Pergamon

Int. J. Intercultural Rel. Vol. 20. No. 3/4, pp. 305–321, 1996

0147–1767/96 $15.00 + 0.00

S0147-1767(96)00021-1

CATEGORY-EXEMPLAR DYNAMICS AND STEREOTYPE CHANGE

MYRON ROTHBART

University of Oregon

ABSTRACT. *Research relevant to Rothbart and John's (1985) [Journal of Social Issues,* **41**, 81–104] *model of stereotype change is examined. Contrary to predictions from the contact hypothesis, the attributes of category members frequently fail to generalize to the category as a whole. To account for this lack of generalization, Rothbart and John proposed that judgments about the attributes of a category are based, in part, on the attributes of the members most strongly activated by the category label. Embedded in this simple assumption, however, is the idea that as a category member becomes increasingly disconfirming of the stereotype, it decreases in its likelihood of being activated by the category. Category members who are too strongly disconfirming of the category are, in effect, not thought of as category members—a view that is consistent with prototype models of category structure. Current evidence for two critical assumptions is examined: (a) at the level of judgment, greater weight will be given to the attributes of good-than poor-fitting members of the category, and (b) typical category members are retrieved from memory more easily than atypical category members. In addition, evidence relevant to two implications of the model is also examined: (a) moderately disconfirming exemplars are more likely to change the stereotype than are strongly disconfirming exemplars, and (b) stereotypes should show considerable stability over time, given the tendency to "functionally isolate" highly disconfirming exemplars. The cultural images of groups tend to be both more extreme and more homogeneous than is warranted by reality. The implications of this discrepancy for understanding resistance to change through contact with group members are discussed. Copyright © 1996 Elsevier Science Ltd.*

Theory and research in the field of intergroup relations does not engender optimism. We find a plethora of theories predicting negative stereotypes and general hostility toward outgroups, but not many ideas for generating favorable attitudes toward other groups. One of the few grounds for optimism has stemmed from the contact hypothesis, predicting that favorable experiences with individual members of a

Requests for reprints should be sent to Myron Rothbart, Deparment of Psychology, 1227 University of Oregon, Eugene, OR 97403-1227, USA.
This research was supported in part by National Institutes of Mental Health Grant MH40662 to the author. I gratefully acknowledge the help of Deborah Day in running the experiments and the editorial comments of Mary Rothbart.

disliked group can generalize to the group as whole, thus changing the unfavorable stereotype.[1] The appeal of the contact hypothesis is due, at least in part, to the apparent truth of two straightforward ideas. First, there is the reasonable assumption that our experiences in the social world ought to change our views of that world. Second, simple logic dictates that there ought to be substantial correspondence between the attributes of a class and the attributes of the individual members that make up the class.

Both propositions, reasonable as they are, have been questioned. The first assumption, that our experiences in the world change our views of the world — presumably by moving our ideas closer to social "reality" — was questioned by Hovland's analysis of the paucity of attitude change found in field settings where subjects can selectively attend to, or even ignore, uncongenial information (Hovland, 1959). Hovland contrasted the high degree of attitude change found in laboratory experiments (usually conducted in university settings) with the very low degree of change associated with attitude surveys conducted in field settings, suggesting that real-world settings allow people to insulate themselves from information that challenges existing beliefs. The second assumption, that there is close correspondence between the attributes of the category and the attributes of the individuals that make up the category, was questioned by Sears (1983), who noted that we tend to be more favorable toward the individual members of the class than we are to the class as a whole. Thus, we may dislike "politicians", but we like particular congressmen; we may dislike Turks, but have favorable feelings toward every individual Turk we have known. This "semi-autonomous" relation between a category and its exemplars is a fundamental phenomenon to be understood by psychological theory.

Given that these two very reasonable assumptions have been challenged, it is not surprising that the research findings on the contact hypothesis itself have been somewhat conflicting. In a very important early summary of the voluminous research on intergroup contact, Amir (1976) noted that some research supported the beneficial effects of contact, but a large number of studies also showed either no effect, or an actual decline in favorable attitudes. Amir (1976), similar to Cook (1984), argued that the conditions under which contact has favorable effects may be constrained by a number of important factors, such as the nature and quality of the contact, the status relations between the groups, and so on. Moreover, Amir then issued a call for a better theoretical understanding of the processes underlying contact:

> Despite a substantial amount of research on ethnic contact, our theoretical understanding of what contact involves as a potential agent of change and

[1] There is no single, definitive form of the contact hypothesis, but the most central tenet seems to be that of generalization from the individual to the group. For a recent thoughtful review, see Pettigrew (1988).

what are the underlying processes is still very limited . . . The lack of basic theory is also exemplified in the little interaction between studies in ethnic contact and general theories of attitude formation and change, or theories in other psychological fields such as perception . . . (p. 289).

