

ACADEMIC  
PRESS

Journal of Experimental Social Psychology 38 (2002) 443–458

Journal of  
Experimental  
Social Psychology[www.academicpress.com](http://www.academicpress.com)

## Contrast effects in stereotype formation and change: the role of comparative context

Natalie A. Wyer,<sup>a,\*</sup> Melody S. Sadler,<sup>b</sup> and Charles M. Judd<sup>b</sup><sup>a</sup> Department of Social Psychology, Free University Amsterdam, van der Boechorststraat 1, 1081 BT Amsterdam, Netherlands<sup>b</sup> University of Colorado, Boulder, CO, USA

Received 24 May 2001; revised 9 November 2001

### Abstract

Two experiments investigated the way in which the presence of a comparative or inter-group context during stereotype formation affects stereotype change, induced by subsequent disconfirming information. Participants learned about a focal group, after learning about one of the two context groups. After reporting their stereotypes about both groups, participants learned additional information about the focal group. This information described new group members who either confirmed or disconfirmed the group stereotype. Consistent with previous research, participants formed more extreme stereotypes about the focal group on dimensions that distinguished it from the context group (i.e., a contrast effect). In response to the subsequently presented disconfirming group members, a greater stereotype change was observed on dimensions that distinguished the focal group from the context group than on dimensions it did not. We argue that these effects are due to differences in perceived typicality of disconfirming group members. © 2002 Elsevier Science (USA). All rights reserved.

Attempts by social psychologists to identify ways in which group stereotypes can be effectively changed have met with mixed success. Most of these attempts have explored the conditions under which stereotype-inconsistent information (in the form of behaviors or traits shown by group members) reduces stereotypic views of the group (e.g., Johnston & Hewstone, 1992; Kunda & Oleson, 1995; Rothbart, 1996; Rothbart & John, 1985; Weber & Crocker, 1983). The focus in this literature has largely been on characteristics of the stereotype-inconsistent information that facilitate or hinder stereotype change.

To gain a better understanding of stereotype change, we argue in this paper that it is important to better understand the factors that affect stereotype formation and content in the first place. That is, we argue that change in response to stereotype-inconsistent information may importantly depend on the nature and content of the stereotype that one is attempting to change.

Group stereotypes have been shown to be context-dependent. Beliefs about a particular group depend

heavily on the comparative context in which those beliefs are salient. Our central thesis is that stereotype change in response to stereotype-inconsistent information also depends in predictable ways upon this comparative context.

To develop the rationale for this hypothesis, we begin by reviewing what is known about the role of context in stereotype formation. In light of these effects, we then outline some ways in which these context effects may influence responses to stereotype-inconsistent information. Finally, we present the results of two studies designed to test these ideas.

### Context effects on group stereotypes

The presence of a salient comparative context during stereotype formation has been shown to have two different, albeit related, effects. First, learning about a social category in the presence of a comparative category leads to the *accentuation of inter-group differences*. This effect is a robust one that numerous studies have shown. Second, learning about a social category in the presence of a comparative category leads to the *accentuation of*

\* Corresponding author.

E-mail address: [n.wyer@psy.vu.nl](mailto:n.wyer@psy.vu.nl) (N.A. Wyer).

*intra-group homogeneity*. As we discuss below, the empirical evidence in support of this second effect is considerably more mixed.

The combination of these two effects has been used to define what theorists in the Social Identity tradition refer to as the meta-contrast ratio (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987): “People use characteristics that differentiate between groups as a crucial piece of information in categorization—people are considered to be members of the same group to the extent that they are different from other groups, but are similar to each other, on some relevant dimensions.” These theorists have argued that the presence of a comparative context group leads to more polarized stereotypes (Oakes, Haslam, Turner, Worchel, & Morales, 1998). This hypothesis has been empirically supported in work by Oakes, Haslam and co-workers (Doosje, Haslam, Spears, Oakes, & Koomen, 1998; Haslam & Turner, 1995; Oakes, Haslam, Reynolds, Abrams, & Hogg, 1999).

How do these contrast effects come about? Thinking about a focal group in the presence of another comparative focal group seems to affect inter-group differentiation at a number of different levels. First, at a relatively early stage, more attention is paid to information that differentiates the target and context groups from each other (Trope & Mackie, 1987). This differential attention means that group stereotypes are more likely to consist of attributes that differentiate a group from its comparative context (see Judd & Park, 1993; McCauley & Stitt, 1978). Ford and Stangor (1992) reported that when participants learned about two novel groups that differed on one trait dimension but not on another, they were more likely to spontaneously mention the differentiating dimension than the non-differentiating one when later describing the groups. In addition, Babey (1999) showed that participants drew trait inferences about a group more quickly and based on less information if they learned about that group in the context of a second one than if they did not have that context.

As a result of greater attention to differentiating dimensions, the presence of a context may lead to more extreme or stereotypic judgments of a group on those dimensions. This accentuation of stereotypic differences has been shown both with purely perceptual categories (e.g., Corneille & Judd, 1999; Tajfel & Wilkes, 1963) and with more complex social categories (e.g., Eiser, 1971; Ford & Tonander, 1992; Krueger & Rothbart, 1990; Schell, 1997).

As we suggested above, the evidence for accentuation of intra-group homogeneity in the presence of a comparative context is considerably more mixed. For example, the classic study of Tajfel and Wilkes (1963) found no evidence that categorization increased intra-category similarity. More recently, Krueger and Roth-

bart (1988) also failed to find such an evidence. On the other hand, McGarty and Penny (1988) showed that opinion statements within a category were judged as more similar to each other when a categorization was made salient to participants. Rothbart, Davis-Stitt, and Hill (1997) reported similar results in research that involved more socially relevant stimuli. Somewhat surprisingly, recent work by Livingston, Andrews, and Harnad (1998) reported evidence for intra-group accentuation of homogeneity in the absence of inter-group differentiation.

### Context and stereotype change

The most typical approach to changing people’s stereotypes about a social group is to expose them to group members who disconfirm the stereotype. This has been done either by exposing perceivers to actual group members in inter-group contact settings (e.g., Brown, Vivian, & Hewstone, 1999; Wilder, 1993) or to written descriptions of hypothetical group members (e.g., Johnston & Hewstone, 1992; Kunda & Oleson, 1995; Weber & Crocker, 1983). Although contact with members of derogated outgroups can sometimes lead to more positive *attitudes* towards those groups (Pettigrew & Tropp, 2000), change in the group *stereotype* seems to be a much more elusive result. For disconfirming information to lead to a stereotype change, a number of important conditions seem to be necessary.

Most importantly for the present argument, the disconfirming group member must actually be seen as a relevant group member. They must be typical of the group as a whole in most ways, even while disconfirming the group stereotype to some extent (Rothbart & John, 1985). Otherwise, the disconfirming member is likely to be “fenced off” (Allport, 1954) or subtyped (Johnston & Hewstone, 1992; Maurer, Park, & Rothbart, 1995; Rothbart & John, 1985; Weber & Crocker, 1983). A subtyped individual is one who is not considered a “real” group member and consequently tends to be excluded from consideration when reporting the group stereotype. Thus, stereotype change in response to a disconfirming group member is unlikely to occur if the individual is subtyped.

What factors influence whether or not a disconfirming group member is subtyped? There is ample evidence to show that an important factor is the degree to which they disconfirm the group stereotype. For instance, Weber and Crocker (1983) found that participants who learned about group members who were partially atypical of the group stereotype changed their beliefs about the group. On the other hand, participants who learned about group members who were entirely atypical of the group did not change their stereotypes at all. Thus, the more extreme the stereotype disconfirmation, the less

likely the stereotype is to change (see also Johnston & Hewstone, 1992; Johnston, Hewstone, Pendry, & Frankish, 1994; Kunda & Oleson, 1995, 1997).

These earlier studies varied the degree of disconfirmation by manipulating characteristics of the disconfirming group member. In other words, those whose stereotype was changed saw a less extreme disconfirming group member than whose stereotype was not changed. We suggest that the exact same disconfirmer may be seen as more or less extreme by manipulating characteristics of the group stereotype, namely the comparative context in which that group stereotype is formed. Thus, the same disconfirmer may be subtyped or not depending on the context in which the group stereotype is learned.

In essence, we are arguing that because a comparative context leads to more extreme and homogeneous group perceptions along context-relevant dimensions, someone who disconfirms the group stereotype along those dimensions will be judged more atypical of the group. Consequently, less stereotype change should ensue. In this case, it is not the content of the disconfirming information that affects its judged typicality; rather it is the context that has altered the group stereotype along the relevant dimensions.

