Journal of Personality and Social Psychology
1994, Vol. 67, No. 3, 414–429

0022-3514/94/$3.00

Illusory Correlation in the Perception of Groups: An Extension of the Distinctiveness-Based Account

Allen R. McConnell, Steven J. Sherman, and David L. Hamilton

The prevailing explanation for illusory correlation in the stereotyping of groups is that distinctive information (minority groups' infrequent behaviors) is salient, receives enhanced encoding, and becomes highly accessible, thus biasing subsequent judgments. This distinctiveness-based explanation (DBE) depends on information distinctiveness at the time of its encoding. Information distinctiveness at encoding was manipulated, while ultimate distinctiveness was kept constant. Experiment 1, contrary to the DBE, found illusory correlations emerge regardless of distinctiveness at encoding. Experiment 2 collected process data that showed that ultimately distinctive behaviors were highly accessible at the time of judgment even when they were not distinctive at encoding. Experiments 3–5 ruled out an alternative account. A basis for illusory correlation that depends on postpresentation, but prejudgment, encoding of distinctive information is suggested.

Illusory correlation refers to the erroneous judgment of a relation between uncorrelated information categories. Chapman's (1967) original demonstration of this phenomenon showed that the co-occurrence of paired distinctive stimuli resulted in an overestimation of the frequency of such pairings. These distinctive pairs were presumably processed more thoroughly than other word pairs because of their salience. Illusory correlation was first proposed as a mechanism in stereotype formation by Hamilton and Gifford (1976). In their study, subjects read behavior-descriptive sentences about members of two groups (labeled A and B) who exhibited the same ratio of desirable to undesirable behaviors (9:4). Thus, undesirable behaviors were far more infrequent than desirable behaviors. In addition, twice as many items of information referred to Group A as to Group B, rendering Group B the "minority group." After reading those sentences, subjects' judgments revealed an illusory correlation between Group B and negative behaviors. That is, rather than estimating equal proportions of desirable and undesirable behaviors by both groups and rather than showing equally positive impressions of the two groups, subjects overestimated the frequency of negative behaviors by Group B members and evaluated Group B more negatively.

Following Chapman's (1967) reasoning, Hamilton and Gifford (1976) proposed that subjects overestimated the frequency of undesirable Group B behaviors (B−) because those exemplars were less frequent during presentation, were more distinctive than frequent items, received more extensive encoding, and therefore were more accessible in memory. This enhanced accessibility presumably led to biases in retrieval, errors of frequency estimation, and biased impressions of the groups during the subsequent judgment tasks.

There are two cornerstones of this distinctiveness-based interpretation. First, infrequent or otherwise salient items are encoded more extensively at exposure and thus are more accessible later, when judgments are made. Second, judgments are memory-based, rather than on-line, in nature. That is, judgments depend on information that is retrieved at the time of judgment rather than on an impression process that occurs at the time the information is initially encoded (Hastie & Park, 1986).

Recent research on illusory correlation has supported this distinctiveness-based explanation (hereafter labeled DBE; for reviews see Hamilton & Sherman, 1989; Mullen & Johnson, 1990). For example, Hamilton, Dugan, and Trolier (1985) tested the role of encoding strength. Subjects who were presented the behavior items in the usual manner showed the expected illusory correlation effects. In a second condition, subjects shown only a summary table revealing the distribution of desirable and undesirable items by Groups A and B did not demonstrate the illusory correlation. Most important, in a third condition, subjects who were presented with the behavior items and then the summary table reported liking Group B less than Group A, despite having access to the summary table. These results support a strength of encoding explanation. Hamilton et al. (1985, p. 10) concluded, "This finding indicates that having the summary information available after the serial presentation was not sufficient to eliminate the bias that developed during the initial encoding of stimulus sentences." Subsequently,

Allen R. McConnell and Steven J. Sherman, Department of Psychology, Indiana University; David L. Hamilton, Department of Psychology, University of California, Santa Barbara.