THEORETICAL MODEL

Rothbart and John (1985) attempted to address Amir's concern by noting the link between stereotype change induced by contact and the dynamic relation that exists between a category and its members. They proposed two formidable obstacles to contact-induced stereotype change. First, there is the problem of establishing the type of contact that may promote favorable and/or counterstereotypic impressions of individual outgroup members, which may be extremely difficult to realize under conditions of historical enmity (cf. Hewstone & Brown, 1986). Even if this necessary condition is satisfied, however, there is the additional problem of having the attributes of individual category members generalize to the category as a whole, which is the key assumption underlying the contact hypothesis. The failure to obtain this generalization is common (Cook, 1984; Kelman, 1992) and may well be a central factor inhibiting stereotype change.

It is the failure of generalization that Rothbart and John took as the fundamental problem requiring explanation. They argued that the problem of generalization could be understood as a result of the dynamic relation between the attributes of the category and the attributes of the exemplars that make up the category. "Dynamic relation" refers to the bidirectional influence that the category has on the exemplar, and that the exemplar has on the category. It is assumed that when people make judgments about the attributes of the category, those judgments are at least in part based upon the attributes of the exemplars that are activated by the category label (cf. Hintzman, 1986). It is further assumed that the degree of activation of an exemplar by the category is proportional to the overall goodness of fit between the attributes of the category and the attributes of the exemplar. Thus the attributes of the category, presumably determined by some complex mixture of cultural learning and experience with category members, are critical in determining which of the exemplars are most likely to be activated. However, it is also true that the attributes of the activated category members will then determine the attributes of the category as a whole. Indeed, it is this last assumption that is at the heart of the contact hypothesis—that the attributes of known individual members generalize to the category as a whole. The problem arises as a result of the premise that activation of an exemplar is proportional to its goodness of fit to the category. The interesting prediction from the contact hypothesis is that when the attributes of the exemplar

disconfirm the attributes of the category (the stereotype), the latter accommodate to the former. The "goodness of fit" assumption, however, which states that typical (good fitting) exemplars are more likely to be activated than atypical (poor fitting) exemplars, makes it unlikely that highly disconfirming exemplars will even be activated by the category label. In effect, it is as if poor-fitting exemplars are not psychologically treated as category members. In Rothbart and John's view, disconfirming exemplars are "functionally isolated" from the category as a whole.

To state the problem in a slightly different way, exemplars whose attributes disconfirm the stereotype have contradictory effects on stereotype change. On the one hand, the more disconfirming an individual exemplar's behavior is of the group stereotype, the greater the potential force exerted on the stereotype to change. On the other hand, stereotype-disconfirming information, associated with an exemplar, also decreases the "goodness of fit" between exemplar and category, reducing the likelihood that the exemplar will be activated (associated with the category), and thus reducing the likelihood of generalization. An exemplar that is highly disconfirming and also a very poor fit to the category may not be activated by the category at all, resulting in no generalization, whereas a moderately disconfirming exemplar, while less disconfirming, will be activated by the category to some degree, and result in some generalization. The Rothbart and John model indicates that the optimal level of disconfirmability for stereotype change is one that jointly maximizes both degree of disconfirmability and category activation. Since disconfirmability and category activation are, themselves, inversely related, however, the tendency for poor fitting members of a category to be dismissed, either as atypical exceptions or as "nonmembers" of the category, will be common, and may be an important way in which stereotypical beliefs remain insulated from change.

In summary, one of the central assumptions underlying this model is that not all exemplars of a category are equal in their ability to modify the attributes of a category. Good- and poor-fitting exemplars differ in their ability to be activated by the relevant category, and thus differ in the likelihood that their attributes will generalize to the category as a whole. Although a number of models of belief change, including the contact hypothesis, assume that degree of generalization should be independent of the degree of match between category and exemplar, that assumption is questioned by prototype models of inference (e.g., Rips, 1975) and by the Rothbart and John analysis.

The remainder of this paper deals with the research evidence relevant to the model. At present, there are a number of experiments that directly test some of the underlying assumptions of the model, while other research examines predictions generally consistent with the model, even though they may have been generated for purposes other than testing this model.

In general, we will first consider research that directly tests some of the model's underlying assumptions, and then examine research whose findings are consistent with the model's predictions. In the first set, we consider two basic assumptions: (a) when making inferences about a category, greater weight is given to the attributes of good- rather than poor-fitting members of the category, and (b) it will be more difficult to retrieve atypical category members than typical category members in response to a category retrieval cue. In the second set we consider the question of (a) which category members are most likely to change the stereotype and (b) the stability of stereotypic beliefs.