### The current research

We report two studies designed to test the hypothesis outlined in the previous paragraphs, i.e., that the same disconfirming information will be more or less influential in producing stereotype change, depending on the comparative context in which the original group stereotype has been learned. In these studies, participants learned about a focal group by encountering stereotypic behaviors of group members along two different attribute dimensions (e.g., academically motivated and politically liberal). They formed an impression of this focal group under one of the two comparative context conditions. In both conditions, participants first learned about a context group, by also encountering stereotypic behavioral information along two different attribute dimensions. One of these context groups was described as academically unmotivated and artistic. The other context group was described as politically conservative and campus leaders. Note that each group provides a comparative context along one of the two dimensions that describes the focal group but not the other.

Participants later encountered additional members of the focal group who disconfirmed the group stereotype on *both* the academic and politically liberal dimensions. We predicted that this disconfirming information would lead to a greater stereotype change on whichever attribute dimension was *not relevant* to the comparative context. Thus, the disconfirming information should

induce more stereotype change on the politically liberal dimension when the context group was academically unmotivated than when it was politically conservative. Similarly, a greater stereotype change on the academically motivated dimension should occur when the context group was politically conservative than when it was academically unmotivated.

In summary, we predict that participants will form more extreme stereotypes about the focal group on the dimension that distinguishes it from the context group than on the dimension that does not. Because of its greater extremity, we expect that participants will be less likely to change their stereotypes of the focal group on the context-relevant dimension than on the context-irrelevant dimension.

## Experiment 1

### Method

#### Participants

Eighty students at the University of Colorado participated in this experiment to fulfill a requirement of their introductory psychology course. Students participated individually or in small groups.

#### Design

The experiment used a 2 (context group: academic vs. political)  $\times$  2 (confirmation: confirm vs. disconfirm)  $\times$  2 (trait dimension rated: academic vs. political) design with the first two factors between- and the last within-participants. Participants formed a stereotype of a focal group of college students who were academically motivated and politically liberal. To provide a social context in which to learn about the focal group, participants first learned the stereotype of one of the two other college student groups. Each context group was described on one dimension that was shared with the focal group but of the opposite pole (e.g., academically unmotivated) and on another dimension that was irrelevant to the focal group stereotype. In the academic context condition, members of the context group were described as academically unmotivated and artistic. In the political context condition, members of the context group were described as politically conservative and leaders on campus. In a subsequent phase of the experiment, participants were exposed to additional focal group members who either confirmed or disconfirmed the stereotype that focal group members were academically motivated and liberal.

#### Procedure

The experiment involved two phases: a stereotype formation phase and a stereotype change phase. Upon arrival at the laboratory, a female experimenter in-

structed participants that they would learn about two fraternities with the goal of forming an impression of what each fraternity was like.

In the *stereotype formation phase* of the experiment, participants were first asked to form an impression of one of the two context fraternities by reading statements that described one behavior for each of the 24 fraternity members. Participants in the academic context condition read 12 statements that described behaviors that were academically unmotivated (e.g., a group member went to a party the night before a test even though he had not studied for it) and 12 statements that described behaviors that reflected interest in the arts (e.g., a group member had season tickets for the local symphony orchestra). Participants in the political context condition read 12 statements that described politically conservative behaviors (e.g., a group member thought that government efforts to preserve national parks were a waste of time) and 12 statements that described leadership behaviors (e.g., a group member made sure students on campus voted in the student union election). The 24 statements were presented on a computer screen in random order and participants were able to read through the statements at their own pace.

After learning about one of the two context fraternities, participants were introduced to the second (focal) fraternity. Participants formed an impression of the focal fraternity by reading 24 statements that described behaviors performed by fraternity members. Twelve of these statements described academically motivated behaviors (e.g., a group member scored 100% on a surprise quiz in his calculus class) and 12 described politically liberal behaviors (e.g., a group member burned an American flag during a protest against military spending). Again, the 24 statements about the focal group were presented in a random order. After learning about both fraternities, participants were allowed as much time as needed to generate descriptions of both groups (most participants wrote 3–4 sentences about each fraternity). The descriptions were subsequently coded to assess the impact of context on focal group stereotypes, prior to presentation of the confirming or disconfirming instances.

In the *stereotype change phase* of the study, participants were told that they would learn about additional members of the focal fraternity. Participants in the confirm condition then read about 12 additional fraternity members whose behaviors were consistent with the stereotype they had formed in the first phase of the experiment. Specifically, they read six statements describing academically motivated behaviors and six statements describing liberal behaviors. Participants in the disconfirm condition also read about 12 additional fraternity members. However, these participants learned about several group members who disconfirmed the group stereotype. Specifically, they read four statements

that described academically unmotivated behaviors and four statements that described politically conservative behaviors.<sup>1</sup> The remaining four statements described stereotype-consistent (two academically motivated and two politically liberal) behaviors.

After learning about the additional members of the focal group, participants completed measures of stereotype central tendency and variability of the focal fraternity as well as measures of perceived typicality of the 12 group members learned in the second phase of the experiment.

#### *Dependent variables*

*Initial open-ended descriptions.* Prior to receiving the additional fraternity members during the stereotype change phase of the study, participants wrote a few sentences that gave their impressions of the focal fraternity. These impressions were rated by judges to assess the perceived extremity of participants' group stereotype based on the initial stereotype formation information.

After the stereotype change phase of the study, participants completed a set of additional measures to assess their stereotypes of the focal group at that point.

*Percentage estimate task.* Participants completed a percentage estimate task in which they were asked to estimate the prevalence of eight attributes among members of the focal fraternity. Two stereotypic and two counterstereotypic items for each dimension of the focal group stereotype were assessed: academically motivated/stereotypic (being good at school, being studious), academically motivated/counterstereotypic (being unmotivated to do well at school, being poor students), politically liberal/stereotypic (supporting liberal causes, being likely to vote for Democrats), and politically liberal/counterstereotypic (being politically conservative, being likely to vote for Republicans).

*Mean/range task.* The perceived mean and range of the focal group on each attribute used in the percentage estimate task were also assessed. Participants estimated the standing of the average member of the focal fraternity on each attribute by placing an 'X' on a 178-mm

<sup>1</sup> We were initially concerned with participants' ability to differentiate between the two groups, particularly, when it came to the disconfirming members of the focal group (some of whom were clearly more similar to members of the context group). To determine whether participants would be able to correctly remember the group to which each individual belonged, we conducted an initial pilot study. In that study, 12 participants completed the stereotype formation phase of Experiment 1 and were then exposed to the disconfirming members of the focal group. After a brief delay, participants completed a recognition task in which they were again presented with the disconfirming group members, as well as confirming members of the focal group and members of the context group. Their task was to identify the group membership of each individual. Participants were highly accurate ( $M = 94\%$ ) for all types of statements.

line that represented the possible range ('not at all' to 'extremely') of that attribute. Participants also put marks on the dimension to indicate where the group members who possessed the least and the greatest amounts of the attribute were located. The perceived range of the group was computed by taking the difference in these two ratings.

*Within-group similarity ratings.* Participants rated how similar members of the focal group were to each other on the academic and political dimensions. They marked their ratings on seven-point scales that ranged from extremely dissimilar to extremely similar.

*Typicality ratings.* The set of behaviors that were included in the stereotype change phase of the experiment was again presented to participants. Participants rated how typical each group member was of the overall group using a seven-point scale (extremely atypical to extremely typical).

## Results<sup>2</sup>

### *Stereotype formation: perceived stereotypicality*

After first learning about the two fraternities and before the stereotype change phase of the experiment, participants wrote open-ended descriptions of each group. Judges' rated these descriptions to assess the effect of context on the group stereotype. Based on prior research, we expected ratings of the focal group to be more stereotypic for the dimension that was context-relevant than for the dimension that was context-irrelevant. The focal fraternity descriptions were rated using five-point scales to assess the extent to which the focal group was depicted as academically motivated (1 = very unstudious, 3 = academic performance not mentioned, 5 = very studious) and politically liberal (1 = very conservative, 3 = political orientation not mentioned, 5 = very liberal). Two judges rated a subset of the descriptions (from 22.5% of the participants). Inter-rater reliability was satisfactory (range for focal and context group ratings of .68–.80) thus, the remaining descriptions were rated by a single judge.