This research was supported by Grant MH 40058 from the National Institute of Mental Health. Portions of this research were presented at the 65th annual meeting of the Midwestern Psychological Association, April 1993, Chicago, Illinois.

We would like to thank Erika Hopper, Dawn Horn, Kay Lawson, Angela Leinenbach, Linda Mendonca, Cynthia Strader, and Erin Winchester for their assistance in conducting these experiments. Also, we appreciate the helpful comments offered by Klaus Fiedler, Igor Gavanski, Brian Mullen, Sharon Shavitt, and Eliot Smith on drafts of this article. Finally, we would like to thank Russ Fazio, Rob Goldstone, Reid Hastie, Ed Hirt, Reed Hunt, Rob Nosofsky, Don Robinson, Dick Sherman, Rich Shiffrin, Teresa Treat, and Ray White for their suggestions.

Correspondence concerning this article should be addressed to Allen R. McConnell or Steven J. Sherman, Department of Psychology, Indiana University, Bloomington, Indiana 47405-1301.

Stroessner, Hamilton, and Mackie (1992) obtained more direct evidence that subjects encoded B− items more extensively than the other three categories. In their experiment, subjects controlled the presentation rate of the stimulus items by advancing to the next item at their own pace. Stroessner et al. found that subjects spent more time reading B− items (the distinctive items) than the other three stimulus classes, indicative of greater encoding for the B− items.

If the B− items are differentially encoded at the time of presentation, then they should be more accessible to influence subsequent judgments. Evidence for this was obtained by Johnson and Mullen (1994), who measured subjects' response latencies in a group-assignment task. Subjects read 36 behavioral items about members of Groups A and B (Group B and undesirable behaviors were the infrequent classes of stimuli) and later were asked to indicate the group membership (A or B) of the person who performed each behavior. Subjects' responses were faster for B− behaviors than the other three group-behavior categories (A+, A−, and B+). Finally, Hamilton et al. (1985) reported free recall data that provided support for the memory-based judgment aspect of the DBE. Recall of behaviors and evaluative judgments for Group B were strongly correlated—an indication of memory-based judgments (Hastie & Park, 1986).

These diverse findings support the argument that information that is distinctive at the time of presentation is processed more thoroughly and encoded more extensively than other information. Later, when judgments based on the totality of information must be made, these distinctive items, having been extensively encoded, are more easily retrieved from memory. These items thus carry disproportionate weight in the judgment process, producing biased perceptions in the form of illusory correlations.

A clear assumption of the DBE is that the information must be distinctive at the time of encoding. That is, the mechanism for illusory correlation formation is an encoding process based on the salience of certain stimulus items. Thus, whenever a B− item is encountered, its distinctiveness due to its infrequency induces greater processing. According to this account, it is the distinctiveness of the B− items at the time they are presented that is responsible for the bias in judgments.

What might be expected if, at the time of presentation of the B− items, these items were not infrequent relative to the other item types and were therefore not distinctive? According to the DBE, these items would not receive extra processing, would not be extensively encoded, and would not play a biasing role in the judgment process. Therefore, an illusory correlation would not be expected in this case. All previous illusory correlation studies have presented stimuli in a randomized order. Hence, all B− items have had a high and equal level of distinctiveness throughout the presentation list. As B− items are randomly dispersed through the list, each B− item is distinctive as it is encountered, regardless of whether the item appears relatively early or late in the list of behaviors.

However, it is possible to vary the serial position of the various item types (A+, A−, B+, and B−) in a manner that would create differential stimulus class frequencies at different points in the serial presentation. In other words, the total number of B− items can be held constant, but the placement of the B− items in the list can make them distinctive or not at the time they are encountered. Because distinctiveness is assumed to vary as a function of a stimulus class's frequency relative to all other classes at the time of encoding (Schmidt, 1991), this manipulation should influence the extent of encoding of the B− items and hence the likelihood of forming an illusory correlation.

