a particular individual to be a typical member of the group *elderly people*. This research suggests that stereotypic preconceptions about this group would be more likely to have a positive, assimilative effect on your judgments of this

* Corresponding author.

E-mail address: alambert@artsci.wustl.edu (A.J. Lambert).

person if he or she were perceived to a relatively typical, rather than atypical, group member.

Given the importance of typicality in these matters, this begs the question of what factors might determine whether single group members are perceived to be relatively typical or atypical in the first place. In the social psychological literature, nearly all of the focus has been on central tendency: perceptions of what the average or prototypic group member is like (Fiske & Neuberg, 1990). In particular, typicality assessment is presumed to involve a matching process between (a) the observed features of the target with respect to one or more trait dimensions (e.g., *honesty*) and (b) perceived central tendency of the group along these dimensions. Nevertheless, research has shown that people are sensitive to and store information not only about the central tendency of categories, but about group variability as well (Guinote, Judd, & Brauer, 2002; Lambert, 1995; Lambert, Barton, Lickel, & Wells, 1998; Lambert & Wyer, 1990; Linville, Fisher, & Salovey, 1989; Park & Hastie, 1987; Park & Rothbart, 1982; Ryan, Judd, & Park, 1996).

Given that typicality is a known moderator of category-based inferences, what is known about the dual roles of typicality and group variability in these matters? Surprisingly, very little. This could be due, in part, to historical timing. In particular, those stereotyping models that most explicitly address typicality (Brewer, 1988; Fiske & Neuberg, 1990; Lord, Lepper, & Mackie, 1984) appeared at a time when prototype-based models of categorization still held sway in the social literature. Hence, the extent to which group variability and typicality might play interactive roles during the impression-formation process, and their implications for category-based inferences, is virtually unexplored.

One of the few studies to systematically investigate the relation between group variability and typicality was conducted by Lambert (1995). Because the present research was designed, in part, to resolve some important issues left unanswered by that study, we consider it in some detail here. That study used an “artificial group” paradigm similar to that used by Park and Hastie (1987; see also Manis, Paskewitz, & Cotler, 1986). Participants were presented with information about the performance of 50 members of a fictional group on a test of spatial skills. Participants were presented in sequential (i.e., one-by-one) fashion the first name of the group member along with their score on the spatial skills test (e.g., Bob: 6). The actual variability of these scores varied across condition such that they were either extremely low (homogeneous) or high (heterogeneous). Afterwards, participants then made judgments about the average cognitive abilities of the group. Although central tendency was actually identical in all cases, there was substantial variability in these judgments. This measure thus captured individual differences in participants’ beliefs about the group (i.e., the “group stereotype”).

Participants were then presented with a counterbalanced array of six “new” group members whose performance on the test of spatial skills varied in terms of whether they were slightly, moderately, or extremely discrepant from central tendency. Participants made two types of judgments of these individuals (again, counterbalanced for order). In one block, participants judged the typicality of all six new group members by pressing one of two keys, one marked *typical* and the other marked *atypical*. The latency of response (RTs) as well as the actual judgment rendered was recorded. (We shall consider these data in more detail presently.) In a second block, participants judged the same six targets, this time rating each person with respect to the continuous dimensions of (a) overall perceptual motor skills, (b) overall cognitive skills, and (c) general intelligence. Each of these three ratings was made along a scale ranging from 0 to 10.

Although this design is undeniably artificial in several respects, it permits analyses that are roughly analogous to previous research on category-based inferences (cf. Fiske & Neuberg, 1990). In particular, it allows investigations of the correspondence between group stereotypes and judgments of single group members, operationalized here as the correlation between (a) participants’ overall judgments of the group with (b) their judgments of each of the six new group exemplars, based on an average rating of the person’s motor skills, cognitive skills, and intelligence. Overall, results showed modest assimilation effects (cf. Manis, Paskewitz, & Cotler, 1986). The more participants believed the group to be cognitively skilled, the higher they rated the cognitive skills of any given group member, and vice versa. However—and more important for present purposes—typicality moderated these assimilative effects in the manner postulated by Fiske and Neuberg (1990), but *only* when the group was homogeneous. In that case, assimilation occurred when the target’s score was near the group average, but not when it was moderately or extremely discrepant from it. When the group was heterogeneous, however, assimilative effects occurred *regardless* of the discrepancy of the target’s features from the group mean.

Two alternative models of typicality and group variability

The results of Lambert (1995) were provocative, because they suggested that group variability and typicality could interact in important ways to influence when and how people use category-based knowledge when responding to single group members. This possibility is important to consider because some very influential and widely accepted models of category-based inferences (Brewer, 1988; Fiske & Neuberg, 1990) are completely silent on the interactive roles of these factors.

However, one major weakness of the Lambert (1995) investigation is that it was unable to distinguish between two competing hypotheses, each of which offers a different interpretation of precisely *how* group variability and typicality play dual moderating roles. We briefly review those hypotheses below.

Latitude of acceptance hypothesis

In many cases, central tendency of the group—conceptions of what the “average” group member should be like—should be the strongest determinant of perceived typicality. As will be seen later this was, in fact, the case in both of the studies reported in this paper. Nevertheless, group variability could conceivably play a role as well. Although few studies have directly investigated the effects of group variability on typicality, at least one investigation suggests that single group members are more likely to be regarded as relatively typical if the group is heterogeneous (Park & Hastie, 1987). This notion makes some intuitive sense. To the extent that heterogeneous categories have a greater frequency of stored exemplars in the tails of the distribution, this would make it more likely that discrepant group members (i.e., persons whose attributes are in the upper or lower distribution of known group members along a given dimension) would be included within the stored representation of the group. Consistent with this “latitude of acceptance” perspective, Lambert (1995) found some evidence, albeit small in magnitude, that participants regarded the moderately and extremely discrepant exemplars as more typical if the group was heterogeneous than if it was not. This provides one fairly simple interpretation of the correlational results described above. If people are indeed more likely to mentally include discrepant exemplars as part of heterogeneous categories, this would make it more likely that the group stereotype would be seen as relevant to judgments of such exemplars. This, in turn, would increase the extent to which the group stereotype would have an assimilative effect across a wider range of known group members.¹

¹ Our use of the term “latitude of acceptance” here bears only loose correspondence to how it is used in a well-known model of assimilation and contrast by Sherif and Hovland (1961). Sherif and Hovland (1961) postulated that when people are asked to judge where a particular attitudinal statement falls along the *pro-anti* continuum, these judgments are assimilated to one’s own attitude when it is sufficiently similar to it (i.e., falls within the latitude of acceptance). However, we use this term to refer to the effects of perceived group variability on perceived typicality of single group members, and the consequences of such judgments for the application and use of group attitudes. Hence, although our framework, like that of the Sherif and Hovland model, has general implications of assimilative effects in social judgment, the specific considerations involved in such effects are rather different. [For a discussion of some unresolved issues associated with the Sherif and Hovland (1961) model of assimilation, see Lambert and Wedell (1991).]

Typicality-functionality hypothesis

Rather than affecting whether a particular exemplar is *judged* to be typical or atypical, group variability could affect whether typicality is actually *used* as an important cue during the impression-formation process. To preview the line of reasoning presented below, this hypothesis suggests that typicality plays the gatekeeper role assigned to it by the Fiske and Neuberg (1990) formulation, but only when the group is perceived to be homogeneous. When the group is perceived to be heterogeneous, however, typicality is less likely to play this kind of gatekeeper role. In this latter case, category-based knowledge (e.g., the group stereotype) is used as a basis for responding to the target, regardless of whether his features are considered to be typical or not. Note that this hypothesis represents a more significant departure from the Fiske and Neuberg (1990) formulation compared to the preceding hypothesis, insofar as typicality is presumed to *function* in quite different ways, depending on group variability.