Judges' ratings were subjected to a two-way analysis of variance (ANOVA) with context condition as a between-participants variable and dimension being rated as a repeated measure. As hypothesized, participants' open-ended descriptions of the groups indicated that

they formed more extreme stereotypes about the focal group on a particular dimension if that dimension differentiated it from the context group than if it did not, as indicated by a significant interaction between context and dimension,  $F(1, 72) = 5.69$ ,  $p = .020$ . The group was perceived as more liberal in the political context ( $M = 3.62$ ) than in the academic context ( $M = 3.26$ ),  $F(1, 72) = 4.07$ ,  $p = .047$ . In contrast, the group was perceived only slightly more academically motivated in the academic context ( $M = 4.09$ ) than in the political context ( $M = 4.00$ ),  $F(1, 72) = 1.63$ ,  $p = .206$ .

### *Stereotype change: measures of stereotypicality*

After reading about the second set of focal group members, participants completed a more extensive set of dependent measures, including measures of the group's central tendency and variability on both context-relevant and context-irrelevant dimensions. Among participants in the disconfirm condition, less stereotype change was expected on the relevant dimension than on the irrelevant dimension. This would be evidenced by a significant three-way interaction between context group, trait dimension, and confirmation condition.

*Percentage estimate task.* Two stereotypicality indices for the focal group stereotype were computed, one for ratings of academic motivation and one for ratings of liberalness. To form stereotypicality indices, the two items on each dimension that were stereotype-inconsistent were averaged and subtracted from the average of the two items that were stereotype-consistent on that dimension. Positive scores indicate more stereotypic than counterstereotypic perceptions of the group while negative scores indicate more counterstereotypic than stereotypic perceptions of the group.

The stereotypicality indices were subjected to the three-way ANOVA (context  $\times$  confirmation  $\times$  trait dimension rated) outlined above. The means for this analysis are given in the top panel of Table 1.<sup>3</sup> Not surprisingly, there was a significant main effect of confirmation condition, such that participants viewed the group more stereotypically in the confirm condition than in the disconfirm condition,  $F(1, 71) = 40.70$ ,  $p < .001$ . Furthermore, there was a significant interaction between dimension and confirmation condition, such that the difference between the confirm and disconfirm conditions was stronger for political ratings than academic ratings,  $F(1, 71) = 5.02$ ,  $p = .028$ .

The critical test of our hypothesis was the predicted three-way interaction, which approached significance,  $F(1, 71) = 3.11$ ,  $p = .082$ . To test for the effect of con-

<sup>2</sup> Five outliers were excluded from the analyses. Outliers were defined, using box and whisker plots, as values lying 1.5 times the interquartile range outside of the range itself in each condition. Four of the outliers were in the academic context, disconfirm condition. The other outlier was in the political context, confirm condition. Analyses conducted with the outliers retained yielded an identical pattern of results.

<sup>3</sup> Throughout all of the analyses reported in this experiment, there was an uninteresting main effect of trait dimension, such that participants expressed stronger stereotypes on the academic dimension than on the political dimension. Because this main effect is of no theoretical interest, it will not be discussed further.

text on the stereotype change, follow-up tests were conducted within each confirmation condition. As hypothesized, there was a significant interaction between context and dimension among participants in the stereotype-disconfirming condition. These participants saw the focal group less stereotypically on the dimension that was irrelevant to the social context than on the dimension that was relevant to the social context,  $F(1, 34) = 6.16$ ,  $p = .018$ . This interaction did not emerge among participants in the stereotype-confirming condition,  $F(1, 37) < 1$ .

*Means from the mean/range task.* A similar pattern of results emerged for the mean ratings of the focal group on the mean/range task, given in the lower panel of Table 1. Stereotypicality indices were formed for mean ratings on each dimension and subjected to a context  $\times$  confirmation  $\times$  trait dimension rated mixed-model ANOVA. A significant main effect of confirmation condition was found,  $F(1, 71) = 33.13$ ,  $p < .001$ . The focal group was again perceived more stereotypically in the confirm condition than in the disconfirm condition. More pertinent to the hypothesis, the three-way interaction between dimension, context, and confirmation condition approached significance,  $F(1, 71) = 3.39$ ,  $p = .070$ . Follow-up tests revealed a significant dimension by context interaction within the stereotype-disconfirming condition,  $F(1, 34) = 4.10$ ,  $p = .051$  but not the stereotype-confirming condition,  $F(1, 37) < 1$ . Consistent with predictions, the focal group was seen less stereotypically on context-irrelevant than context-relevant dimensions by participants in the disconfirm condition. Therefore results from the percentage estimate and mean ratings tasks provide converging evidence that the central tendency component of stereotypic perceptions is more amenable to change in dimensions that are irrelevant to a social context, compared to dimensions that are relevant to a social context.

#### *Stereotype change: measures of within-group variability*

Disconfirming information should affect not only the perceived stereotypicality or extremity of a group stereotype but also the perceived within-group variability or dispersion. If disconfirming information is taken into consideration and the stereotype changed, then the perceived dispersion of group members about their central tendency is expected to increase to accommodate differences among group members. We expected that the effect of context on perceived variability would mirror the effects found for perceived stereotypicality: disconfirming information would be associated with more dispersed perceptions of the group on context-irrelevant dimensions compared to context-relevant dimensions.

*Group range from the mean/range task.* For each attribute, we computed the perceived range by subtracting the minimum scale value marked from the maximum scale value marked. Since we did not expect context to differentially impact perceptions of variability on stereotypic and counterstereotypic dimensions, we formed a composite range value by averaging the ranges across all attributes relevant to each trait dimension. Results from the three-way mixed model ANOVA (means given in the top panel of Table 2) revealed only a significant main effect of confirmation condition, such that participants perceived a greater range among group members when they had received stereotype-disconfirming information in the second phase of the experiment than when they had learned additional stereotype-confirming information only,  $F(1, 69) = 36.99$ ,  $p < .001$ .

While the overall ANOVA yielded no significant interactions, analyses conducted within each confirmation condition partially supported our predictions. Within the disconfirm condition, participants viewed the group as more variable on context-irrelevant dimensions than context-relevant dimensions,  $F(1, 33) = 4.60$ ,  $p = .039$ . This was not true in the condition where they only received confirming information,  $F(1, 37) < 1$ .

Table 1  
Experiment 1: Perceived stereotypicality means (standard deviations)

| Confirmation condition            | Context condition | Ratings on academic dimension | Ratings on political dimension |
|-----------------------------------|-------------------|-------------------------------|--------------------------------|
| <i>Percentage estimation task</i> |                   |                               |                                |
| Disconfirm                        | Academic          | 51.97 (25.64)                 | -3.31 (19.80)                  |
|                                   | Political         | 41.15 (27.60)                 | 7.75 (29.55)                   |
| Confirm                           | Academic          | 74.80 (24.96)                 | 52.75 (38.41)                  |
|                                   | Political         | 72.97 (19.45)                 | 43.53 (49.10)                  |
| <i>Mean rating task</i>           |                   |                               |                                |
| Disconfirm                        | Academic          | 58.47 (39.43)                 | -6.47 (29.97)                  |
|                                   | Political         | 44.48 (33.04)                 | 12.73 (55.78)                  |
| Confirm                           | Academic          | 95.13 (42.68)                 | 75.23 (61.69)                  |
|                                   | Political         | 91.58 (38.23)                 | 57.42 (66.50)                  |

Note. Means are difference scores on which higher values are more stereotypic.

Table 2  
Experiment 1: Perceived within-group variability means (standard deviations)

| Confirmation condition              | Context condition | Ratings on academic dimension | Ratings on political dimension |
|-------------------------------------|-------------------|-------------------------------|--------------------------------|
| <i>Range task</i>                   |                   |                               |                                |
| Disconfirm                          | Academic          | 117.00 (38.13)                | 125.63 (29.21)                 |
|                                     | Political         | 129.83 (23.78)                | 115.98 (30.44)                 |
| Confirm                             | Academic          | 82.93 (40.16)                 | 65.53 (33.18)                  |
|                                     | Political         | 89.16 (49.32)                 | 72.04 (40.40)                  |
| <i>Within-group similarity task</i> |                   |                               |                                |
| Disconfirm                          | Academic          | 4.94 (0.77)                   | 4.13 (1.35)                    |
|                                     | Political         | 4.60 (1.09)                   | 4.85 (0.93)                    |
| Confirm                             | Academic          | 6.00 (1.12)                   | 5.60 (0.67)                    |
|                                     | Political         | 6.00 (1.31)                   | 5.21 (1.27)                    |

Note. In the top panel, higher values indicate a greater perceived range (i.e., more variability). In the lower panel, higher values indicate a greater perceived similarity (i.e., less variability).