Although the DBE proposes that illusory correlation is based on the distinctiveness of information at the time of encounter, other findings suggest that illusory correlation formation will not be dependent on the infrequency of B− items at the time of presentation but rather on the ultimate infrequency of B− items. Von Restorff's work (cited in Koffka, 1935) has typically been interpreted as showing that items that are distinctive or perceptually–conceptually salient at the time of presentation produce better recall due to encoding depth or strength. Recently, Hunt and McDaniel (1993) have noted that, in several experiments, von Restorff found that items that were not distinctive at the time of presentation were better remembered if they were made distinctive only later, because of the nature of the remaining items on the presentation list.

For example, a nonsense syllable embedded in a series of numbers would be salient (because of its unique properties relative to the series of numbers) when presented in a one-at-a-time fashion. However, if the nonsense syllable is presented early in the series of numbers (e.g., second on the list), its uniqueness would not be apparent at the time of presentation, and thus it would not be perceptually salient. Despite a lack of distinctiveness at the time of encounter, von Restorff found that such items (in this example, nonsense syllables) were better recalled than other items (Koffka, 1935). Hunt and McDaniel (1993, p. 422), commenting on this finding, said, "The subject had no reason to perceive the isolate (using the above example, the nonsense syllable) as unusual. This point is relevant to contemporary ideas about distinctiveness in that most theories assume that distinctiveness is a product of perceptual salience." In other words, salience may not become apparent until after presentation. Extended to the illusory correlation paradigm, von Restorff's work predicts that the infrequency of B− items (relative to A+, A−, and B+) at the time of encounter may not be a necessary condition for illusory correlation formation. Hence, illusory correlations might emerge regardless of the distinctive category's infrequency at the time of presentation, as long as the B− items are infrequent in the overall list.

The question, then, is whether the formation of an illusory correlation is necessarily limited to a differential encoding process based on the distinctiveness of certain stimulus information at the time of presentation, or whether a postencoding process could produce the same kind of bias in the perception of groups.

Experiment 1

Experiment 1 examined the role of item distinctiveness in illusory correlation formation by varying the presentation order of B− stimuli in a way that altered their relative frequency at the time they were encountered. Three conditions were compared: a balanced presentation, a primacy-loaded condition, and a recency-loaded condition. In the balanced condition, the B− items were distributed throughout the stimulus list at a constant rate of infrequency (relative to the three other classes of stimuli, A+, A−, and B+) and thus served to replicate standard illusory correlation experiments. In the primacy-loaded condition, all the B− items were presented early in the list, such that all four classes of stimuli had the same frequency at the time that all the B− items were presented. Thus, in this condition, B− items were not infrequent at the time they were encountered (relative to the three other classes), and given their lack of distinctiveness (i.e., infrequency), this arrangement should eliminate the illusory correlation effect according to the standard DBE. However, if postencoding processes can operate on ultimately distinctive stimuli, an illusory correlation should emerge even in this condition.

In the recency-loaded condition, no B− items were presented until late in the list. Thus, when encountered, B− items were especially distinctive relative to the other two experimental conditions. Because of this special distinctiveness of B− items at the time they are encountered, the DBE would predict that illusory correlations in this condition would be accentuated relative to the balanced condition.

In summary, the focus of Experiment 1 is the following question: Is illusory correlation necessarily based on an encoding process or can postencoding mechanisms produce the effect as well? If distinctiveness at the time of encoding is an essential basis of the process, then illusory correlation should be replicated in the balanced condition (where B− items are infrequent at the time of encoding), strengthened in the recency-loaded condition (where B− items are especially infrequent at the time of encoding), and eliminated in the primacy-loaded condition (where B− items are not infrequent at the time of encoding). Alternatively, if the distinctiveness of the B− items can be achieved through postencoding processes, as suggested by von Restorff's work, then illusory correlations would be evident in all three conditions.