Extant evidence for the “typicality-functionality” hypothesis

The preceding line of reasoning can be framed more formally in terms of an empirically testable proposition: Perceivers who view the category in relatively homogeneous terms should be more likely to spontaneously consider the typicality of a single group member to be an important/meaningful cue compared to perceivers who view the category as more heterogeneous. One way to test this hypothesis is through a RT (reaction-time) methodology in which one measures the amount of time required to assess typicality. Our framework suggests that such assessments should be faster if the group is homogeneous than if it is not. Recall that Lambert (1995) found that group variability affected whether individual exemplars were judged as typical or atypical. Importantly, and more relevant to this model, this study also found that participants were faster to render these typicality ratings if the group was homogeneous than if it was not. One plausible explanation for this finding is that participants may have been more likely to have spontaneously assessed typicality (and/or considered it to be more important) in the former compared to the latter case.

Several other studies have provided additional support for this line of reasoning. Moreover, this evidence has arisen in a number of different types of paradigms, all of which were quite different from the Lambert (1995) experiment, suggesting that that the impact of group variability on typicality function is fairly generalizable (i.e., not idiosyncratic to one experimental design). In a study by Verplanken, Jetten, and van Knippenberg (1996; see also Vonk & van Knippenberg, 1995), for example, researchers showed that participants devoted more attention (measured in terms of reading times) to atypical

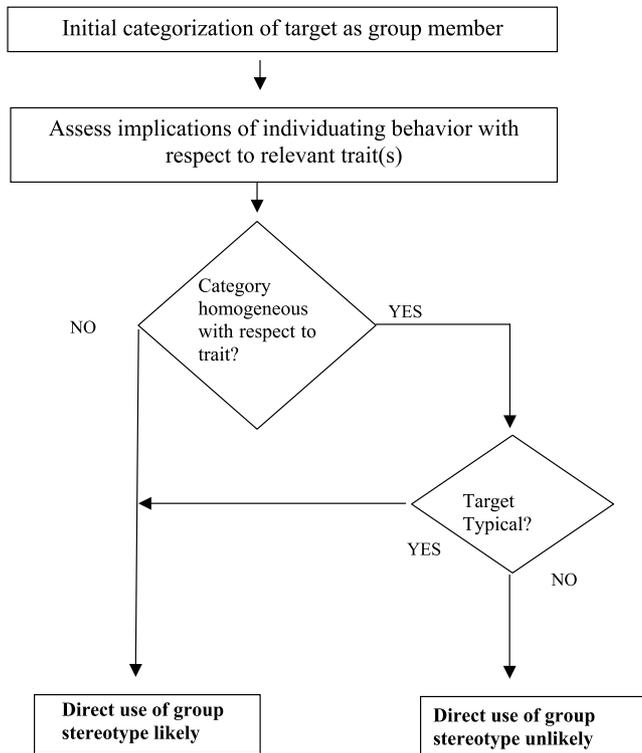


Fig. 1. Summary of predictions arising from the typicality-functionality hypothesis.

compared to typical information, but this was only true when the category was homogeneous. When the group was heterogeneous, typicality did not moderate reading times in any manner whatsoever, suggesting either that participants did not assess the typicality of the information in the first place or, alternatively, simply considered it to be an unimportant feature. Along these same lines, Bardach and Park (1996) showed that people have better memory for stereotype-inconsistent compared to stereotype-consistent information but, again, this was only true when the group was homogeneous. Taken together, this suggests that “deviations from expectations,” including perceptions of target typicality, looms larger when the category is homogeneous.

This view suggests, therefore, an important qualification on existing models of category-based processing that postulate the “gatekeeper” role of typicality (Brewer, 1988; Fiske & Neuberg, 1990). In particular, it suggests that typicality plays this kind of role when the group is perceived as homogeneous, but not when it is heterogeneous. In the latter case, people should be more likely to use/apply their personal evaluation of the group to the target, regardless of typicality. Fig. 1 provides a relatively simple, descriptive summary of this hypothesis.

Summary

In the preceding section, we considered two hypotheses, each of which make different assumptions regard-

ing the interactive roles of group variability and typicality. The latitude of acceptance hypothesis stresses the impact of group variability on *judged* typicality, whereas the typicality-functionality hypothesis focuses more on the *function* of typicality. We will review the more specific predictions that flow out of each hypothesis presently,

Despite the differences between these hypotheses, the observant reader may have already noted that they both lead to a rather counterintuitive implication. In particular, both suggest that higher levels of group variability should lead to *more* widespread use/application of the group stereotype across a wide range of different kinds of group members. At first blush, this seems at odds with a more intuitive view of group variability, which suggests that higher levels of perceived variability might be associated with “weaker” stereotyping effects compared to cases in which the group is perceived in homogeneous terms. In other words, when considering the potential impact of category-based knowledge on social judgment, one might expect this impact to be less pervasive if the group is perceived as heterogeneous than if it is not.

Relatedly, one might suppose that people would be less confident in their overall expectations about the group if its members were perceived to be heterogeneous in their attributes than if they were not. In fact, some of our own past work (Lambert et al., 1998) as well as previous research (Dijksterhuis & van Knippenberg, 1999; Ryan et al., 1996) provides evidence for this notion. Lambert et al. (1998) found that experimental manipulation of group variability led participants to be less confident in making assertions about the category if its members were heterogeneous in their attributes. In a second study, Lambert et al. (1998) also found that participants were relatively slow to make assertions about the category as a whole if its members were heterogeneous in their attributes (see also Dijksterhuis & van Knippenberg, 1999). A conceptually similar finding was reported by Ryan et al. (1996), who found a negative correlation between group variability and the confidence with which participants made trait-based judgments of single group members.

One of the goals of the present research was to show that these two views of group variability are not at all contradictory. Indeed, in Experiment 2, we report supplemental data showing that these two consequences of group variability—differential use of typicality and self-reported confidence in beliefs about the category as a whole—can be demonstrated within the same experimental design.

Overview of the present research

Although the two experiments reported in this paper differ in some important ways, both share some common

elements. Both studies first had participants complete a 25-page mass-testing packet containing several different types of questionnaires, most of which were unrelated to present concerns. In the middle of the packet, we included a general measure of participants' overall evaluations of the elderly and, in addition, measures of perceived group variability and central tendency with respect to traits associated with this group. (After these measures, participants spent the next 5–10 min on unrelated tasks.) Next, in an ostensibly unrelated study, participants were then asked to form an impression of a single elderly individual on the basis of a one-page written description of his individuating characteristics. After presentation of this information, participants were asked to express their general impressions of the target and, following this, rated the extent to which the target's features were consistent with their expectations about elderly people as a whole.

Predictions

In the paradigm described above, there are four constructs that are relevant to our model: (a) participants' stereotypic evaluation of the group as a whole (hereinafter, group attitudes); (b) perceived variability of the group with respect to one or more stereotypic dimensions; (c) perceived typicality of the single group member; and (d) overall impression of this individual. In our analyses, we treat participants' overall impression of the target as the criterion variable, and the other three factors as predictors.

Consider first the latitude of acceptance hypothesis. Critical to this hypothesis is that group variability increases the extent to which targets with relatively rare/infrequent features (e.g., a physically active elderly man) are perceived as relatively typical. Beyond this effect on judged typicality, however, there is nothing in this hypothesis to suggest that group variability would affect the function of typicality, in terms of how it might moderate the relation between group attitudes and judgments of the target. Hence, this hypothesis leads to a prediction of a Group Attitudes \times Typicality interaction, independent of Group Variability. Another way of framing this prediction is that typicality should moderate the relation between stereotypic attitudes and judgments of the target (i.e., the Group Attitude \times Typicality interaction), regardless of group variability.