*Within-group similarity.* Finally, ratings of within-group similarity on each dimension were subjected to a three-way ANOVA, with repeated measures on the dimension being rated (means are given in the lower panel of Table 2). There was an uninteresting significant main effect of dimension,  $F(1, 71) = 7.23$ ,  $p = .009$ , such that the group was perceived as more homogeneous on the academic dimension than on the political dimension. Not surprisingly, there was also a significant main effect of confirmation condition,  $F(1, 71) = 29.91$ ,  $p < .001$ , such that participants in the disconfirm condition perceived less within-group similarity among focal group members than did participants in the confirm condition.

Most importantly, there was a significant three-way interaction between context, confirmation condition, and dimension,  $F(1, 71) = 4.97$ ,  $p = .029$ . To further explore this interaction, we calculated simple two-way interactions within each confirmation condition. Consistent with predictions, participants in the stereotype-disconfirming condition perceived significantly less similarity among group members on context-irrelevant dimensions than on context-relevant dimensions,  $F(1, 34) = 6.07$ ,  $p = .019$ . The context by dimension interaction did not approach significance in the stereotype-confirming condition,  $F(1, 37) < 1$ .

#### *Stereotype change: perceptions of disconfirming group members*

*Typicality.* We have presumed that the differential effect of context on stereotype change for context-relevant and

context-irrelevant dimensions occurred because the perceived extremity shift on context-relevant dimensions made it easier to subtype disconfirming group members on those dimensions. We assessed perceptions of typicality of the disconfirming members to investigate this subtyping interpretation. If differences in perceived stereotypicality on relevant versus irrelevant dimensions are due to differences in subtyping, then we would expect disconfirming group members to be seen as particularly atypical of the group on context-relevant dimensions.

Typicality indices were computed by subtracting the average typicality ratings of stereotype-disconfirming group members from the average typicality ratings of stereotype-confirming group members (thus controlling for differences in how typical individual group members were perceived). Since disconfirmers were presented only in the disconfirm condition, these typicality indices were computed only in that condition.<sup>4</sup> These typicality differences were submitted to a two-way mixed model ANOVA with context and dimension as between- and within-participant factors, respectively (means given in Table 3). A significant two-way interaction between context and dimension was found,  $F(1, 34) = 10.95$ ,  $p < .001$ . Participants saw a greater difference in the typicality of confirming and disconfirming group members when those group members were relevant to the context dimension. Specifically, relative to confirming group members, conservative group members were seen as more atypical of the liberal group to which they belonged when the social context was political. Likewise,

Table 3  
Experiment 1: Differences in judged typicality of confirming and disconfirming group members

| Context   | Academic dimension | Political dimension |
|-----------|--------------------|---------------------|
| Academic  | 3.06 (1.17)        | -0.13 (1.08)        |
| Political | 2.66 (1.87)        | 1.14 (1.78)         |

Note. More positive numbers indicate that the confirming members were seen as more typical than the disconfirming members.

<sup>4</sup> Although we did not make specific predictions regarding the perceptions of the confirming group members presented in the stereotype change phase of the experiment, we did collect typicality data in the confirm condition. In analyzing these data, we found no significant results either in the confirm conditions or in the perceived typicality of the few confirming group members presented in the disconfirm condition.

academically unmotivated group members were seen as more atypical of the academically motivated group to which they belonged when the social context was academic.

### *Discussion*

Several interesting results emerged from this experiment. First, we hypothesized that the context in which a group stereotype was formed would affect the extremity of that stereotype, similar to contrast effects reported in past research. Based on the written descriptions completed prior to receiving the additional (either confirming or disconfirming) information, our results supported this prediction. Participants who learned about a focal fraternity (which was described as both liberal and academically motivated) in the context of a conservative fraternity came to view it as more liberal than did those for whom the context was an academically unmotivated fraternity.

Second, the context in which participants formed the focal group stereotype also influenced the malleability of the stereotype on different dimensions. On both of the stereotypicality measures collected following exposure to the additional group members, participants who received disconfirming information were more likely to change their stereotypes about the focal group on the context-irrelevant dimension than on the context-relevant dimension. Thus, it appears that when people form stereotypes that help them to distinguish two groups from each other along some dimension, they are less sensitive to disconfirming information along that dimension.

Participants' perceptions of within-group variability were also consistent with this conclusion. Namely, participants who received disconfirming information saw the focal group as more variable than did those who only received confirming information, particularly on the context-irrelevant dimension. In other words, the presence of disconfirmers may have increased the perceived dispersion of the group, consistent with stereotype change, but they did so more strongly on the dimension that failed to differentiate the focal group from its context than on the differentiating dimension.

Finally, participants' typicality ratings also suggest that individuals who disconfirmed the group stereotype on the context-relevant dimension were perceived as more atypical than those who disconfirmed it on the context-irrelevant dimension. This provides evidence for our subtyping explanation. Disconfirming group members are more likely to be subtyped when they disconfirm on a dimension that differentiates a group from its context than when they disconfirm on a context-irrelevant dimension.

It is important to note that the presence of disconfirming information led to less extreme stereotypes of

the group both on context-relevant and context-irrelevant dimensions. Thus, it was not the case that the disconfirming information had no impact when it was context-relevant. Rather, it simply had less impact than what the same disconfirming information had when it was context-irrelevant. This suggests that subtyping is not an all-or-none process. Rather, there is a continuum here: to the extent that disconfirming information is judged to be atypical of the group, the impact of that information on subsequent views of the group will be reduced. Our results here are thus consistent with what Johnston and Hewstone (1992) refer to as the "weak version" of subtyping, in which disconfirmers have an impact on stereotype change, but this impact is curtailed to the extent that they are subtyped.

### **Experiment 2**

Although the results of Experiment 1 were generally consistent with our predictions, there were a number of weaknesses in the data that we felt should be addressed in a subsequent experiment. First, and most obviously, a number of predicted three-way interactions reached only marginal significance levels. The somewhat weak effects may have been due in part to the lack of success in producing a context effect on the academic dimension. Thus, we felt that a replication of the experiment was appropriate to examine, if the expected three-way interactions would be significant when an initial context effect was found for both trait dimensions.

Second, there were consistent main effects of trait dimension in all of our measures, such that the focal group was rated more stereotypically on the academic dimension than on the liberal dimension. We speculated that this main effect might have been derived from the fact that we used hypothetical fraternities (organizations about which many students have pre-existing stereotypes) as our groups. If participants held the stereotype that fraternities in general are not particularly studious, our focal fraternity may have been seen as relatively extreme (for a fraternity) on that trait dimension. This may also be partially responsible for the absence of an initial context effect on this trait dimension: the group was already seen to be quite high on this dimension and, as a result, little effect of context could be detected.

To deal with these issues, we decided to present participants in Experiment 2 with information about two 'student groups.' We felt that the term 'student groups' was sufficiently vague in that participants would not try to apply any previous stereotypic knowledge to these groups.

The third issue that we wanted to address in Experiment 2 concerned the fact that our measurement of the focal group stereotype prior to the stereotype change phase of the study was based only upon short written



descriptions provided by participants. We intentionally chose not to give them the full range of stereotype measures at this point, for fear that their ratings would become anchored, prior to learning the additional confirming or disconfirming group members. But the problem with this approach is that it provided us with only weak baseline measures of the stereotype. An alternative, which we implemented in the second study, is to include a control condition in which participants completed the full set of dependent variables without receiving additional information after the initial stereotype formation stage of the study.

Including this additional control condition would also help us address an additional theoretical point raised by the result reported in Experiment 1. Our argument is that disconfirming information is differentially effective in producing stereotype change depending on whether or not it is context-relevant. But, in the absence of a control condition, it is unclear whether the disconfirming information produced any stereotype change at all. Perhaps, it is the presence of additional confirming information that causes previous group stereotypes to become more extreme. The addition of the aforementioned control condition should help eliminate this alternative explanation for the results from Experiment 1.

### Method

#### Participants

One hundred and sixty students at the University of Colorado participated in this experiment in exchange for partial credit towards a course requirement. Participants were run individually or in small groups.

#### Design

The experiment employed a 2 (context: academic vs. political)  $\times$  3 (confirmation: confirming vs. disconfirming vs. no additional information)  $\times$  2 (trait dimension rated: academic vs. political) mixed design, with the first two factors varying between-participants and the last varying within them. As in Experiment 1, participants learned about the focal group in either an academic or political context. In the stereotype change phase of the experiment, participants were either given no additional information about the focal group (no information condition) or were given stereotype-confirming or stereotype-disconfirming information. Participants then rated the focal group on dimensions of academic achievement and political orientation.