Method

Subjects

Subjects were 102 undergraduate students at Indiana University who participated in return for research experience credit in introductory psychology courses. They were randomly assigned (34 per condition) to the three experimental conditions.

Stimuli

A list of 36 stimuli was used: 16 Group A desirable (A+), 8 Group A undesirable (A−), 8 Group B desirable (B+), and 4 Group B undesirable (B−) behavior descriptions were presented. As in Hamilton and Gifford (1976), each stimulus item consisted of a male name, a group designation (either A or B), and a behavior. For example, "Jim, a member of Group A, visited a sick friend in the hospital." Pilot testing of stimuli ensured that the overall evaluation of both desirable and undesirable behaviors ascribed to each group was equivalent.

Design

The experiment featured three conditions: balanced, primacy-loaded, and recency-loaded presentations. The 36 stimulus items were subdivided into nine partitions, each containing four items. Within each presentation, four of the nine partitions (designated as "critical partitions") contained one item from each of the four stimulus classes (A+, A−, B+, and B−). Thus, all of the B− items (4) appeared in these four critical partitions (one B− in each). Within each critical partition, the two positive and the two negative behaviors were rated as equivalent in desirability. During presentation, group membership was randomly assigned to both the positive and negative behaviors in these critical partitions. The serial presentation orders for the balanced, primacy-loaded, and recency-loaded conditions are shown in Table 1.

Balanced condition. The balanced condition featured critical partitions in Partitions 2, 4, 6, and 8. All other partitions had a random distribution of the remaining A+, A−, and B+ items. This condition served as a replication of previous studies where the relative infrequency of B− items remains essentially high and constant throughout the stimulus presentation.

Primacy-loaded condition. The primacy-loaded condition featured critical partitions in Partitions 1, 2, 3, and 4. Items 17–36 (partitions 5–9) consisted of a random distribution of the remaining A+, A−, and B+ items. In the primacy-loaded condition, the B− items were not infrequent relative to the other three stimulus classes at the time they were encountered, and thus B− items should not be distinctive at the time of encoding.

Recency-loaded condition. The recency-loaded condition featured critical partitions in Partitions 6, 7, 8, and 9. Items 1–20 (Partitions 1–5) had a random distribution of the remaining A+, A−, and B+ items. In this condition, the B− items were especially infrequent at the time they were encountered by the subject. Such a configuration should make B− items very distinctive at the time of encoding relative to the other two conditions.

Procedure

Methodology and instructions followed the Hamilton and Gifford (1976) paradigm with the exception that subjects were not told that Group B (the minority group) would appear less frequently than Group A. This omission was necessary to prevent subjects (especially in the primacy-loaded condition) from knowing that a particular group was infrequent. Subjects were run at individual computer workstations located in individual rooms. They were told that the experiment concerned "how people process and retain information that is presented to them visually" (Hamilton & Gifford, 1976, p. 395). In addition, they were told that they would read about behaviors of members of Group A and Group B, that both groups represented groups in the real world, and that their task in the experiment was to "simply read each statement carefully." Each of the 36 stimulus items was presented on the computer monitor for 8 s. After the display of the appropriate stimuli (as dictated by the experimental condition), subjects completed a 4-min filler task, a free-recall task, a group-assignment task, a frequency-estimation task, and an evaluation of the likability of members of each group. Stimulus presentation and data collection were controlled by the computer.

Free recall. After the filler task, subjects were provided with a blank piece of paper and were instructed by the computer to write down as many of the behaviors as they could recall. Also, they were asked to write down the group that engaged in the behavior. If they could recall a behavior but not the group that was associated with it, they were encouraged to guess (Hamilton et al., 1985).

Group assignments. The 36 behaviors (without group association) were presented to the subjects in a randomized order. Subjects read each item, then indicated using the keyboard whether a member of Group A or Group B performed the action.