The typicality-functionality hypothesis is different in two important ways. First, the issue of whether group variability affects judged typicality is largely irrelevant in this case. In other words, this hypothesis certainly allows for the fact that group variability could affect judged typicality, but the presence (or absence) of this kind of effect does not bear on its validity. More relevant to this hypothesis is the idea that typicality should function in different ways, depending on group variability. That is,

typicality should moderate the relation between group attitudes and judgments of the target—strong assimilation when typical, weak when atypical—but only when the group is perceived as homogeneous. In contrast, typicality should not act as a moderator when the group is perceived as heterogeneous, producing assimilation effects, regardless of how the target is judged. Statistically, this prediction corresponds to a three-way Group Attitudes \times Group Variability \times Typicality interaction.

Study 1

Method

Participants

Participants included 193 undergraduates who either received course credit or payment of \$7 for their participation. Of these participants, 78 were male and 94 were female; gender was not recorded for 21 participants. Separate analyses on the 172 participants for whom gender was recorded revealed a pattern of results that was very similar to the entire sample and none of the analyses using this subset revealed any significant effects of gender. Hence, the analyses to be reported below are collapsed over this variable.

Measurement of group-related variables

Group attitudes. Group attitudes were measured on the basis of participants' responses to a balanced 16-item measure of individual differences in elderly attitudes similar to that used by Lambert and Chasteen (1997). We deliberately selected our items to capture a diverse assortment of different factors relevant to people's overall feelings about the elderly (e.g., *It is irritating to have to listen to older people talk about the 'good old days'; It's wrong to assume that senior citizens are wise and should always be respected; Social security benefits for the elderly need to be reduced; Young adults should care for their parents when their parents get old; I think that older people have a lot to contribute to society*). After reverse scoring anti-elderly items, a composite measure of group attitudes ($\alpha = .78$) was formed on the basis of these items, such that higher numbers indicated more favorable attitudes toward the group. Previous work by Lambert and Chasteen (1997; Lambert, 1998) has generated support for the construct and discriminative validity of this measure, showing (for example) that it is uncorrelated with other attitudinal measures (e.g., attitudes towards Blacks) once shared overlap in relevant ideology, such as political conservatism, is accounted for.

Group variability and central tendency. Several pages later in the packet, perceived group variability was as-

sessed using a “frequency distribution” procedure (Judd & Park, 1988; Lambert & Wyer, 1990; Wyer, 1973). In this procedure, participants were provided with a single trait construct along with an 11-point scale ranging from *not at all* (0) to *extremely* (10). Participants were then asked to estimate how many group members, out of 100, they think would fall into each of the 11 scale categories with respect to that trait dimension. Each participant’s responses form a frequency distribution, from which a variability (standard deviation) parameter can be computed for each participant. In this study, participants were given the trait “physically healthy” and were told to complete this task with respect to their beliefs about “elderly persons.” (In order to make sure that participants knew what we meant by “healthy,” participants were explicitly informed at the beginning of the task that “by healthy, we mean in good physical condition without major illness or serious medical complications.”) Note that this methodology also permits assessment of perceived central tendency with respect to the focal trait. Central tendency could be considered analogous to a type of “stereotypic index” of the group (i.e., capturing one’s beliefs about the average group member) and, as we discuss later, could also be used in a meaningful way to predict people’s reactions toward a single group member.

Target impression task

Target description. Participants were told that they would be presented with some information about an individual and that they should form the most accurate impression that they could about this person. (At no point during this experiment were participants told that the study was about elderly people, nor were they informed that we had some special interest in their attitudes toward this group.) At the top of the next page of the experimental packet, participants were provided with some background information about the target, including his name, home address, place of birth, gender, race, and age. In all cases, the target was identified as a white male who was 69 years of age. Importantly, the target’s age was presented incidentally along with the other information about him and participants were given no special indications that they should either consider or ignore this information while forming their impressions. Below this background information appeared a short vignette about the target that was designed to convey that this individual was relatively healthy and physically active (e.g., “As Jim headed for his morning shower, he thought about the positive results from the physical he took yesterday. He was very fortunate to be in such good health at his age. His favorite activities were playing tennis and jogging, but he enjoyed many other events such as day trips to the nature center, where he could hike along the trails.”)

Target judgments. Following the target description, participants were asked to indicate their overall reaction to the target along a scale ranging from -5 (*very unfavorable*) to $+5$ (*very favorable*) and were asked to indicate how much they would look forward to actually meeting him along a scale from 0 (*not at all*) to 10 (*very much*). On the next page of the packet, participants were asked to rate the target with respect to 8 trait dimensions along a scale ranging from 0 (*not at all*) to 10 (*extremely*), including *likeable*, *active*, *healthy*, *friendly*, *wise*, *forgetful*, *intelligent*, and *frail*. Given participants’ explicit goal of forming an overall impression of the target, we anticipated and found that their responses to the various target ratings would tend to be correlated with one another, essentially yielding a general likeableness cluster. This was confirmed by principal component analyses, the largest of which corresponded to a main evaluative component on which almost all of the items loaded highly (average loading = .65).² One exception to this trend was the rating of “frail”, which yielded, compared to the other items, the lowest factor loading (.34). Reliability analyses confirmed this fact, as this item yielded the lowest corrected item-total correlation amongst all the items used. Thus, in our primary analyses we used as our main criterion variable an average of the remaining nine items ($\alpha = .83$) as an index of participants’ overall impression of the target person. (Analyses using all 10 items yielded a very similar, although slightly weaker, pattern of results.)

Typicality ratings. On the next page of the experimental packet, participants were asked to assess the perceived typicality of the target by responding to the following statement: “Sometimes we meet people whose behaviors fit our expectations about the group to which they belong (are considered to be typical) but other times their behaviors do not fit those expectations (are atypical). Using the scale below, please indicate the extent to which you consider John’s behaviors to be typical or atypical to your general expectations about people the same age as John.” Participants responded to this question along a scale ranging from 0 to 10, with higher numbers indicating greater perceived typicality. Following their completion of this measure, participants were completely debriefed, thanked for their participation, and dismissed.

Results and discussion

Correlational analyses

Table 1 shows the zero-order correlations between the relevant group parameters, namely, group attitudes,

² In this and other principal component analyses reported in this paper, the only a priori constraint was that components have eigenvalues of 1.0 or greater.

Table 1
Correlation among category and target-person variables (Study 1; $N = 193$)

	1	2	3	4	5
1. Group attitudes	1.0				
2. Group variability	.16*	1.0			
3. Central tendency	.12	.03	1.0		
4. Overall target judgment	.33**	.04	.04	1.0	
5. Typicality	.29**	.10	.38**	.00	1.0

* $p < .05$.

** $p < .01$.

perceived variability/central tendency with respect to the focal dimension of physical health, and the two classes of judgments of the target (composite impression and typicality). In this study, group attitudes were uncorrelated with central tendency, showing that overall liking for the group was not correlated with the extent to which the group was seen, on average, as relatively healthy/unhealthy. Second, analyses showed a small but reliable relation between group attitudes and perceived group variability, indicating a tendency for the group to be seen as more variable among participants who held favorable attitudes toward the group. Finally, group variability was uncorrelated with central tendency along the focal dimension of health.