#### Procedure

The procedure of this experiment was similar to that used in Experiment 1, with three exceptions. Most significantly, this experiment included a no additional information condition in which participants did not

complete the stereotype change phase of the experiment and instead completed the dependent measures (except for the typicality measure), immediately after the stereotype formation phase of the experiment.

Two other methodological changes were made. First, the groups were identified as ‘student groups on a college campus’ rather than as fraternities to reduce the application of pre-existing stereotypes to the groups. Second, the mean/range task described in the second set of dependent measures used in Experiment 1 was replaced with a histogram task. In this task, participants were asked to generate a perceived frequency distribution of group members along each of the eight attributes used in the first experiment. Participants were presented with five blank bars of equal height that were labeled to reflect a continuum that ranged from *very few group members* to *very many group members*. Participants were asked to fill in each bar to the height that would indicate the proportion of group members they believed belonged in that segment of the continuum.

### Results<sup>5</sup>

#### *Stereotype formation: measures of stereotypicality*

*Open-ended descriptions.* As in Experiment 1, all participants wrote brief descriptions of the groups, prior to the stereotype change phase of the study (or, in the case of the no-information participants, prior to completing the other dependent variables). As in Experiment 1, these open-ended descriptions were rated on five-point scales to assess the degree to which the focal group was portrayed as academically motivated and politically liberal. Two judges rated a subset of the descriptions (18% of participants). Inter-rater reliability was again satisfactory (range for context and focal group ratings of .67–.94), so one judge rated the remaining descriptions.

Ratings of the focal group were entered in a two-way ANOVA with context as a between-participant factor and dimension rated as a within-participant factor. A significant context  $\times$  dimension rated interaction emerged,  $F(1, 158) = 18.00, p < .001$ . As in Experiment 1, the focal group was perceived as more liberal when learned in the political ( $M = 3.78$ ) than in the academic context ( $M = 3.33$ ),  $F(1, 158) = 16.10, p < .001$ . In contrast to Experiment 1, the context effect was also obtained for the academic dimension,  $F(1, 158) = 5.08, p = .026$ , the focal group was perceived as more academically motivated when learned in the academic

<sup>5</sup> Six outliers were excluded from the analysis. Outliers were defined as in Experiment 1. Four of the outliers were in the academic context, disconfirm condition. One outlier was in the political context, confirm condition. The final outlier was in the political context, disconfirm condition. Analyses conducted with the outliers included yielded the same pattern of results as those reported.

( $M = 4.38$ ) than in the political context ( $M = 4.07$ ). Thus, participants formed more extreme stereotypes of the focal group on dimensions that were context-relevant than on dimensions that were context-irrelevant.

#### *Stereotype change: measures of stereotypicality*

We again collected measures of central tendency and variability. As before, we expected less stereotype change on context-relevant than on context-irrelevant dimensions. This result should be seen in significant three-way interactions between context group, trait dimension, and confirmation condition.

We examined this hypothesis by separate three-way mixed-model ANOVAs involving context (academic vs. political), confirmation (no information vs. confirm vs. disconfirm), and trait dimension rated (academic vs. political), treating the first two factors as between-participants and the latter within. Because our primary hypothesis concerned differences between the disconfirm condition and the other two conditions, we tested planned contrasts for the effect of confirmation condition in each analysis. The first contrast compared ratings in the disconfirm condition to ratings in the confirm and no-information conditions. The second orthogonal contrast compared the no-information and confirm conditions. Our hypotheses would be supported by significant interactions involving the first contrast, but not the second.

*Percentage estimate.* Stereotypicality scores were computed for each dimension as in Experiment 1. The three-way ANOVA revealed a significant interaction between context and dimension rated,  $F(1, 148) = 19.37$ ,  $p < .001$ . Overall, ratings of liberalness tended to be higher in the political context condition, whereas ratings of studiousness tended to be higher in the academic context condition. This is consistent with the context effects found during stereotype formation. Not surprisingly, mean stereotypicality scores differed by condition. Participants in the disconfirm condition made less stereotypic judgments than those in the other two conditions ( $F(1, 150) = 19.19$ ,  $p < .001$ ) but the confirm and no-information conditions did not differ significantly,  $F(1, 101) = 2.42$ ,  $p = .123$ .

The predicted three-way interaction between dimension, context, and the first planned contrast for confirmation was significant,  $F(1, 148) = 6.03$ ,  $p = .015$  (means presented in the top panel of Table 4). As expected, participants who learned about stereotype-disconfirming group members were more likely than participants in the other two conditions to alter their stereotypes of the group, particularly on dimensions that were not relevant to the context. Fig. 1 presents the means for this triple interaction, averaging across the confirm and no-information conditions and collapsing across the specific dimensions. Thus, comparisons can be made between the context-relevant and context-irrele-

Table 4  
Experiment 2: Perceived stereotypicality means (standard deviations)

| Confirmation condition            | Context condition | Ratings on academic dimension | Ratings on political dimension |
|-----------------------------------|-------------------|-------------------------------|--------------------------------|
| <i>Percentage estimation task</i> |                   |                               |                                |
| Disconfirm                        | Academic          | 61.35 (22.33)                 | 6.26 (18.06)                   |
|                                   | Political         | 50.21 (28.37)                 | 50.46 (32.42)                  |
| No-information                    | Academic          | 71.42 (43.04)                 | 29.62 (43.47)                  |
|                                   | Political         | 69.27 (22.82)                 | 55.88 (36.75)                  |
| Confirm                           | Academic          | 73.07 (28.85)                 | 41.41 (42.81)                  |
|                                   | Political         | 83.15 (14.37)                 | 60.22 (37.12)                  |
| <i>Mean rating task</i>           |                   |                               |                                |
| Disconfirm                        | Academic          | 1.59 (0.72)                   | 0.19 (0.75)                    |
|                                   | Political         | 1.26 (0.86)                   | 1.44 (1.00)                    |
| No-information                    | Academic          | 2.16 (1.45)                   | 1.19 (1.40)                    |
|                                   | Political         | 2.32 (0.96)                   | 1.93 (1.45)                    |
| Confirm                           | Academic          | 2.28 (0.99)                   | 1.55 (1.65)                    |
|                                   | Political         | 2.44 (0.87)                   | 2.19 (1.25)                    |

*Notes.* Means are difference scores on which higher values are more stereotypic.

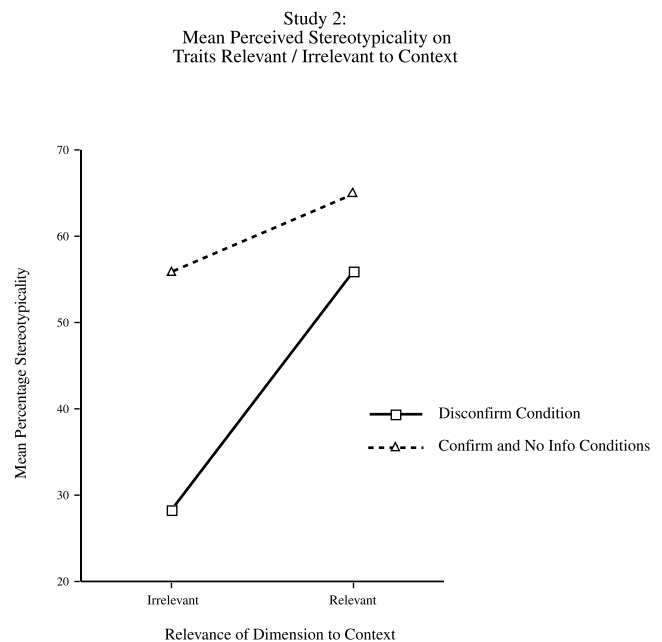


Fig. 1. Study 2: Mean perceived stereotypicality on traits relevant/irrelevant to context.

vant dimensions (regardless of which one was the academic dimension and which one was the political dimension). As this figure clearly shows, the decrease in perceived stereotypicality as a function of receiving disconfirming information was substantially larger on the context-irrelevant dimension than on the relevant one.

The three-way interaction involving the second confirmation contrast was not significant,  $F(1, 101) = 1.15$ ,

$p = .286$ , thus indicating that the context  $\times$  dimension rated interaction did not differ in the confirm and no-information conditions.