RESEARCH EVIDENCE

Typicality and Inferential Weight

To test the hypothesis that peoples' judgments about a category are disproportionately based on the attributes of the most typical members, it would be possible to present subjects with equal proportions of typical and atypical members of a social category (e.g., "corporate lawyers"), in which the typical members (e.g., who "enjoy jogging") have attributes different from atypical members (e.g., who "enjoy swimming") and then have subjects estimate the frequency of jogging and swimming for "corporate lawyers" as a whole. The problem with such a design centers around the "fuzziness" of social categories, in which the presence of some critical attributes may psychologically remove atypical group members from the category. That is, when using social categories, pairing a category member with a highly atypical attribute may remove that individual from the category (e.g., by pairing the attribute "poor" with corporate lawyer may seem oxymoronic—perhaps it is an imprisoned or disbarred corporate lawyer, but in either case, it is a nonpractising corporate lawyer).

It should be clear that, even for fuzzy categories, not all atypical attributes "remove" the exemplar from the category (corporate lawyers from Iowa are less typical than those from New York), but the effect of pairing an exemplar with atypical attributes is likely to reduce the perceived strength of category membership. The strongest test of the model would occur when even the atypical members are unambiguously perceived as belonging to the category.

To test a strong form of the assumption that judgments about the attributes of a category are disproportionately influenced by the "good-fitting" or typical members of the category, Rothbart and Lewis (1988) presented subjects with "good" and "poor" examples of rectangles, triangles, ellipses and pentagons, with one color associated with the good (typical) and a different color associated with the poor (atypical) shapes

within each category. When subjects were later asked to estimate the frequency of colors associated with, say, triangles, they estimated the color associated with the good-fitting triangles to be more common than the color associated with the poor-fitting triangles, even though both colors were presented equally often. The use of geometric shapes was considered a stringent test of the idea that good-fitting members are given disproportionate weight in judging category attributes, since mathematical categories, unlike social categories, possess defining rather than characteristic features. By using mathematical concepts, such as rectangle and triangle, there is no doubt about the membership class of the individual exemplars.

In the experiment on geometrical shapes, subjects were presented with all members of the category and asked to summarize the attributes of the category as a whole. In another experiment, using a social category, subjects were given information about the prototypicality of a single category member and then asked to estimate the likelihood that the category, as a whole, possessed the same attribute as the individual category member. Subjects read about a single University of Oregon fraternity member, who was elaborately described in a way that indicated a poor, moderate, or good fit to the fraternity stereotype and who had voted for either Mondale or Reagan in the United States' presidential election of 1980.[2] The tendency for subjects to infer that the fraternity, as a whole, voted in the same manner as the individual fraternity member was proportional to the prototypicality of that group member. For both experiments, using nominal as well as social categories, subjects accorded typical category members greater importance than atypical category members, for both frequency judgments and inductive inference.

Typicality and Retrieval

The previously described experiments suggest that greater weight is given to good- versus poor-fitting examples in the judgment process. However, the model clearly states that typical members of the category ought to be retrieved more easily than atypical category members. A recent paper by Rothbart, Sriram, and Davis-Stitt (1996) presented three types of experiments in support of this assumption. In their first experiment, subjects learned about the category membership of 12 male stimulus persons. Four of the men were Black, four were gay, four were Asian, and four were fraternity members. After correctly learning the category of each member, the strength of association between category

[2]In general, typical fraternity members are viewed at the University of Oregon as boisterous, gregarious, athletic and academically nonserious. Prototypicality was manipulated by varying the number of stereotypic traits present in the description of the stimulus person.

and exemplar was tested by first presenting on a computer screen a category name (e.g., "Asian") for a brief duration, followed by one of the 12 names that either was or was not a member of the category. The subject was to respond, in a reaction time paradigm, by pressing a "yes" key if the pairing was correct, and a "no" key if the pairing was incorrect. Since the category name acts as a priming stimulus that activates the names of the category members, the speed with which the subjects correctly verify the category member is taken as a measure of the strength of association between category and category member. Following this first reaction time measure, subjects then learned new information about each stimulus person, which was stereotypically consistent with, inconsistent with or irrelevant to the category. Again, subjects were tested for the strength of association between the category and the exemplar, using the same reaction time procedure as before, in which the category name is followed by the name of a category member. The difference between the first and second reaction time was calculated and taken as a measure of the change in strength of association between the category and the category member as a function of learning of new information. It was predicted, consistent with the dynamic view of category–exemplar relations, that new stereotypical information should strengthen the category exemplar association (i.e., speed up the second over the first RT), while the inconsistent information should weaken that link (i.e., slow down the second over the first RT), in comparison to the irrelevant control. This is exactly the pattern of results that was found (see Figure 1). These results have two implications for thinking about the issue of stereotype change. First, they corroborate the assumption that the effect of pairing an exemplar with information disconfirming of the stereotype distances the exemplar from its parent category, making it less likely to be activated in response to the category name. Second, the data say something about the effects of individuating information on stereotypic perception. There has been considerable speculation that stereotypical thinking manifests itself only in the absence of individuating information (e.g., Locksley, Borgida, Brekke & Hepburn, 1980; Locksley, Hepburn & Ortiz, 1982). Research on the dilution effect (Nisbett, Zukier & Lemley, 1981), for example, suggests that irrelevant information dilutes the power of diagnostic information to predict a criterion. To use an example from the dilution research, if we learn that Joe is an alcoholic, we are likely to predict that Joe physically abuses his children, since the attribute "alcoholic" is diagnostic of the criterion "child abuse." The dilution effect occurs when information that is irrelevant to the criterion, such as "works as a shoe salesman," is paired with diagnostic information ("is alcoholic"), to reduce the predicted likelihood of child abuse. The putative mechanism for this effect is that nondiagnostic or irrelevant information has the effect of reducing the degree of similarity between the agent (Joe) and the action (child abuse)