Turning now to judgments of the target, group attitudes were significantly correlated with target judgments. Thus, independent of other considerations, there was a reliable relation between participants' feelings about the group and their impressions of a single member of it. Also, given that this person was described in fairly positive and healthy terms, we expected and found that the perceived typicality of the target would be positively correlated both with group attitudes as well as the perceived central tendency of the category along the health dimension. This shows that participants who tended to like elderly people, and expected them to be relatively healthy, viewed our healthy target as more typical. Importantly, typicality ratings were not correlated with group variability. Thus, whatever consequences of group variability to be reported below, they cannot be attributed to differential ratings of typicality as a function of group variability. This provided an initial indication that the latitude of acceptance hypothesis would turn out to be less viable than the typicality-functionality hypothesis.

Aside from its relation to typicality, perceived central tendency was not related to our measure of group attitudes. Nor was it correlated with judgments of the target. These null findings could be due, in part, to the fact that central tendency was based only on one numerical index, derived from the distribution task performed on the single trait *physically healthy*. This could limit the usefulness of this measure as a tool to predict reactions toward the target. Given the lack of relation between

this measure to group attitudes and judgments of the target, we shall defer further consideration of this measure until Study 2.

Regression analyses

Test of the predicted three-way Group Attitudes \times Group Variability \times Typicality interaction was conducted in regression analyses, after standardizing all variables (Aiken & West, 1991). These analyses yielded support for the predicted three-way Group Attitudes \times Group Variability \times Typicality interaction $F(1, 185) = 3.03$, $p = .084$. Although this effect was marginally significant, subsequent theory-driven analyses of the data strongly confirmed the assumptions of our framework. When the group was seen as relatively homogeneous (as determined on the basis of a median split), regression analyses yielded strong support for the expected Group Attitudes \times Typicality interaction ($p < .001$). When the group was heterogeneous, however, entry of group attitudes and typicality in the first step yielded only a significant effect of group attitudes ($p < .001$), no effect of typicality ($p > .25$) and no evidence of the two-way interaction in the second step (F change < 1.0).

The formal tests of these interactions, reported above, were based on differences between standardized β weights (Aiken & West, 1991). Nevertheless, research and theory in the attitude literature has historically emphasized the correspondence between general attitudes and specific judgments in terms of correlations (Ajzen, 1988; Fiske & Neuberg, 1990; Lord et al., 1984). To establish links with this literature, and to descriptively clarify the nature of our results, Table 2 displays the pattern of correlations as a function of participants who (a) viewed the group as relatively heterogeneous or homogeneous, and (b) who judged the target as relatively typical or atypical, based on median splits of these

Table 2
Correlation between group attitudes and overall judgments of target person as a function of perceived group variability and perceived typicality: Study 1 ($N = 193$)

Perceived group variability			
Heterogeneous		Homogeneous	
Perceived typicality		Perceived typicality	
High ($n = 48$)	Low ($n = 49$)	High ($n = 48$)	Low ($n = 48$)
<i>Mean typicality ratings</i>			
7.0 ($SD = 1.02$)	3.84 ($SD = 1.13$)	6.97 ($SD = .98$)	3.25 ($SD = 1.09$)
<i>Correlations of group attitudes with composite target judgment</i>			
.43**	.39*	.54***	-.01

* $p < .05$.

** $p < .01$.

*** $p < .001$.

variables. As seen in this table, when participants viewed the group as homogeneous, group attitudes were strongly correlated with target judgments among participants who judged the target as typical (above the median) but not atypical (below the median). This confirms the implication of the significant simple effects Group Attitude \times Typicality interaction reported earlier. In contrast, correlations were significant regardless of typicality among participants viewing the group as heterogeneous. These data also confirm the implications of Table 1, showing that the influence of group variability did not involve differences in the degree as to *whether* the target was perceived as typical or atypical.³

Study 2

Although Study 1 yielded support for the typicality-functionality hypothesis using several different analytic techniques, the predicted three-way interaction did not achieve conventional levels of significance and hence it seemed prudent to replicate this effect in an additional study. More important, we wished to extend and replicate our findings in a second study that differed from the first in five important respects.

First, we presented participants with a completely different description of the target person. Recall that in Study 1, we presented participants with a relatively active, physically robust individual whose features were generally inconsistent with the traditionally negative stereotype about the elderly. In Study 2, we presented participants with a description of a person who was more likely to be perceived, on average, as more typical of the cultural stereotype of elderly people: a person with moderate physical (e.g., arthritis) and mental (e.g., early senility) deficits. This change is important for our purposes, because it allowed us to test the viability of our model across two target persons who, even though they were both elderly, were described in dramatically different terms. (In the preliminary results section, we report supplemental analyses confirming this assumption.)

Second, we expanded our measure of group variability to include two, and not just one, distinct stereotypic dimensions, *physical health* and *mental health* (mental sharpness). This allowed us more leverage in constructing more reliable measures of variability and

central tendency. A third and fourth change was that (a) we improved our group attitude measure, expanding the number of items in it from 16 to 21, and (b) we expanded the number of items used to assess participants' impressions of the target from nine to twenty three, including a "thermometer rating" designed to tap participants' affective reaction to the target. Finally, a fifth change pertained to our measure of target typicality. Instead of using a global measure of typicality (as was used in Study 1), we framed our questions to more specifically reflect the behavioral implications of the target, asking participants to rate the target with respect to the stereotypic dimensions of physical health and mental sharpness.

Method

Participants

Participants included 110 undergraduates (28 male and 82 female) who either received course credit or payment of \$7 for their participation. Analyses involving gender did not produce any significant effects (either as a main effect or in combination with the other relevant variables) and hence analyses are collapsed over this variable.

The methodology of Study 2 was identical to the first study in all respects, except as follows:

Mass-testing phase

As noted above, we expanded the number of items in the group attitudes task to 21, yielding excellent internal reliability ($\alpha = .88$). Also, participants completed the distribution task with respect to the trait construct "healthy/physically active" and "mentally sharp/quick witted." The standard deviations derived from these two distributions were highly correlated ($r = .66$), and the same was true of the central tendency parameters from these tasks ($r = .49$, $p < .01$). We thus formed a composite index of perceived variability and central tendency derived from the two distributions.

Target judgment phase

We presented participants with a new target description, one whose features were more aligned with the traditional negative stereotype of the elderly. In the first paragraph, the target was described as becoming "more and more forgetful" and that he "has had similar lapses while driving, during which he has been temporarily disoriented." The target was also described as being resistant to having his car keys taken away from him by his family. The second paragraph provided a detailed description of his failing health (e.g., "the debilitating effects of his arthritis was complicated by a slipped disk in his back, thus limiting his mobility") and described an outcome of a recent visit with his doctor, who "expressed some guarded pessimism about John's general

³ Two points regarding these analyses should be noted. First, formal tests of interactions in regression were always conducted with continuous variables. Second, correlation coefficients can be more sensitive to differences in variance across groups compared to unstandardized beta weights. However, supplemental analyses revealed that the variance in group attitudes and target judgments were very similar across conditions and, more important, similar implications obtained when the relation between group attitudes and target judgments are represented in terms of unstandardized beta weights.

health and suggested that he consider monitoring his health more closely over the coming year.”

Following this description, participants expressed their judgments of the target. In addition to the overall liking and “desire to meet” ratings used in Study 1, participants completed a “thermometer” rating of their affective reaction to the target along a scale from 0 to 100, with higher numbers indicating greater positive feelings toward the target. Participants also rated the target with respect to the following 22 traits: *likeable, active, healthy, friendly, wise, forgetful, intelligent, mentally sharp, energetic, kind, sociable, compassionate toward others, gentle, impatient, physically healthy, courageous, optimistic, senile, patient, physically weak, hostile, and physically unhealthy*. Note that many of these items pertained to aspects of the target that were relatively ambiguous, i.e., permitted some degree of differences in interpretation. This would include the evaluative items (e.g., thermometer ratings) as well as trait ratings for which the target behaviors had no direct implications (e.g., kindness). However, a few items pertained to stereotypic features of the target that were relatively unambiguous (e.g., his physical and mental health). We expected that group attitudes would play the more important role in shaping impressions along ambiguous, rather than unambiguous, features of the target.