*Histogram mean.* Stereotypicality indices were also computed based on the histogram task. The mean of each histogram was computed and the average of counterstereotypic items was subtracted from the average of stereotypic items for each dimension. The three-way ANOVA again revealed a significant interaction between context and dimension,  $F(1, 147) = 15.55$ ,  $p < .001$ , generally replicating the pattern found in the percentage estimate data. On average, the focal group stereotype was more extreme on context-relevant than on context-irrelevant dimensions. As was the case in the percentage estimate ratings, the mean group stereotype differed as a function of the confirmation condition. Participants in the disconfirm condition generated less stereotypic judgments than did those in the other two conditions,  $F(1, 149) = 30.38$ ,  $p < .001$ , which did not differ from each other,  $F(1, 100) = 1.10$ ,  $p = .298$ .

The most important result is provided by the three-way interaction between trait dimension, context, and the first planned contrast for confirmation condition,  $F(1, 150) = 19.19$ ,  $p < .001$ . The means, presented in the lower half of Table 4, nicely replicate the results from the percentage estimate task. Namely, participants were less likely to adjust their stereotypes in response to disconfirming information when that information was related to a dimension that differentiated the focal group from the context group. The triple interaction that focused on the other contrast (comparing the no-information and confirm conditions only) was not significant.

#### *Stereotype change: measures of within-group variability*

Perceptions of within-group variability were examined with two measures: the variance computed from the subjective histograms and the within-group similarity ratings. We expected disconfirming information to increase perceptions of group variability more on context-irrelevant dimensions than on context-relevant dimensions.

*Histogram variance.* The variance of each subjective histogram was computed and averaged, separately for the academic and political dimensions. The three-way ANOVA yielded a significant interaction between context and dimension rated that partially supported predictions,  $F(1, 148) = 9.20$ ,  $p = .003$ . While the focal group was perceived as less variable on the academic dimension than on the political dimension, this was especially true in the academic context. Perceptions of group variability also differed systematically by confirmation condition. Results of the planned contrasts revealed that participants in the disconfirm condition perceived the group as more variable than did those in the other two conditions,  $F(1, 146) = 65.59$ ,  $p < .001$ ,

which did not significantly differ from each other,  $F(1, 148) = 1.27$ ,  $p = .262$ .

The hypothesized three-way interaction that focused on the planned contrast between the disconfirm condition and the other two was marginally significant,  $F(1, 148) = 2.82$ ,  $p = .095$ . The pattern of results is consistent with expectations. Namely, the difference in perceived variability between the disconfirm condition and the other two was larger on the context-irrelevant dimension than on the context-relevant one. In other words, increases in perceived within-group dispersion as a function of encountering disconfirming group members were especially large on the context-irrelevant dimension.

*Within-group similarity.* The three-way ANOVA on perceived within-group similarity revealed a significant two-way interaction between context and trait dimension,  $F(1, 147) = 11.83$ ,  $p < .001$ . The pattern of results suggests that participants viewed the group as more homogeneous (less variable) on dimensions that distinguished it from the focal group. With regard to confirmation condition, planned contrasts indicated that participants in the disconfirm condition perceived the group as generally less homogeneous (more variable) than did those in the other two conditions,  $F(1, 149) = 78.07$ ,  $p < .001$ . The no-information and confirm conditions did not differ,  $F(1, 100) = 2.25$ ,  $p = .137$ .

In contrast to the results for variance ratings, the three-way interaction using the first planned contrast was significant for within-group similarity ratings,  $F(1, 149) = 8.57$ ,  $p = .004$ . The pattern of results mirrored that from the histogram variance measure. Participants saw group members as less similar to each other if they had encountered stereotype-disconfirming information, but this was much less likely to be the case of a dimension that differentiated the focal and context groups than of a context-irrelevant dimension. Fig. 2 presents a graph of this result, again collapsing across the confirm and no-information condition, and across whether the specific dimension was an academic or a political one. The other component of the triple interaction, comparing the confirm and no-information conditions, was not significant,  $F(1, 100) < 1.0$ .

#### *Stereotype change: perceptions of disconfirming group members*

*Typicality ratings.* We have suggested that the difference in the impact of disconfirming information on context-relevant and context-irrelevant dimensions is due to subtyping disconfirming exemplars on context-relevant dimensions. In particular, we expected disconfirmers on context-relevant dimensions to be perceived as more atypical of the group than disconfirmers on context-irrelevant dimensions. Typicality ratings from the disconfirm condition were analyzed to assess how

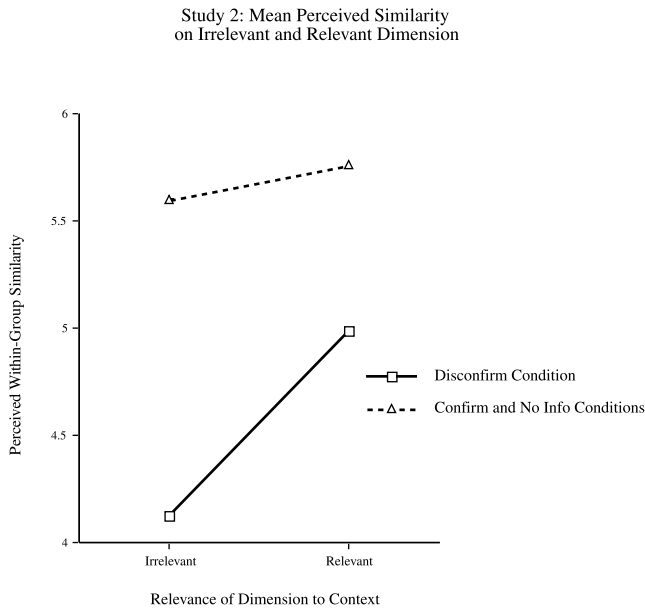


Fig. 2. Study 2: Mean perceived similarity on irrelevant and relevant dimension.

stereotype-disconfirming group members were perceived relative to the rest of the group. Typicality indices were computed by subtracting the average typicality of stereotype-disconfirming group members from the average typicality of stereotype-confirming group members (thus, controlling for differences in how typical individual group members were perceived to be on average).

A two-way ANOVA revealed a significant two-way interaction between context and dimension,  $F(1, 47) = 20.86$ ,  $p < .001$ . As expected, participants perceived a bigger discrepancy between the academically motivated and academically unmotivated group members in the academic context ( $M_{\text{difference}} = 3.90$ ) than in the political context ( $M_{\text{difference}} = 3.11$ ). Similarly, participants perceived a bigger discrepancy between conservative and liberal group members in the political context ( $M_{\text{difference}} = 1.92$ ) than in the academic context ( $M_{\text{difference}} = 0.29$ ).

#### Mediational analyses

To examine whether typicality differences were responsible for the stereotype change effects we observed, a mediational analysis was undertaken, following the procedures laid out by Judd and Kenny (1981) and Baron and Kenny (1986). We have already shown that stereotypes on context-irrelevant dimensions changed more than stereotypes on context-relevant dimensions in the disconfirming condition. Additionally, we have shown that disconfirmers on context-irrelevant dimensions were perceived as more typical of the group than disconfirmers on context-relevant dimensions. To demonstrate that the typicality difference mediated stereotype change, we examined whether the predicted effects

on stereotype change were reduced in magnitude, once we controlled for the perceived atypicality of the disconfirming instances. More specifically, our analyses examined whether the difference in stereotypicality between the context-relevant and context-irrelevant dimensions was reduced, once differences in the disconfirmer atypicality ratings (confirmers minus disconfirmers) were controlled. Since dimensions varied within-participants, procedures for assessing mediation of within-subject effects were used (Judd, Kenny, & McClelland, 2001). Using the Sobel test (Kenny et al., 1998), there was a significant reduction in the effect of context once typicality was controlled, in the case of both percentage estimates ( $Z = 3.37$ ) and means from the histogram task ( $Z = 3.01$ ).<sup>6</sup> These analyses thus suggest that the effect of context on stereotype change was partially mediated by impressions of disconfirmer atypicality.

#### Discussion

The results of Experiment 2 replicated those of our first experiment. We again observed contrast effects of the social context in which a stereotype was learned on the initial stereotypes formed about the focal group. Importantly, the contrast effect was found for both trait dimensions in the second experiment. The focal group was perceived as more academically motivated and less liberal when it was presented in the context of an academically unmotivated group than in the context of a conservative group. We also replicated several important results in the stereotype change phase of the experiment. As in Experiment 1, stereotype-disconfirming information had a smaller impact on the context-relevant dimension than on the context-irrelevant dimension. In addition, participants perceived the group as less variable on the context-relevant dimension, following stereotype disconfirmation. Finally, stereotype-disconfirming group members were perceived as more atypical if their behavior disconfirmed the context-relevant stereotype than the context-irrelevant stereotype. Note that the pattern of results is again suggestive of a graded model of subtyping in which disconfirmers can have a greater or lesser impact on stereotype change, depending on how atypical they are perceived to be from the group average. This conclusion is drawn from the finding that even stereotypes on context-relevant dimensions, for whom disconfirmers were thought to be relatively less typical of the group, reflected some change

<sup>6</sup> In the percentage estimate task, the slope coefficient for the effect of context was 27.67. When controlling for the typicality measure, the slope was reduced to 19.42. Similarly, in the mean ratings, the slope coefficient for the effect of context was .79. When typicality was controlled for, the slope was reduced to .63.

compared to confirming and no-information conditions (see Fig. 1).