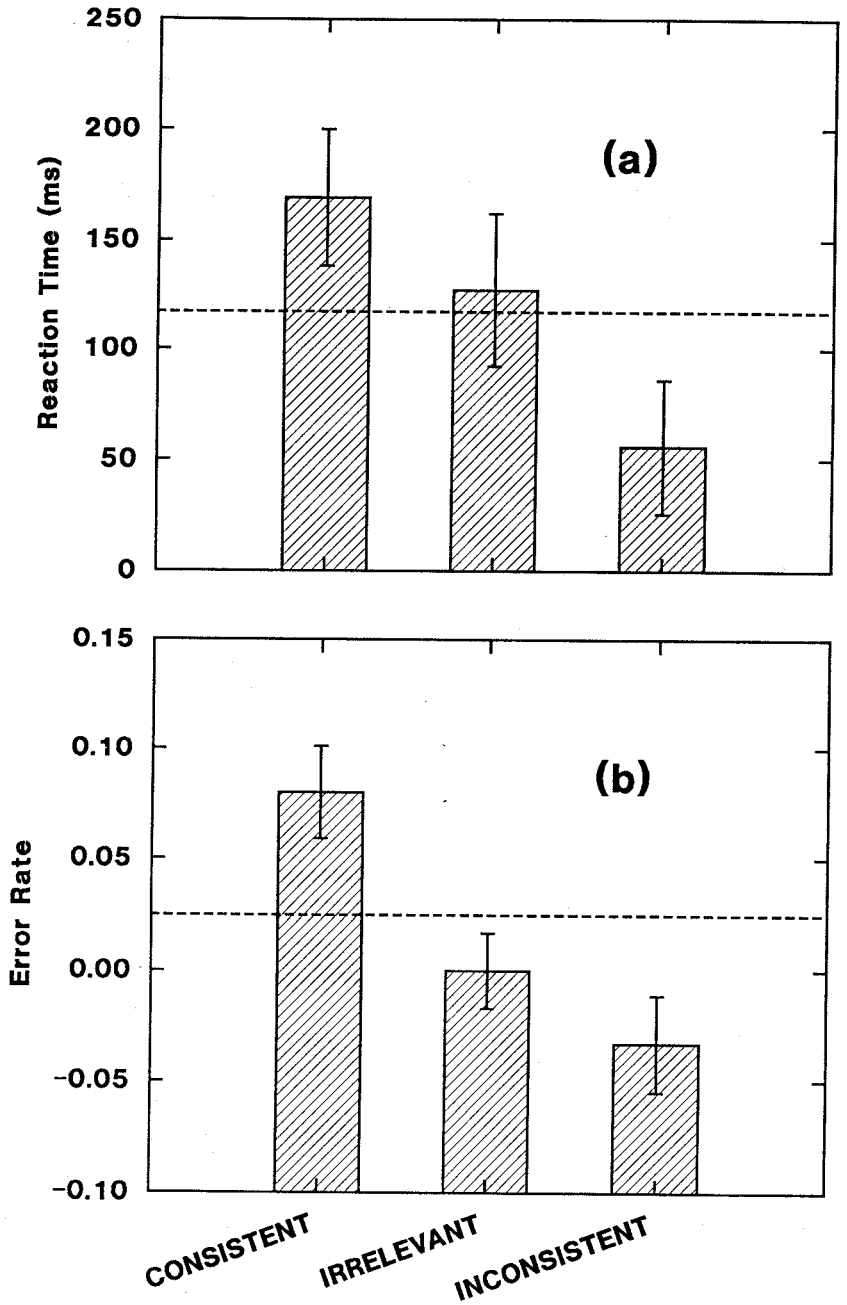

FIGURE 1. Reaction time and error rates for differences between Phase 1 and Phase 2 as a function of type of information associated with exemplar (experiment 1).

by increasing the presence of non-overlapping attributes. The present findings suggest a different interpretation of the dilution effect. Although the irrelevant information (e.g., "works as a shoe salesman") is not predictive of the criterion (child abuse), it may be quite relevant to interpreting the strength and meaning of the diagnostic information ("is alcoholic"). The stereotype of the alcoholic (in contrast to the reality of alcoholics) is one that includes binge drinking, memory loss, inability to hold a job, incoherence, etc. Information that is related to this stereotype, such as holding a job as a shoe salesman, may render Joe a less strong, or less typical alcoholic, than would be the case if that information were absent. The findings from our first experiment suggest that information irrelevant to the criterion should be diluting only when that information is inconsistent with the stereotype (as holding a job is inconsistent with our image of alcoholics). If that information were consistent with the stereotype (e.g., in the way that "has trouble remembering names" would be consistent with alcoholic amnesia), we would expect not a dilution of the stereotype, but an increased predictability based on the stereotype.