In fact, principal components analyses revealed two interpretable components, the first appearing to capture a general evaluative component (eigenvalue = 8.2) and a second that pertained more specifically to stereotypically negative (and, in our target scenario, unambiguous) aspects of the target (eigenvalue = 2.42). Most of the items loaded highly on the first component, except for four items (ratings of *physically weak, physically unhealthy, forgetful, and impatient*) that loaded much less strongly on the first component (loadings between $-.01$ and $+.01$) than the second (loadings between $+.38$ and $+.76$). Given our interest, as in Study 1, in forming an overall index of participants’ evaluative reaction to the target, we formed an index based on all of the remaining 18 target ratings; this yielded excellent internal reliability ($\alpha = .91$).

As for the measure of typicality, we framed our questions to more specifically reflect the behavioral implications of the target. Participants were told “compared to your overall expectations about elderly men, please rate how typical you see John with respect to the following traits.” This was followed by the phrases “mentally sharp/quick witted” and “physically active/healthy”, both of which were accompanied by scales ranging from 0 (*atypical*) to 10 (*typical*). Responses to these questions were highly correlated ($r = .50, p < .01$), suggesting that participants who saw the target as typical (or not) along one dimension tended to feel very similarly along the other. We thus formed a composite

index of trait-based typicality perceptions based on the average of these two items.

Results and discussion

Preliminary analyses

Initial correlational analyses, shown in Table 3, revealed some key similarities with Study 1. Consistent with Study 1, group variability was not correlated with either trait-based central tendency or perceived typicality. Also, group attitudes were, as expected, again correlated with judgments of the target, collapsed over other variables. Along with these similarities, however, there were a few differences. Unlike the first study, group attitudes were correlated with central tendency, indicating that participants who liked elderly people as a whole tended to view them to be physically and mentally healthy. Also, typicality was not, as it was in Study 1, correlated with either group attitudes or central tendency. The reason for these differences is not entirely clear although they could have been due, in part, to changes in our measures and/or the description of the target person. However, none of these differences qualify our analytic technique or, more importantly, the implications of our findings as they bear on the viability of the hypotheses under consideration.

Comparison of target implications in Study 1 vs. Study 2

We assumed that the target description used in Study 2 conveyed a picture of an elderly man who would, on the average, be perceived as (a) less physically healthy, and (b) more typical compared to the target in Study 1. Supplemental analyses supported this assumption. First, focusing on how participants rated the target with respect to the single trait *healthy*, participants obviously regarded the target in Study 2 ($M = 3.33; SD = 1.59$) as far less healthy compared to the target in Study 1 ($M = 8.86, SD = .96$), $p < .001$. In addition, the target in Study 2 was perceived as more typical ($M = 6.38, SD = 1.48$) compared to the target in Study 1 ($M = 5.16, SD = 2.04$), $p < .001$. Despite the large differences in means here, some caution should be exercised in drawing conclusions from these analyses given that participants were not randomly assigned to judge the two

Table 3
Correlation among category and target-person variables (Study 2; $N = 110$)

	1	2	3	4	5
1. Group attitudes	1.0				
2. Group variability	.07	1.0			
3. Central tendency	.41**	.04	1.0		
4. Overall target judgment	.46**	.18	.25**	1.0	
5. Typicality	.06	-.02	-.06	.22*	1.0

* $p < .05$.

** $p < .01$.

targets, and to the fact that the typicality composite in Study 2 was slightly different than the typicality rating used in Study 1 (although they did use the same underlying 0–10 metric). Nevertheless, the more important point is that we obtained support for our predictions across both studies, as we show below. This suggests that the predicted interaction involving group attitudes, group variability, and target typicality was not an artifact of the particular description presented to participants.

Regression analyses on overall impressions of the target

As in Study 1, all variables were standardized prior to conducting regression analyses.

Group attitudes. Replicating the overall pattern found in Study 1, strong support was found for the predicted three-way Group Attitudes \times Group Variability \times Typicality interaction, $F(1, 105) = 5.82$, $p = .018$. Moreover, subsequent analyses revealed that when the group was homogeneous, the predicted Group Attitudes \times Typicality interaction emerged ($p < .01$ for entry of the interaction in the second step). In contrast, analyses for the heterogeneous group revealed a significant effect only of group attitudes ($p = .001$). In order to more clearly show its correspondence with Study 1, the top panel of Table 4 shows the pattern of correlations between group attitudes and judgments of the target as a function of group variability and perceived typicality, derived from median splits on the latter two variables. These results strongly replicate the pattern found in the first study. Note again that these effects were not due to differences in ratings of typicality per se,

thus supporting the typicality-functionality hypothesis. In particular, high levels of heterogeneity led to increased correspondence between group attitudes and judgments of the target, independent of *how* typical or atypical the target was rated.

Central tendency. Would our results replicate using a measure other than the group attitude index? In order to show the robustness of our findings, we conducted a parallel set of regression analyses, substituting the central tendency measure for the group attitude measure. (Recall that this measure was based on an average of the perceived central tendency of the elderly with respect to the traits *mentally healthy* and *physically healthy*.) Regression analyses revealed a marginally significant three-way Central Tendency \times Group Variability \times Typicality interaction ($p = .09$) corresponding to a pattern of correlations very similar to those shown with the group attitude measure, as shown in the bottom panel of Table 4. The relative weakness of this pattern is almost certainly due to the fact that our composite group attitude measure was better suited to predict general evaluative reactions to the target. Nevertheless, as with the group attitude measure, the correlation between the central tendency parameter and overall impressions of the target was moderated by typicality, but only when the group was seen as homogeneous.

Supplemental analyses

Group variability can have consequences other than those reported here. As noted earlier, at least two previous studies (Lambert et al., 1998; Ryan et al., 1996) found that people are less confident in making assertions about social categories if its members are perceived in heterogeneous terms. Such effects are quite different from the findings reported in this paper. Both Lambert et al. (1998) and Ryan et al. (1996) were concerned with the relation between group variability and confidence in making judgments about categories and the persons who belong to them. In contrast, the present findings are focused on the moderating effects of group variability insofar as it affects spontaneous use of typicality during the impression-formation process.

It would be useful, therefore, to show that these two consequences of group variability could co-exist, simultaneously, in the same design with the same group of participants. In order to investigate this possibility, we had a subset of our participants from Study 2 ($N = 49$) complete a battery of confidence measures following their impressions of the target. These included (a) six questions pertaining to participants' overall feelings about the elderly (e.g., "In general, to what extent are you confident in your overall impressions of this group and what its members are like?") and (b) eight questions

Table 4

Perceived typicality (top panel) and correlation between group attitudes (middle panel) and central tendency index (bottom panel) and overall judgments of target person as a function of perceived group variability and perceived typicality: Study 2 ($N = 110$)

Perceived group variability			
Heterogeneous		Homogeneous	
Perceived typicality		Perceived typicality	
High ($n = 28$)	Low ($n = 26$)	High ($n = 29$)	Low ($n = 26$)
<i>Mean typicality ratings</i>			
6.84 ($SD = .93$)	4.58 ($SD = .61$)	7.21 ($SD = .83$)	4.62 ($SD = .86$)
<i>Correlation of group attitudes with composite target judgment</i>			
.57**	.70**	.43*	-.01
<i>Correlation of central tendency index with composite target judgment</i>			
.33 [†]	.46*	.36 [†]	-.14

[†] $p < .10$.