Beyond replicating and strengthening the effects of the first study, the results of our second experiment also extend our conclusions in two important ways. In particular, results from the no-information condition provided a number of important insights. Data from the no-information condition in Experiment 2 allowed us to disentangle the locus of stereotype change effects by comparing stereotypes formed on the basis of the initial set of information with stereotypes held after receipt of additional disconfirming or confirming information. Without exception, there were no differences in the stereotypes held by participants who received additional stereotype-consistent information during the second phase of the experiment and those who received no additional information about the group. In contrast, stereotypes were consistently weaker (less extreme and more dispersed) after disconfirming information than after either confirming or no additional information, especially on context-irrelevant dimensions. Given that the results from the confirming and no-information conditions were virtually identical, it is reasonable to conclude that effects we observed in both experiments were due to learning disconfirming information about a group, as was hypothesized.

Second, data from the no-information condition allow us to speculate about the role of context in inducing a greater perceived within-group homogeneity. The context literature that we reviewed in the introduction suggested that contexts sometimes lead to increases in within-group homogeneity and sometimes not. If we look at measures of within-group homogeneity or variability in Table 5, we see that this assimilation effect of

context was obtained only very weakly for the no-information condition. That is, the group was judged to be only slightly (and non-significantly) more homogeneous on the context-relevant dimension than on the context-irrelevant one. However, after receiving additional disconfirming information about the group (on both attribute dimensions), the differences in perceived homogeneity as a function of context emerged. This perhaps suggests that the effect of context on perceived within-group homogeneity depends critically on the presence or absence of relatively disconfirming information about the group. We return to this point in the general discussion.

### General discussion

In the current research, we extended inter-group context effects on stereotyping to the arena of stereotype change. We hypothesized and found that stereotype change is less successful in situations where the stereotypic attribute to be disconfirmed is relevant to distinguishing a stereotyped group from its social context. This is because (1) stereotypes formed in an inter-group context are more extreme than those formed without a context and (2) disconfirmers are seen as more atypical of a stereotype that has been formed in context and are therefore more likely to be subtended.

Results from multiple measures of the perceived stereotypicality supported these predictions. Consistent with prior research, our participants formed more extreme stereotypes about the focal group on dimensions that distinguished it from a context group than on dimensions that did not. More important for the current research, the focal group stereotype was more resistant

Table 5  
Experiment 2: Perceived within-group variability means (standard deviations)

| Confirmation condition              | Context condition | Ratings on academic dimension | Ratings on political dimension |
|-------------------------------------|-------------------|-------------------------------|--------------------------------|
| <i>Variance from histogram task</i> |                   |                               |                                |
| Disconfirm                          | Academic          | 1.64 (0.52)                   | 2.20 (0.60)                    |
|                                     | Political         | 1.38 (0.52)                   | 1.33 (0.53)                    |
| No-information                      | Academic          | 0.78 (0.44)                   | 1.12 (0.54)                    |
|                                     | Political         | 0.83 (0.51)                   | 0.87 (0.60)                    |
| Confirm                             | Academic          | 0.93 (0.76)                   | 1.36 (0.87)                    |
|                                     | Political         | 0.71 (0.40)                   | 1.03 (0.73)                    |
| <i>Within-group similarity task</i> |                   |                               |                                |
| Disconfirm                          | Academic          | 5.30 (1.06)                   | 3.87 (0.97)                    |
|                                     | Political         | 4.38 (1.24)                   | 4.69 (0.97)                    |
| No-information                      | Academic          | 6.08 (0.86)                   | 5.18 (0.97)                    |
|                                     | Political         | 5.81 (0.80)                   | 5.27 (1.15)                    |
| Confirm                             | Academic          | 5.96 (0.94)                   | 5.35 (1.00)                    |
|                                     | Political         | 6.08 (1.06)                   | 5.73 (1.00)                    |

Note. In the top panel, higher values indicate a greater perceived range (i.e., more variability). In the lower panel, higher values indicate a greater perceived similarity (i.e., less variability).

to change in context-relevant dimensions compared to context-irrelevant dimensions, after exposure to disconfirming information. Thus, both experiments provided evidence that stereotypes formed in an inter-group context are more extreme than those formed in isolation and that such context-based stereotypes are more likely to be preserved in the face of disconfirmation.

Context effects on perceptions of within-group variability were a bit more complicated to interpret. Collapsing across conditions, our data provided no evidence that the focal group was seen more homogeneously on context-relevant than on context-irrelevant dimensions in either experiment. Our failure to find overall within-group assimilation effects is, however, consistent with the mixed results found from previous research (Krueger & Rothbart, 1988; Tajfel & Wilkes, 1963).

However, the predicted effect of context on perceptions of within-group variability was found in both experiments for those participants in the disconfirm conditions, who saw the focal group as less diverse on context-relevant than on context-irrelevant dimensions. Given that context effects on variability did not occur in non-disconfirm conditions, this set of results suggests that stereotype-disconfirming information is a necessary ingredient for context effects to emerge on perceived within-group variability. The presence of disconfirming information may have led participants to ‘chop off’ the part of their perceived distribution for the group closest to the disconfirmers, thus, resulting in less overall perceived variability.

Our argument for the effect of context on stereotype change hinged on the degree to which disconfirming group members on context-relevant versus context-irrelevant dimensions were more likely to be subtyped. We predicted a mediational role for disconfirmer typicality with more extreme stereotypes (those formed in context) harder to change because disconfirmers would be perceived as more atypical of the group and would be more likely to be subtyped. Partial support for this hypothesis was garnered from both experiments from the finding that the tendency to see disconfirmers as less typical of the group than confirmers was especially true when they disconfirm context-relevant stereotypes than context-irrelevant stereotypes.

Our mediational analyses offer further support for the mediational role of judged typicality in producing these stereotype change effects. However, we also tested a reverse mediational model in which stereotype change mediated the effects of confirmation condition on judged typicality. Consistent with Maurer et al. (1995) (see also Hewstone & Hamberger, 2000), we found evidence for both of these mediational models. One should probably conclude our manipulation effects on typicality and stereotype change are partially independent effects.

While much of the previous research has acknowledged the role of exemplar typicality (e.g., Rothbart,

1996; Rothbart & John, 1985; Rothbart & Lewis, 1988) and degree of exemplar disconfirmation (Weber & Crocker, 1983) in facilitating or blocking stereotype change, our results highlight the importance of considering the nature of the stereotype targeted for change. Note that our argument for the consideration of characteristics of the stereotype to be changed is similar to that of Kunda and Oleson (1997). Those authors posited that stereotype change ought to depend not only on the extremity of the attributes of disconfirming members but also on the extremity of the stereotype one is trying to alter. Perceivers with moderate and extreme stereotypes were exposed to a counterstereotypic group member who was either moderately or extremely disconfirming of the stereotype. Results showed that extreme-stereotype participants perceived moderately and extremely disconfirming exemplars as less stereotypic than did moderate-stereotype participants. Thus, consistent with the current research, the same disconfirmer was seen as more or less typical of the group based on differences in the mean group stereotype.

Furthermore, the success of stereotype change depended on the extremity of the stereotype participants held. Extreme-stereotype participants continued to see the group in more stereotypic terms after disconfirmation than did moderate-stereotype participants. Interestingly, extreme-stereotype participants also exhibited a boomerang effect (compared to control extreme-stereotype participants) when exposed to an extremely disconfirming group member, perceiving the group as a whole to be even more stereotypic.

#### *Possible mechanisms underlying context effects on stereotype change*

It is useful at this point to be very concrete about the underlying mechanism that we believe to be responsible for the effects that we have reported. Additionally, we want to briefly discuss an alternative mechanism that seems to us to merit further empirical consideration.

From our data, it seems clear that a context influences the perceived stereotypicality of a focal group on those dimensions that are relevant to the context. Thus, a group is seen more extremely or more stereotypically on context-relevant than irrelevant trait dimensions. This effect emerges clearly from the analysis of the initial summary paragraphs written by participants in both studies and from the ratings of participants in the no-information condition of Experiment 2. Interestingly, however, there is very little evidence from the no information participants that context induces a greater within-group similarity on context-relevant than context-irrelevant dimensions.