A second set of studies addressed the issue of the ease of retrievability of stereotypically neutral information, when that information is paired with individuals who are typical or atypical of the category (Rothbart et al., 1996, experiments 3 and 4). The argument is that when subjects search for information (even neutral information) associated with individuals, they will be able to access that information best when it is paired with individuals typical of the category. In these experiments, typicality is defined by a particular combination of social categories. Subjects learned about the hobbies and occupations of a number of men and women, where half the occupations were scientific in nature (physicist, metallurgist, etc.) or involved teaching (kindergarten, music, etc.). Pretesting indicated that, for scientists, males were considered typical and females atypical, while for teaching, females were considered typical, while males were atypical. The hobbies engaged in by the stimulus persons, were chosen to be independent of both gender and occupation, and constituted the neutral information later to be retrieved by subjects. After the learning phase, subjects were given a category prime (male, female, scientist, teacher) followed by a specific hobby. The subjects' task was to verify whether the person who engaged in the specific hobby (e.g., swimming) was or was not a member of the primed category. It was predicted, and found, that subjects could verify the hobby more quickly and accurately when the hobby was associated with a typical, rather than an atypical, member of the primed category.

Both the first and second set of studies show a greater ease of retrieving typical than atypical exemplars. In the first study, the "distance" between category and exemplar itself was influenced by experimentally pairing an exemplar with information that confirmed, disconfirmed or was irrelevant

to the stereotype. In the second set of studies, the typicality of the exemplar was determined at encoding by the intersection of gender and occupation, and even neutral information was more difficult to retrieve when associated with atypical category members. Both experiments presented information to subjects in a highly artificial setting, and there is a legitimate issue about the applicability of these laboratory findings to real life settings.

It is possible, of course, to examine the relationship between retrievability and typicality in more natural settings, but there are limitations to that design as well. For very large groups, we may only have knowledge of the typical examples of the category, thus creating a spurious relation between recall and typicality. The virtue of the laboratory setting, as artificial as it is, is that we can control the amount and frequency of information for all group members, typical and atypical alike, so that subjects have the same information about, and are equally familiar with, all group members. In real-life settings, we can be sure that there is not equal information about all group members, and at least for liked groups, there is a natural discrepancy between familiarity and typicality. That is, for a liked group, the typical members will be liked more than the atypical members, and since we are more likely to become familiar with liked than disliked, persons, the typical group members will be more familiar as well. Since familiarity is a strong determinant of retrieval, it will be important to partial out the effects of familiarity and liking when assessing the relation between typicality and retrieval.

By examining the retrieval of members of a disliked group, however, it is easier to unconfound familiarity, liking, and typicality, since the most typical group member was probably the least liked (and therefore the least familiar). This was in fact the second study conducted by Rothbart et al. (1996). Members of a fraternity were asked to think about a specific rival, disliked fraternity with which they had considerable contact, and then to list the members of that fraternity. After completing this recall task, subjects were given the entire list of members of the target fraternity and were asked to rate each member on liking, familiarity and typicality for their group. They were asked to rate those members about whom they had some knowledge. First, it is worth noting that, even for this relatively small target fraternity, with which subjects had a great deal of contact, subjects could list only a relatively few group members. Out of a maximum possible total of 44 group members, the range in number of group members recalled was from 3 to 13, with a mean of 4.9. Thus, the issue of differential knowledge of group members is real. A total of 22 different target persons were recalled, and the correlation between the number of subjects who recalled a given target and the average judged typicality of the target was $r = .49$ ($p < .05$), after the combined effects of liking and familiarity had been partialled out. Thus the predicted

correlation between magnitude of recall and typicality was confirmed, at least when retrieving members of a disliked group.

Which Group Members are Most Likely to Change the Stereotype?

Given the circular relationship between a category's tendency to activate the most typical exemplars, and the tendency to infer a category's attributes from the attributes of the most activated exemplars, how does disconfirming information ever become introduced into the stereotype? The general answer accorded by the model is that the exemplar to which the disconfirming information is associated must be strongly activated by the category. Given that the disconfirming information reduces the goodness of fit, and therefore the level of activation, this can occur in a number of different ways. First, the disconfirming attributes can be embedded in an exemplar that is an otherwise good fit to the category. This suggests that moderately disconfirming exemplars should have a stronger disconfirming effect on the stereotype than do highly disconfirming exemplars. Second, aspects of the exemplar that are confirmatory of the stereotype can be made particularly salient, also increasing the likelihood of association with the category. Third, it may be possible through a number of contextual variables, some quite subtle, to influence the perceived link between exemplar and category. Social context, task demands, and linguistic usage often serve to delineate group membership, which is one way of strengthening the link between an individual and its parent category.