* $p < .05$.

** $p < .01$.

pertaining to beliefs about the physical or mental health/alertness of the group (e.g., “To what extent do you think that it is easy, or difficult to form a summary judgment of the group with respect to its overall level of physical health?”) Because answers to these items tended to be highly correlated with each other ($p < .05$), we formed a “confidence index” based on an average of all 14 items ($\alpha = .76$).

Replicating our own previous work (Lambert et al., 1998), participants expressed relatively lower confidence on this index if they perceived the category as relatively heterogeneous ($r = -.28$, $p < .05$) for the correlation with the group variability measure. In fact, group variability was the *only* predictor of confidence, as this index was uncorrelated with group attitudes ($r = -.03$), central tendency ($-.16$), and judgments of the target person ($r = .07$), all *ns*. These findings are important, insofar as they show that group variability can simultaneously exert two rather different kinds of consequences. First, increased perceptions of heterogeneity can decrease the self-reported confidence with which people form inferences about the category as well as the persons who belong to them (Lambert et al., 1998; Ryan et al., 1996). At the same time, however, group variability can have a more complex consequence, having ramifications for the correspondence between group attitudes and judgments of individual members of the group. As we have shown, the influence of group variability in this regard is somewhat indirect, through its influence on the subjective importance of typicality (deviation from expectations) of single group members.

General discussion

The overall purpose of this paper was to shed more light on a surprisingly under-studied issue, namely, the consequences of group variability for social judgment and behavior. As we have shown, group variability moderates this relation in conjunction with typicality. When the group was perceived as homogeneous, the well-known typicality (“gatekeeper”) effects of the sort discussed by Lord et al. (1984; see also Fiske & Neuberg, 1990) emerged: stereotypic attitudes played a role in guiding reactions toward typical, but not atypical, group members. When the group was perceived as heterogeneous, however, the correlation between group attitudes and judgments of the target were the same, regardless of typicality.

We obtained converging support for our conclusions, regardless of whether the target’s attributes were typical (Study 2) or atypical (Study 1) of the general stereotype of the elderly. Moreover, we showed generalizability of our effects across different operationalizations of the “predictor variable” (group attitudes vs. central tendency), and across two different operationalizations of

typicality. In addition, Study 2 successfully addressed one potential criticism of Study 1, which had focused on only one stereotypic trait in the context of measuring individual differences in group variability and central tendency. This convergence is important, because it shows that our results cannot easily be explained in terms of some idiosyncratic or peculiar aspect of our methodology or stimulus materials.

It is true that the two studies reported here both relied on natural (non-manipulated) variation in both group variability as well as typicality. Ordinarily, this might raise some “third variable” concerns, that is, that effects associated with group variability (or typicality) might have actually been due to some other variable with which that factor is correlated. However, our theoretical model builds upon key findings from Lambert (1995), who experimentally manipulated group variability as well as typicality. Hence, although correlational designs cannot, of course, offer definitive evidence on causality, the convergence of two lines of research—one experimental, one correlational—on the same theoretical conclusions makes these issues somewhat less of a concern than they otherwise might be.

Finally, it is worth noting that previous research on a variety of groups other than the elderly generated support for one of our key assumptions, regarding the importance of typicality/deviation from expectations for heterogeneous vs. homogeneous groups (Bardach & Park, 1996; Lambert, 1995; Verplanken et al., 1996; Vonk & van Knippenberg, 1995). This suggests that the implications of the present research are fairly generalizable, and not at all peculiar to something about the elderly stereotype. In the section to follow, we review some additional studies showing indirect evidence for the effects of group variability on the use/application of group stereotypes, all focusing on groups other than the elderly.

On the counterintuitive consequences of group heterogeneity

One provocative implication of the present research is that it shows more generalized use and application of stereotypic attitudes when the group was perceived as heterogeneous. We stress this aspect of our results because it suggests a highly counterintuitive and rather ironic consequence of group variability. An increased perception of diversity within the group is typically seen as one of the primary goals of combating prejudice. Indeed, this is one reason why the contact hypothesis (Miller & Brewer, 1984) seems so intuitively appealing: increased contact should increase the perceived diversity of the group members which, according to the typical line of reasoning in this area, would “dilute” the impact of any previously held stereotypes (see also Judd, Ryan, & Park, 1991). Our data show that high levels of heterogeneity can foster the exact opposite effect, in the

sense that perceivers more readily use their group attitudes as a basis for responding to relatively “poor fitting” members of the group. This finding in no way challenges the overall validity of the contact hypothesis. There may be many reasons why the contact hypothesis might indeed reduce prejudice, for reasons not having anything to do with group variability. Nonetheless, our data show that increasing the perceived diversity of group members should not be seen as a “quick fix” to prejudice and that increased heterogeneity per se may not always decrease stereotyping effects (in the form of application and use of group attitudes) and indeed, may have effects different from those that are usually expected.

Two process-level explanations

As we have shown, the counterintuitive findings demonstrated in this study arise from the different role/impact of exemplar typicality during the impression-formation process, depending on the perceived variability of the group. Although our data provide good support for the typicality-functionality hypothesis (see Fig. 1), these findings beg the question as to *why* the typicality of the target person might be more important if the group is homogeneous. In the space below, we consider two potential candidates, beginning with our preferred interpretation.

Subjective importance of typicality

Both Fiske and Neuberg (1990) and Lord et al. (1984) assume that people spontaneously assess typicality, making it available for its subsequent role as a gatekeeper of the application/use of category-based knowledge structures (e.g., group attitudes prejudice). The present research qualifies this assumption, insofar as this might be true only when the group is viewed as homogeneous. Why might this be? Note that typicality is an “emergent” property of the target, whose existence depends on a meaningful comparison of (a) features of the target, and (b) features associated with the group. Such comparisons are presumably easier if the category has clear/unambiguous implications, which is more likely when the group is viewed as homogeneous. For example, if people believe that most birds fly, then non-flying birds (e.g., penguins) are readily recognized as atypical. By the same token, if all elderly people are believed to be highly sedentary, a spry 69-year-old springing about on the tennis court is more easily judged to be atypical. Another way of framing this idea is that for homogeneous categories, features either occur very frequently, or very infrequently. Accordingly, features are either clearly typical, or clearly atypical [see Smith, Shoben, and Rips (1974) for a related discussion.]

On the other hand, the typicality of any given group member is not so clear-cut when the group is heteroge-

neous. As the importance/meaningfulness of typicality fades, this could make it less likely that people spontaneously *assess and use* typicality during the natural flow of events during the impression-formation process. This could explain why typicality did not moderate the attitude-behavior relation among participants viewing the group as heterogeneous (see Tables 2 and 4). Although it is true that the present research did not directly assess the importance of typicality and its contingency on group variability, there is a plethora of previous work supporting this assumption (Bardach & Park, 1996; Lambert, 1995; Verplanken et al., 1996; Vonk & van Knippenberg, 1995).

A representation-based explanation

In a well-known paper, Lord et al. (1984) extended the correspondence principle of Fishbein and Ajzen (1975) in order to explain why group attitudes might be more predictive of reactions toward typical, rather than atypical, group members. Lord et al. (1984) argue that one should expect strong correspondence between group attitudes and judgments only to the extent that there is a “match” between (a) the cognitive representation of the group that is considered at the time that attitudes are assessed, and (b) the salient features of the person. This suggests, in turn, that people’s general attitudes toward the category might better predict reactions toward typical exemplars. This could, of course, be true of social as well as non-social categories. For example, “if an individual loved birds...then that individual would probably behave positively toward all birds, but especially so toward a robin or a sparrow” (Lord, Desforges, Ramsey, Trezza, & Lepper, 1991, p. 552).