Frequency estimates. After the group-assignment task, subjects were told that Group A performed 24 behaviors and were asked to estimate how many of them were undesirable. Next, subjects were told that Group B performed 12 behaviors and were asked to estimate how many of them were undesirable.

Likability ratings. Finally, subjects were asked to rate, on a scale ranging from 1 to 10 (where 1 represents *strong disliking* and 10 repre-

Table 1
Sequential Presentation of Stimuli as a Function of Experimental Condition

Condition	Partition 1	2	3	4	5	6	7	8	9
Balanced	x	CRIT	x	CRIT	x	CRIT	x	CRIT	x
Primacy	CRIT	CRIT	CRIT	CRIT	x	x	x	x	x
Recency	x	x	x	x	x	CRIT	CRIT	CRIT	CRIT
Items	1–4	5–8	9–12	13–16	17–20	21–24	25–28	29–32	33–36

Note. x = random mix of A+, A−, B+; CRIT = contains one A+, A−, B+, and B−.

sents *strong liking*), how much they thought they would like members of Group A and Group B.

Results

Measures of Illusory Correlation

Likability ratings. A 3 (experimental conditions: balanced, primacy, or recency) × 2 (group membership: A vs. B, a within-subjects factor) mixed-design analysis of variance (ANOVA) was conducted on the subjects' evaluations of the groups. As Table 2 illustrates, subjects preferred members of Group A to members of Group B (M = 6.48 vs. M = 5.27, respectively), $F(1, 99) = 10.21$, $p < .001$, replicating previous findings. In addition to this group effect, likability varied as a function of experimental condition, $F(2, 99) = 4.19$, $p < .02$, with greater total liking for Groups A and B in the recency condition (M = 6.26) than in the balanced (M = 5.75) or primacy conditions (M = 5.62). This finding suggests some degree of on-line processing in the formation of group evaluations because the recency-loaded condition puts more desirable behaviors toward the beginning of the sequential presentation, and an on-line processing account would predict primacy effects for impressions.

Table 2
Means for Likability Ratings, Frequency Estimates, and Group Assignments as a Function of Experimental Conditions in Experiment 1

Variable	Experimental condition: Balanced	Primacy	Recency
Likability ratings			
Group A	6.53	6.15	6.77
Group B	4.97	5.09	5.74
difference = *0*	1.56	1.06	1.03
Frequency estimates			
$p(-\mid B) = .33$	0.54	0.57	0.51
$p(-\mid A) = .33$	0.39	0.41	0.46
difference = *0*	0.15	0.16	0.05
Group assignments			
$p(A\mid +) = .67$	0.62	0.57	0.62
$p(A\mid -) = .67$	0.51	0.46	0.52
difference = *0*	0.11	0.11	0.10

Note. Italicized figures indicate correct baserate values. Likability scale ranges from 1 to 10, with higher values indicating more favorable impressions of group members.

Most important, however, is the absence of the Group × Condition interaction predicted by the standard DBE, $F(2, 99) = .20$, *ns.* The illusory correlation was strong in all cases and it was neither attenuated in the primacy condition nor enhanced in the recency condition. The two conditions that should differ most according to the DBE (primacy and recency) were virtually identical in their level of preference for Group A over Group B. Thus, although these likability data replicate previous illusory correlation findings, results for the primacy and recency conditions do not support expectations based on the standard DBE.

Frequency estimates. A second measure of illusory correlation is the frequency estimates of negative behaviors engaged in by each group. Across all conditions, Table 2 shows that subjects viewed Group B as engaging in proportionately more undesirable behaviors than Group A (M = .54 vs. M = .42), $F(1, 99) = 13.34$, $p < .001$, in a Condition × Within-Groups mixed-design ANOVA, even though both groups had the same base rates (.33) for undesirable behaviors. Most important, this finding was not moderated by a Group × Condition interaction, $F(2, 99) = 1.08$, *ns.* As with the likability ratings, the frequency estimates demonstrated strong illusory correlations, but the magnitude of those effects did not vary across the three experimental conditions.