The prediction that moderately disconfirming exemplars have a stronger influence on stereotype change than do strongly disconfirming exemplars has considerable support. Weber and Crocker (1983, experiment 3) presented subjects with information about "a random sample" of 30 corporate lawyers. Ten of the 30 lawyers were associated with information disconfirming of the stereotype, and for one condition, the 10 were otherwise typical of the category (white and affluent), and in the other case were atypical of the category (black and poor). They found that the disconfirming attributes were most likely to be attributed to corporate lawyers "as a whole" when the disconfirming information was associated with the typical (representative) than the atypical (unrepresentative) group members. In a fourth experiment using a somewhat similar design, Weber and Crocker (1983) presented all subjects with the same amount of disconfirming information about corporate lawyers, but presented in one of two ways: (a) in a dispersed condition, every group member possessed one counter-stereotypic attribute, or (b) in a concentrated condition, one-third of the group members each possessed three counter-stereotypical attributes. Comparable to the previous study, the disconfirming attributes were most likely to generalize to corporate lawyers as a whole when the

disconfirming information was dispersed rather than concentrated. Both studies support the prediction that information associated with moderately disconfirming rather than extremely disconfirming individuals will be more likely to generalize to the group as a whole. In experiment 3, the disconfirming information attached to the atypical individuals was least likely to generalize, ostensibly because the presence of atypical attributes allowed subjects to "subtype", functionally treating the atypical group members as nonmembers of the category. However, experiment 4, which did not attach the disconfirming information to exceptional group members, also found less generalization when disconfirming information was concentrated in a few group members. It is certainly possible that subtyping also occurred in experiment 4, where extremity or atypicality itself became the basis for subtyping, but it is also possible to interpret the findings of both experiments as supporting the idea that generalization is greater for moderately than for extremely disconfirming category members. More generally, it should be clear that "functionally isolating" members of a group and "subtyping" are not necessarily identical processes. Conceptually, it is possible, for example, to have three group members who are each atypical of the category, but atypical in different ways that do not allow grouping under a common rubric. Subtyping should involve, in addition to atypicality, a common grouping of the atypical members, as would be the case when the atypical members share common attributes. Both involve isolation of poor-fitting cases, and both have the effect of preserving the existing stereotype, but subtyping involves, additionally, a grouping of the atypical exemplars.

Hewstone, Macrae, Griffiths, Milne and Brown (1994) extended the paradigm developed by Weber and Crocker (1983), and generally confirmed the findings that generalization was more likely to take place when disconfirming information was dispersed across many group members, rather than concentrated in a few members. They also reported that the judged typicality of the group members predicted the likelihood of generalization of disconfirming information, confirming one of the central predictions of the Rothbart and John model. Kunda and Oleson (1995) proposed what appears to be a more motivational model of subtyping, in which any attribute will be used to isolate group members associated with disconfirming information. Thus, disconfirming information associated with lawyers will be less likely to generalize when those group members are associated with either a large law firm or small law firm, in contrast to a condition where no information about the size of the firm is presented.

The prediction that contextual factors can influence the perceived "goodness of fit" to the group was borne out in an experiment conducted by Maurer, Park and Rothbart (1995). They presented all subjects with identical information about the members of a group. In one condition, subjects were encouraged to think about the group in terms of each

member's "goodness of fit" to the group as a whole; in another condition, they were asked to sort the members into subcategories that minimized differences within subcategories, and maximized differences between subcategories. In a baseline control condition, subjects were given no organizing instructions. In the "goodness of fit" condition, subjects viewed the group in the most stereotypical and most homogeneous way, and viewed the group in the subcategorizing condition in the least stereotypical and most heterogeneous way. It was hypothesized that subjects in this study were defining the implicit group boundaries in different ways, with the "goodness of fit" subjects functionally excluding the disconfirming exemplars from the superordinate category. Subcategorizing subjects, on the other hand, were presumably retaining the disconfirming exemplars within the superordinate category, and viewing the group in an appropriately complex manner. One measure of "psychological exclusion" is the difference between the typicality ratings for the good and poor fitting group members, which was found to be much larger in the "goodness of fit" condition than in the subcategorizing condition. The important role of typicality in defining the boundaries of the group is consistent with the Hewstone et al. (1994) findings. One important implication of the results of this study is that typicality itself is not invariant, and can be influenced by the subject's implicit or explicit organizing strategies.

Does the Model Predict too Little Stereotype Change?