It is important to note that previous work by Lord and colleagues did not focus on group variability. Nevertheless, their formulation is useful insofar as it provides an alternative, “representation-based” account of our findings. When answering our group attitude measure, for example, suppose that one of our participants was thinking (implicitly or explicitly) of an arthritis-ridden elderly man, incapable of walking without the aid of a walker. If so, group attitudes might do a good job of predicting reactions toward a frail, but not a physically robust, elderly person. This is because the former, but not the latter, target was likely to be part of that person’s cognitive representation of the group. A moment’s reflection suggests that people who view the elderly as more variable might be considering a relatively wide range of exemplars during the initial attitudinal assessment stage. To this extent, it stands to reason that group attitudes would be more strongly predictive of a diversity of targets in such cases. In short, the relative breadth of exemplars that are included in the group representation during the initial attitude assessment stage could be directly proportional to the breadth

of “new” group members subsequently predicted by the group attitude measure.⁴

On the one hand, we believe that this explanation is reasonable and it could well account, at least in part, for the results obtained in this paper. However, there is one serious problem with it. This argument suggests that group variability moderated the attitude–behavior relation by virtue of affecting whether any given target was actually perceived as relatively typical or atypical. In other words, it assumes that high levels of group variability make it more likely that any given group member is considered to be a relatively good-fitting member of the group, thereby increasing the magnitude of the attitude–behavior relation. Our data were not entirely consistent with this argument. For one thing, group variability was not correlated with typicality ratings in either study. Even more important, the results of our regression analyses (see also Tables 2 and 4) clearly showed that group variability moderated the attitude–behavior relation, holding constant whether the target was perceived as typical or atypical. Hence, although we do not wish to completely dismiss the merits of a representation-based account, our findings would appear to be more consistent with the “subjective importance” explanation, described earlier.

Relation of the present research to Lambert and Wyer (1990)

Although earlier findings reported by Lambert and Wyer (1990) offer only indirect support for our model, they are helpful insofar as they provide further evidence of the generalizability of the current findings, with two stereotyped groups other than the elderly. In one study, participants holding either exceptionally strong pro-sorority or anti-sorority attitudes were asked to form an impression of a single group member who either was, or was not, identified as a group member. (In this study, the nature of group attitudes was confounded with ingroup/outgroup status, as none of the anti-sorority participants were themselves group members.) Results showed that the perceived typicality of the group member moderated application of group attitudes only among the anti-sorority participants. Among the “pro” participants, group attitudes were applied, regardless of target typicality. These findings are compatible with the implications of the present paper, if one assumes that the “pros” perceived the group to be more heterogeneous than the “antis”, perhaps due to their greater personal knowledge about the group (Linville, 1999). Nevertheless, because this study confounded three critical variables (group attitudes, perceived variability, and ingroup/outgroup status), it was impossible to discern

the unique and/or interactive effects of these variables in this study.

In a second study, group variability was measured more directly, this time regarding the perceived variability of the category *priests*. Although this study avoided some of the confounds noted above, the target behaviors in this study were highly extreme, as the priest was convicted of stealing. This resulted in a near-base-ment effect for perceived typicality such that most participants judged the target to be extremely atypical. (This study was conducted more than a decade prior to the recent set of allegations regarding the possible misconduct of Catholic priests.) Hence, this study did not allow a critical test of how group variability might vary as a function of whether the target was judged as relatively typical or atypical, a consideration critical to the present model. Nevertheless, this study, like the first one described above, provides some indirect support for our framework. Participants having relatively favorable expectations of priests, but believing them to be heterogeneous, were less harsh in their judgments of the target compared to other participants holding homogeneous views. In other words, pro-priest participants partially assimilated this highly atypical priest to their own initial expectations, but only when the group was perceived to be heterogeneous. This finding is compatible with our model, insofar as participants might be more likely to partially overlook the extremely atypical implications of his behaviors if they perceived the group to be relatively heterogeneous.

Relation of the present findings to the “Attitude strength” literature

As with many studies in the attitude literature, a major focus of the present research is on the extent to which generalized attitudes can subsequently predict more specific reactions, such as toward other persons (e.g., Lord et al., 1984). Research over the last 30 years has provided ample evidence that the predictive leverage of attitudes is likely to be moderated by many different factors. “Attitude strength” represents one important class of moderators, in the sense that strong attitudes should naturally be expected to better predict judgments and behavior than weak attitudes (Petty & Krosnick, 1995). There is currently some disagreement as to the exact conceptualization of attitude strength, but it appears to be a multidimensional construct, consisting (at least) of two dimensions, certainty and importance (Visser, Krosnick, & Simmons, 2003). Researchers have used different types of measures to operationalize attitude strength, including subjective reports as well as reaction-time (RT) measures (cf. Fazio, 1995).

On the one hand, it seems sensible to infer that people’s attitudes and opinions of a particular group might be stronger if its members are perceived to be homogeneous than if they are perceived to be heterogeneous.

⁴ We would like to thank an anonymous reviewer for alerting us to this possibility.

This seems especially true in the sense of confidence, because homogeneous categories should naturally foster a more cohesive and evaluatively uniform set of opinions/beliefs about the group than would otherwise be the case. Indeed, previous studies in the literature, as well as the supplemental data from Experiment 2, are consistent with this idea, insofar as people are less confident (and slower) in expressing their opinions about social categories if its members are heterogeneous than if they are homogeneous (Lambert et al., 1998; Ryan et al., 1996; see also Dijksterhuis & van Knippenberg, 1999).

In light of these findings, it would seem counterintuitive, or even paradoxical, to presume that heterogeneous categories might be associated with *more* pervasive use and application of group attitudes. Yet, this is precisely what the present research showed (cf. Tables 2 and 4). However, this apparent paradox is, in fact, illusory. Group variability appears capable of exerting two different kinds of effects on social judgment that appear to operate relatively independently of one another. One effect has to do with the confidence with which participants formulate beliefs and opinions about the group as a whole (cf. Lambert et al., 1998; Ryan et al., 1996). On the other hand, the other effect is more complicated, and has to do with the importance of a single exemplar's typicality during the impression-formation process. Stated differently, asking how group variability might affect the confidence of people's sentiments about the category as a whole is quite different from asking how this factor might spontaneously influence the importance of a single exemplar's typicality when perceivers are attempting to form an impression of this person.⁵

Directions for future research

We see at least two areas of research that merit further attention. First, social psychologists have recently

shown a great deal of interest in perceptions of entitativity, which appears to be correlated with, but distinct from, perceived variability (Lickel et al., 2000). Entitativity appears to be the more general of the two constructs, referring not only to the perceived variability of the group along a given trait dimension, but also to a number of other variables (e.g., perceptions of whether all the members of the group share a common goal or purpose). As researchers gain a better understanding of the similarities and differences between perceived variability and entitativity (along, perhaps, with familiarity and expertise), it will be important to document the unique consequences of these factors.

Second, the relevance of group variability to current work on implicit processes also arises in the context of a recent line of research by Livingston and Brewer (2002). In the aspect of their results most germane to current concerns, these researchers found that when White participants were presented with African American faces, it was only the faces containing typical—but not atypical—features that automatically activated negative evaluations. In one respect, this finding is broadly consistent with the work of Lord et al. (1984) insofar as a priori, category-based evaluations are more likely to be activated in the presence of typical features of the category. Nevertheless, as our research shows, the importance of typicality may depend on the perceived variability of the group. This suggests that the kinds of effects reported by Livingston and Brewer (2002) could, in principle, be moderated by perceived group variability, a variable not considered in their work. These and other lines of inquiry should be useful as social psychologists seek to further understand the complex and sometimes surprising consequences of group variability.