When disconfirming information is subsequently encountered both on context-relevant and context-irrelevant trait dimensions, that disconfirming information is judged to be less typical of the group as a whole on

context-relevant dimensions than irrelevant ones because the group stereotype as a whole is more extreme or stereotypic on the context-relevant dimensions. As a result, while the group stereotype becomes less extreme as a result of the disconfirming information, that shift is less dramatic when the counterstereotypic information is context-relevant than when it is irrelevant (see Fig. 1).

Additionally, it is at this point, only after counterstereotypic information has been encountered, that the effects of a context on perceived within-group homogeneity emerge (see Fig. 2). Learning disconfirming information on both dimensions leads to a less within-group homogeneity, but this effect is attenuated on context-relevant dimensions.

In addition to the above mechanism, it seems plausible that there are others that could be invoked to explain the effects we have documented. For instance, a context not only may cause stereotypes to be more extreme on relevant dimensions but it may also make those relevant dimensions somehow more useful or diagnostic than irrelevant dimensions. In other words, even without an extremity shift on context-relevant dimensions, it may be harder to disconfirm a stereotype on those dimensions simply because those dimensions are more likely to be the focus of spontaneous attention and are presumed to be more diagnostic in defining what the group is about (Ford & Stangor, 1992; Trope & Mackie, 1987).

Obviously, our data cannot be used to make this argument, since in fact we do have extremity shifts as a function of context. Nevertheless, it seems worthwhile to try to tear apart this explanation, due to the relative salience of a dimension, from the extremity mechanism that seems most plausible in our data. Further research along these lines seems warranted.

Allowing for multiple mechanisms that may produce the effects we have observed, it seems clear from our data that stereotype change depends not only on the nature of the disconfirming information that is encountered but also on the nature of the stereotype itself. In retrospect, this conclusion seems hardly surprising: stereotypes that are more extreme or confidently held are harder to change and contexts make stereotypes more extreme and confidently held on trait dimensions relevant to the context. Yet, we believe that our conclusions are considerably more nuanced than this argument, since contexts are malleable and so are group stereotypes. The same group will be thought about a bit differently in two different context situations and this difference can be affected by relatively transitory manipulations. If this is the case, then stereotype change that is targeted along one dimension would best be accomplished if one temporarily made salient a given context group that was relevant along another stereotypic dimension than the one targeted for change. Just as one is able to retrieve multiple representations about focal groups, depending

on the context in which that retrieval takes place (Wittenbrink, Judd, & Park, 2001), so to stereotype change may depend on the particular group representation that is retrieved at the particular time when counterstereotypic information is encountered.

## Acknowledgments

The research was supported by Grant R01 MH 45049 from the National Institute of Mental Health to Charles M. Judd and Bernadette Park.

## References

- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Babey, S. H. (1999). Intergroup effects in the mental representation of stereotypes. Unpublished doctoral dissertation, University of California, Santa Barbara.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.
- Brown, R., Vivian, J., & Hewstone, M. (1999). Changing attitudes through intergroup contact: The effects of group membership salience. *European Journal of Social Psychology*, *29*, 741–764.
- Cornille, O., & Judd, C. M. (1999). Accentuation and sensitization effects in the categorization of multifaceted stimuli. *Journal of Personality and Social Psychology*, *77*, 927–941.
- Dooje, B., Haslam, S. A., Spears, R., Oakes, P. J., & Koomen, W. (1998). The effect of comparative context on central tendency and variability judgements and the evaluation of group characteristics. *European Journal of Social Psychology*, *28*, 173–184.
- Eiser, J. R. (1971). Enhancement of contrast in the absolute judgment of attitude statements. *Journal of Personality and Social Psychology*, *17*, 1–10.
- Ford, T. E., & Stangor, C. (1992). The role of diagnosticity in stereotype formation: Perceiving group means and variances. *Journal of Personality and Social Psychology*, *63*, 356–367.
- Ford, T. E., & Tonander, G. R. (1992). The role of differentiation between groups and social identity in stereotype formation. *Social Psychology Quarterly*, *61*, 372–384.
- Haslam, S. A., & Turner, J. C. (1995). Context-dependent variation in social stereotyping: III. Extremism as a self-categorical basis for polarized judgement. *European Journal of Social Psychology*, *25*, 341–371.
- Hewstone, M., & Hamberger, J. (2000). Perceived variability and stereotype change. *Journal of Experimental Social Psychology*, *36*, 103–124.
- Johnston, L., & Hewstone, M. (1992). Cognitive models of stereotype change: (3) Subtyping and the perceived typicality of disconfirming group members. *Journal of Experimental Social Psychology*, *28*, 360–386.
- Johnston, L., Hewstone, M., Pendry, L., & Frankish, C. (1994). Cognitive models of stereotype change: (4) Motivational and cognitive influences. *Journal of Experimental Social Psychology*, *24*, 237–265.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, *5*, 602–619.
- Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods*, *6*, 115–134.

- Judd, C. M., & Park, B. (1993). Definition and assessment of accuracy in social stereotypes. *Psychological Review*, *100*, 109–128.
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (vol. 1) (fourth ed., pp. 233–268). Boston: McGraw-Hill.
- Krueger, J., & Rothbart, M. (1988). Use of categorical and individuating information in making inferences about personality. *Journal of Personality and Social Psychology*, *55*, 187–195.
- Krueger, J., & Rothbart, M. (1990). Contrast and accentuation effects in category learning. *Journal of Personality and Social Psychology*, *59*, 651–663.
- Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviates. *Journal of Personality and Social Psychology*, *68*, 565–579.
- Kunda, Z., & Oleson, K. C. (1997). Extremity of deviance and subtyping. *Journal of Personality and Social Psychology*, *72*, 965–979.
- Livingston, K. R., Andrews, J. K., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology: Learning, Memory Cognition*, *24*, 732–753.
- Maurer, K. L., Park, B., & Rothbart, M. (1995). Subtyping versus subgrouping processes in stereotype representation. *Journal of Personality and Social Psychology*, *69*, 812–824.
- McCauley, C., & Stitt, C. L. (1978). An individual and quantitative measure of stereotypes. *Journal of Personality and Social Psychology*, *36*, 929–940.
- McGarty, C., & Penny, R. E. (1988). Categorization, accentuation and social judgement. *British Journal of Social Psychology*, *27*, 147–157.
- Oakes, P. J., Haslam, S. A., & Reynolds, K. J. (1999). Social categorization and social context: Is stereotype change a matter of information or of meaning? In D. Abrams, & M. A. Hogg (Eds.), *Social identity and social cognition* (pp. 55–79). Malden, MA: Blackwell Publishers.
- Oakes, P., Haslam, S. A., & Turner, J. C. (1998). The role of prototypicality in group influence and cohesion: Contextual variation in the graded structure of social categories. In S. Worchel, & J. F. Morales et al. (Eds.), *Social identity: International perspectives* (pp. 75–92). Thousand Oaks, CA: Sage Publications.
- Pettigrew, T. F., & Tropp, L. R. (2000). Does intergroup contact reduce prejudice: Recent meta-analytic findings. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 93–114). Mahwah, NJ: Lawrence.
- Rothbart, M. (1996). Category-exemplar dynamics and stereotype change. *International Journal of Intercultural Relations*, *20*, 305–321.
- Rothbart, M., Davis-Stitt, C., & Hill, J. (1997). Effects of arbitrarily placed category boundaries on similarity judgments. *Journal of Experimental Social Psychology*, *33*, 122–145.
- 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.
- Schell, T. L. (1997). A decision-bound model of stereotypes: Implications for stereotype representation, relativity, and accuracy. Unpublished doctoral dissertation. Santa Barbara: University of California.
- Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British Journal of Psychology*, *54*, 101–114.
- Trope, Y., & Mackie, D. M. (1987). Sensitivity to alternatives in social hypothesis-testing. *Journal of Experimental Social Psychology*, *23*, 445–459.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. New York: Basil Blackwell.
- Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, *45*, 961–977.
- Wilder, D. A. (1993). The role of anxiety in facilitating stereotypic judgments of outgroup behavior. In D. M. Mackie, & D. L. Hamilton (Eds.), *Affect cognition and stereotyping: Interactive processes in group perception* (pp. 87–109). Academic Press: San Diego.
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Evaluative versus conceptual judgments in automatic stereotyping and prejudice. *Journal of Experimental Social Psychology*, *37*, 244–252.