Since the attributes ascribed to a group are based, in part, on the attributes of the most highly activated exemplars, and the most highly activated exemplars are, in turn, based on "goodness of fit" to the existing stereotype, the model depicts a highly inertial system—one that is slow to change in the face of disconfirming information. Does the model predict too much stability? Hovland's analysis of the difference between laboratory and field studies in both terms of selective exposure, and differences in time course may provide an answer to this question. Studies in the tradition of Weber and Crocker (1983), Hewstone et al. (1994), and Kunda and Oleson (1995), often suggest that even a single exposure of a few (often as few as a single) exemplars results in stereotype change.

A 4-yr naturalistic longitudinal study conducted by Rothbart and John (1993) presents a less optimistic picture of stereotype change. The stereotypes of entering college freshman were assessed on four occasions, with the final assessment completed during the students' senior year. In all, impressions of 14 different stimulus groups were assessed (Blacks/Black students, Asians/Asian students, Jews/Jewish students, gay men, lesbians, fraternity men, sorority women, male/female professors, men

and women), although any given subject judged only three of the groups. For each target group, subjects judged the applicability of 40–45 trait descriptive terms relevant to that group (a different set of trait terms was constructed for each group).

The issue of change can be addressed in a number of ways, but one measure is the test–retest reliability, computed over the 4-yr period, between subjects' mean trait attributions between their freshman and senior years. That is, for the 40 + traits relevant to each group, the mean attributions for the traits during the first assessment were correlated with the same mean attribution during the last assessment (where the number of cases is equal to the number of trait terms). For comparison, an independent group of subjects made the same ratings of the same 14 groups, but the test–retest period was separated by 1 week. Assuming the reliability over 1 week reflects an error of measurement rather than systematic change, we use the 1-week results as a baseline measure. Averaged across the 14 groups, the average correlation was .96 for the 1-week test–retest condition. For the longitudinal sample, the correlation over the 4-yr period was .92. Although test–retest reliability is only one measure of change — consistency in ordering of the trait terms — it indicates, none the less, tremendous stability (1) over a rather long period, and (2) under conditions in which many entering students are being exposed to groups and group members whom they had not known previously.

Results from the longitudinal sample are reminiscent of Hovland's analysis of survey data suggesting little attitude change — in contrast to frequently observed change in laboratory settings. It is possible that the same factors, such as the presence of selective exposure in naturalistic settings and the longer time course of field studies, all contribute to a lack of change. At present, the longitudinal data are generally consistent with the suggestions of Rothbart and John, who argue for an "inertial" belief system that is slow to change in the presence of disconfirming information. A major problem, of course, is that we do not know to what degree our longitudinal subjects encountered group members who disconfirmed their original stereotypes. It is plausible to assume that, for largely unfamiliar groups, such as blacks, jews, gays and lesbians, the stereotypes would be subject to significant disconfirmation, but those stereotypes were, if anything, more stable than was the case for familiar groups. We now turn to recent research on the issue of stereotype accuracy, which addresses the question of discrepancy between reality and perception more directly.

Speculations on the Discrepancy Between Perception and Reality

Research on the accuracy of social stereotypes (Judd, Ryan & Park, 1991; Judd & Park, 1993) supports the idea that there is a notable discrepancy between social reality and the perception of that reality.

Stereotypes tend to be too simple and too extreme, with respect to both perceived central tendency and perceived variability. If there is a discrepancy between perception and reality, should the perception not change, however slowly, to become concordant with reality? Lewis (1990) argues that such discrepancies are not necessarily self-correcting. If the image of the group is too idealized and extreme, then most exemplars, being highly discrepant from the stereotype, might be too distant from the stereotype to be activated by the category, and thus "functionally isolated" from the stereotype in all the ways previously indicated. Of course, the ability of the stereotype to remain uninfluenced by the exemplars depends on the magnitude of discrepancy between category and exemplar, which is not precisely spelled out by the model.

A current trend that may exacerbate the disparity between category and exemplar is the highly distilled form in which we receive information about groups through the media. Even under the best circumstances, it is difficult to convey the full complexity of a group's attributes. The tremendous time constraints placed on the media when describing complex groups, be it intergroup conflict in the Balkans or the characteristics of Americans on welfare, insure that it will be difficult to accurately convey a "modal" group member along with some notion of deviation from that mode. This is not so much the issue of bias in central tendency, where, for a number of reasons, the media may be predisposed to report the most dramatic, or most negative aspects of a group, but rather it is the question of how much range or variability around a central tendency can be conveyed by news reports that have to report the characteristics of variegated human groups in a matter of minutes or even fractions of a minute. It is daunting to consider the changes, over the past century, in the amount, nature, and quality of information that reach us about other groups. The competition between news organizations for a mass audience, the speed of modern global communication and possibly the short attention span of the modern television audience, combine to give us brief, strong, if not inaccurate, images of groups with whom we will likely never have contact. In the past, the groups we were most likely to learn about were also groups with whom we had a modicum of contact, allowing for the possibility that contact would challenge the simplicity of culturally generated stereotypes.