Acknowledgments

This research was supported in part by Grant No. SBR 9817554 from the National Science Foundation awarded to Alan Lambert and by a Social Sciences and Humanities Research Council of Canada research grant (410-2000-0017) awarded to Alison Chasteen.

References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage.
- Allport, G. W. (1954). *The nature of prejudice*. Cambridge, MA: Addison-Wesley.
- Ajzen, I. (1988). *Attitudes, personality, and behavior*. Chicago, IL: Dorsey Press.
- Bardach, L., & Park, B. (1996). The effects of in-group/out-group status on memory for consistent and inconsistent behavior of an individual. *Personality and Social Psychology Bulletin*, 22, 169–178.

⁵ If these two consequences of group variability are truly independent of each other, readers might wonder why, then, we did not consistently obtain stronger correlations between group attitudes and judgments of the target when the group is homogeneous, collapsing over perceived typicality of the target. In other words, setting aside differential attention to typicality, one might have expected stronger group-target correlations to emerge when the group is homogeneous. Although the present research did not clearly show evidence of this kind of effect, we have recently provided stronger evidence for it across two studies in the realm of racial attitudes (Lambert, Payne, Shaffer, & Ramsey, 2004). In particular, racial attitudes towards Blacks more reliably predicted reactions toward a single group member if the group was perceived to be homogeneous than if it was perceived as heterogeneous. (Interestingly, this moderation effect was more pronounced for implicit compared to explicit measures of racial attitudes.) Of more interest to present concerns, supplemental analyses on typicality ratings revealed a pattern generally similar to that shown in this paper. That is, these data, similar to the present findings, showed that typicality was more likely to play a “gatekeeper” role when the group was perceived as homogeneous. Further details regarding these latter analyses may be obtained from the first author.

- Brewer, M. B. (1988). A dual process model of impression formation. In T. K. Srull & R. S. Wyer (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.
- Dijksterhuis, A., & van Knippenberg, A. (1999). On the parameters of associative strength: Central tendency and variability as determinants of stereotype accessibility. *Personality and Social Psychology Bulletin*, 25, 527–536.
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences, and correlates of attitude accessibility. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude Strength: Antecedents and consequences* (pp. 247–282). Mahwah, NJ: Erlbaum.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior. An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation: from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). New York: Academic Press.
- Guinote, A., Judd, C. M., & Brauer, M. (2002). Effects of power on perceived and objective group variability: Evidence that more powerful groups are more variable. *Journal of Personality and Social Psychology*, 82, 708–721.
- Judd, C. M., & Park, B. (1988). Out-group homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology*, 69, 460–481.
- Judd, C. M., Ryan, C. S., & Park, B. (1991). Accuracy in the judgment of in-group and out-group variability. *Journal of Personality and Social Psychology*, 61, 366–379.
- Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin*, 129, 522–544.
- Lambert, A. J. (1995). Stereotypes and social judgment: The consequences of group variability. *Journal of Personality and Social Psychology*, 68, 388–403.
- Lambert, A. J. (1998). [Group attitudes, values, and group variability.] Unpublished raw data.
- Lambert, A. J., & Chasteen, A. L. (1997). Perceptions of disadvantage versus conventionality: Political values and attitudes toward the elderly versus Blacks. *Personality and Social Psychology Bulletin*, 23, 469–481.
- Lambert, A. J., Barton, L. R., Lickel, B., & Wells, J. (1998). The influence of group variability and processing goals on the ease of making judgments about social categories. *Personality and Social Psychology Bulletin*, 24, 807–820.
- Lambert, A. J., Payne, B. K., Shaffer, L. M., & Ramsey, S. (2004). On the predictive validity of implicit attitude measures: The moderating effect of perceived group variability. Manuscript submitted for publication.
- Lambert, A. J., & Wedell, D. H. (1991). The self and social judgment: Effects of affective reaction and “own position” on judgments of unambiguous and ambiguous information about others. *Journal of Personality and Social Psychology*, 61, 884–898.
- Lambert, A. J., & Wyer, R. S. (1990). Stereotypes and social judgment: The effects of typicality and group heterogeneity. *Journal of Personality and Social Psychology*, 59, 676–691.
- Lickel, B., Hamilton, D. L., Wierzchowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, 78, 223–246.
- Linville, P. (1999). The heterogeneity of homogeneity. In J. Cooper & J. Darley (Eds.), *Attribution processes, person perception, and social interaction: The legacy of Ned Jones*. Washington, DC: American Psychological Association.
- Linville, P. W., Fisher, G. W., & Salovey, P. (1989). Perceived distributions of the characteristics of ingroup and outgroup members: Empirical evidence and a computer simulation. *Journal of Personality and Social Psychology*, 57, 165–188.
- Lippman, W. (1922). *Public opinion*. New York: Macmillan.
- Livingston, R. W., & Brewer, M. B. (2002). What are we really priming? Cue-based versus category-based processing of facial stimuli. *Journal of Personality and Social Psychology*, 82, 5–18.
- Lord, C. G., Lepper, M. R., & Mackie, D. (1984). Attitude prototypes as determinants of attitude-behavior consistency. *Journal of Personality and Social Psychology*, 46, 1254–1266.
- Lord, C. G., Desforges, D. M., Ramsey, S. L., Trezza, G. R., & Lepper, M. R. (1991). Typicality effects in attitude-behavior consistency: Effects of category discrimination and category knowledge. *Journal of Experimental Social Psychology*, 27, 550–575.
- Manis, M., Paskewitz, J., & Cotler, S. (1986). Stereotypes and social judgment. *Journal of Personality and Social Psychology*, 50, 461–473.
- Miller, N., & Brewer, M. B. (1984). *Groups in contact: The psychology of desegregation*. New York: Academic Press.
- Park, B., & Hastie, R. (1987). The perception of variability in category development: Instance versus abstraction-based stereotypes. *Journal of Personality and Social Psychology*, 53, 621–635.
- 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.
- Petty, R. E., & Krosnick, J. A. (1995). *Attitude strength: Antecedents and consequences*. Mahwah, NJ: Erlbaum.
- Ryan, C. S., Judd, C. M., & Park, B. (1996). Effects of racial stereotypes on judgments of individuals: The moderating role of perceived group variability. *Journal of Experimental Social Psychology*, 32, 71–103.
- Sherif, M., & Hovland, C. I. (1961). *Social judgment: Assimilation and contrast effects in communication and attitude change*. New Haven: Yale University Press.
- Smith, E. E., Shoben, E. J., & Rips, L. J. (1974). Structure and process in semantic memory: A featural model for semantic decisions. *Psychological Review*, 8, 214–241.
- Verplanken, B., Jetten, J., & van Knippenberg, A. (1996). Effects of stereotypicality and perceived group variability on the use of attitudinal information in impression formation. *Personality and Social Psychology Bulletin*, 22, 960–971.
- Visser, P. S., Krosnick, J. A., & Simmons, J. (2003). Distinguishing the cognitive and behavioral consequences of attitude importance and certainty. *Journal of Experimental Social Psychology*, 39, 118–141.
- Vonk, R., & van Knippenberg, A. (1995). Processing attitude statements from in-group and out-group members: Effects of within-group and within-person inconsistencies on reading times. *Journal of Personality and Social Psychology*, 68, 215–227.
- Wyer, R. S. (1973). Category ratings as “subjective expected values”: Implications for attitude formation and change. *Psychological Review*, 80, 446–467.