Group assignments. As shown in Table 2, subjects were much more likely to assign positive behaviors (M = .60) as opposed to negative behaviors (M = .50) to Group A, $F(1, 99) = 7.76$, $p < .01$, analyzed in a Conditions × Valence mixed-design ANOVA, where valence was a within-subjects measure, once again replicating previous illusory correlation findings. And although there was an effect of condition, $F(2, 99) = 3.49$, $p < .04$, on group assignments, demonstrating a greater tendency to assign more behaviors in general to Group A in the balanced (M = .56) and recency conditions (M = .57) than in the primacy condition (M = .51), there was no indication of the Condition × Valence interaction, $F(2, 99) = 0.01$, *ns,* predicted by the DBE. Thus, once again, subjects showed strong illusory correlations across all conditions, rather than attenuated effects in the primacy condition and enhanced effects in the recency condition.

In addition to examining the overall group-assignment data, it is also important to examine the group-assignment data from the critical partitions only. Because the presentation frequency

Table 3
Mean Proportion of Items Recalled in Free Recall Task as a Function of Stimulus Category in Experiment 1

Stimulus category	Type of analysis	
	General	Correct
A+	0.19	0.09
A−	0.27	0.14
B+	0.24	0.10
B−	0.28	0.13

of B− items was skewed in the primacy and recency-loaded conditions, the group-assignment data presented above might not allow for generalizations between B− items and other stimulus classes. In the recency-loaded condition, for instance, the B− behaviors occurred in the last four partitions (i.e., the last 16 items). It is conceivable that comparing these B− behaviors with other classes that occurred in the first five partitions (as well as the last four partitions) may produce a confound with serial presentation. For example, primacy effects might benefit A+, A−, and B+ items in the recency-loaded condition, providing them with better recall at the time of judgment.

To control for this possibility, we analyzed the group-assignment data by looking at only the 16 items (4 from each stimulus class) in the four critical partitions. This analysis allowed us to compare items that were precisely matched in terms of desirability from pretested norms as well as presentation location. The analyses of the critical partition group-assignment data exactly mirrored the finding of the overall group-assignment analysis. Thus, it seems that the overall effects in the group-assignment data were present in the critical partitions.[1]

Free recall. Two judges, blind to the experimental hypotheses, scored free-recall data by using a gist criterion. Interjudge agreement was high (95%), reflecting good reliability in stimulus assessment. In cases of disagreement, the decision of a third judge, also blind to the experimental hypotheses, prevailed. For each of the four stimulus categories, the proportion of behaviors recalled by each subject was computed and analyzed by a 3 × 2 × 2 (Experimental Condition × Group Membership × Behavioral Valence, with the last two factors being within-subjects variables) mixed-design ANOVA. Because subjects did not always associate a group with the behaviors they listed, two separate ANOVAs were conducted. First, the general free-recall ANOVA examined the behaviors recalled by subjects regardless of whether they correctly associated the group with the behavior. Second, the correct free-recall ANOVA examined only those behaviors recalled by the subject and correctly assigned to the group that engaged in the behavior.

As Table 3 illustrates, for general free recall, there was an effect of group membership, $F(1, 99) = 4.40$, $p < .04$, and behavioral valence, $F(1, 99) = 16.83$, $p < .001$.[2] Note that subjects in general recalled proportionately more Group B than Group A behaviors ($M = .26$ vs. $M = .23$) and proportionately more undesirable behaviors than desirable behaviors ($M = .28$ vs. $M = .22$). When looking at the correct free-recall data, only an effect of behavioral valence, $F(1, 99) = 14.00$, $p < .001$, was detected, with negative items being recalled better than positive items ($M = .14$ vs. $M = .10$).