The model of stereotype change presented in this paper does not suggest that intergroup contact is ineffective, but it does suggest that the amount of change that can be expected through contact may be limited, given the mechanisms that allow us to "functionally isolate" or compartmentalize our experience with individual group members. Indeed, this limitation seems to be consistent with some of the major reviews of the effects of contact (Amir, 1976; Brewer & Miller, 1984; Cook, 1984; Hewstone & Brown, 1986). Assuming, however, that the discrepancy between the

stereotype and the individual members comprising the stereotype is of modest size, there is reason to be optimistic that the stereotype will move slowly, but inexorably, in the direction of reality. There is a concern, however, that if the discrepancy becomes too large, individual group members, by virtue of their atypicality, may fail to influence the general stereotype, which then remains insulated from experience. Recent trends toward the dramatization of news reporting lead us to be quite concerned about the effects of extreme and simple reporting, particularly when describing groups that are quite diverse.

REFERENCES

AMIR, Y. (1976). The role of intergroup contact in change of prejudice and ethnic relations. In P. Katz (Ed.), *Toward the elimination of racism* (pp. 245–308). New York: Pergamon Press.

BREWER, M. B., & MILLER, N. (1984). Beyond the contact hypothesis: Theoretical perspectives on desegregation. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 281–302). New York: Academic Press.

COOK, S. W. (1984). Cooperative interaction in multiethnic contexts. In N. Miller & M. B. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 155–185). New York: Academic Press.

HEWSTONE, M., & BROWN, R. (1986). Contact is not enough: An intergroup perspective on the 'Contact Hypothesis.' in M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 1–44). Oxford: Basil Blackwell.

HEWSTONE, M., MACRAE, C. N., GRIFFITHS, R., MILNE, A. B., & BROWN, R. (1994). Cognitive models of stereotype change: (5) Measurement, development and consequences of subtyping. *Journal of Experimental Social Psychology*, **30**, 505–526.

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

HOVLAND, C. I. (1959). Reconciling conflicting results derived from experimental and survey studies of attitude change. *The American Psychologist*, **14**, 8–17.

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

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.

KELMAN, H. C. (1992). Coalitions across conflict lines: The interplay of conflicts within and between the Israeli and Palestinian communities. In S. Worchel & J. A. Simpson (Eds.), *Conflict between people and groups* (pp. 236–258, 293–294). Chicago: Nelson-Hall.

KUNDA, Z., & OLESON, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, **8**, 565–580.

LEWIS, S. B. (1990). *Ideals as prototypes: The role of representational complexity.* Doctoral dissertation. University of Oregon, Eugene.

LOCKSLEY, A., BORGIDA, E., BREKKE, N., & HEPBURN, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology*, **39**, 821–831.

LOCKSLEY, A., HEPBURN, C., & ORTIZ, V. (1982). Social stereotypes and judgments of individuals: An instance of the base-rate fallacy. *Journal of Experimental Social Psychology*, **18**, 23–42.

MAURER, K. L., PARK, B., & ROTHBART, M. (1995). Subtyping versus subgrouping processes in stereotype representation. *Journal of Personality and Social Psychology*, **69**, 812–824.

NISBETT, R. E., ZUKIER, H., & LEMLEY, R. E. (1981). The dilution effect: Nondiagnostic information weakens the implications of diagnostic information. *Cognitive Psychology*, **13**, 248–277.

PETTIGREW, T. F. (1988). The intergroup contact hypothesis reconsidered. In M. Hewstone & R. Brown (Eds.), *Contact and conflict in intergroup encounters* (pp. 169–195). Oxford: Basil Blackwell.

RIPS, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, **14**, 665–681.

ROTHBART, M., & JOHN, O. P. (1993). Intergroup relations and stereotype change: A social-cognitive analysis and some longitudinal findings. In P. M. Sniderman, P. E. Tetlock and E. G. Carmines (Eds.), *Prejudice, politics, and race in America* (pp. 32–59, 307–332). Stanford: Stanford University Press.

ROTHBART, M., & JOHN, O. P. (1985). Social categorization and behavioral episodes: A cognitive analysis of the effects of intergroup contact. *Journal of Social Issues*, **41**, 81–104.

ROTHBART, M., & LEWIS, S. (1988). Inferring category attributes from exemplar attributes: Geometric shapes and social categories. *Journal of Personality and Social Psychology*, **55**, 861–872.

ROTHBART, M., SRIRAM, N., & DAVIS-STITT, C. (1996). Typicality, category–exemplar relations, and stereotyping. *Journal of Personality and Social Psychology*, **32**, 1–29.

SEARS, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology*, **44**, 233–250.

WEBER, R., & CROCKER, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, **45**, 961–977.