Although subjects recalled more Group B behaviors and more undesirable behaviors in general free recall across all 36 behaviors, recall did not vary as a function of the three experimental conditions. The DBE would predict greatest B− recall in the recency-loaded condition (when the distinctiveness of B− items is highest) and least B− recall under the primacy-loaded condition (where the B− items are not at all distinctive during presentation). These results did not emerge.

The DBE posits that illusory correlations are memory-based judgments. If so, then free recall and judgments should be correlated. To explore this prediction, an index was computed for both general and correct free recall by subtracting the proportion of negative behaviors recalled for each group from the proportion of positive behaviors recalled. A positive index thus indicates proportionally greater recall for desirable behaviors than undesirable behaviors. This index was then used as a predictor for likability ratings for each group. For Group A, neither general, $F(1, 100) = 2.06$, $p > .15$, nor correct, $F(1, 100) = 0.18$, *ns*, free recall predicted likability. Likewise for Group B, neither general, $F(1, 100) = .36$, *ns*, nor correct, $F(1, 100) = .33$, *ns*, free recall predicted likability. In short, there was no evidence of a memory-based judgment despite the strong illusory correlations observed. As discussed in the group-assignment results, it is also beneficial to examine free recall for critical partition items in isolation from the other five partitions because it controls for the skewed quality of the primacy and recency-loaded conditions. We conducted the same ANOVAs (general and correct) and found the same strong main effect of valence in all three conditions.[3]

[1] We examined these critical partition group-assignment data in the same manner in which we examined the group-assignment data from all 36 items. First, we conducted a 3 (conditions) × 2 (behavioral valence, a repeated measure) mixed-design ANOVA on the group assignments of desirable and undesirable behaviors to Group A. In accord with our findings in the overall group-assignment data, subjects assigned more desirable than undesirable behaviors to Group A ($M = .59$ vs. $M = .47$), $F(1, 99) = 8.58$, $p < .01$. Also, as with the overall group-assignment data, their assignments were moderated by the experimental condition, $F(2, 99) = 3.40$, $p < .04$, indicating that subjects assigned more behaviors to Group A in the balanced ($M = .55$) and recency-loaded ($M = .55$) conditions than the primacy-loaded condition ($M = .48$).

[2] No effect of condition was found when analyzing our free-recall data. For the sake of clarity and because it had no main or interactive effects, only the means across the three experimental conditions are presented in Table 3.

[3] In these critical partition free-recall ANOVAs, we found that undesirable behaviors were better recalled than desirable behaviors in both the general ($M = .26$ vs. $M = .18$), $F(1, 99) = 17.92$, $p < .001$, and correct ($M = .10$ vs. $M = .06$), $F(1, 99) = 8.13$, $p < .01$, free-recall analyses. These findings mirrored the previous free-recall findings with all 36 behaviors. In addition to these effects, a Group × Condition interaction emerged, $F(2, 99) = 3.41$, $p < .04$, with the correctly assigned free-recall data. This effect indicates that subjects made more accurate assignments in free recall for Group A than Group B behaviors in the balanced ($M = .10$ vs. $M = .06$) and primacy-loaded ($M = .08$ vs. $M = .06$) conditions; however, they were more correct in assigning Group B than Group A behaviors in the recency-loaded condition during free recall ($M = .11$ vs. $M = .07$).

Discussion

In the balanced presentation condition, Experiment 1 replicated illusory correlation findings for evaluative judgments, frequency estimates, and group-assignment reports. However, the primacy-loaded and recency-loaded presentation conditions produced results that would not be anticipated by the standard DBE. Although the presentation orders were designed to eliminate or enhance the distinctiveness of B− items at the time of encounter, this manipulation had no effect on the strength of illusory correlations. These findings are consistent with von Restorff's conclusion that the ultimate isolation of a stimulus and not simply distinctiveness at presentation can impact on memory and judgment.[4]

What kinds of theories or mechanisms can account for the finding that illusory correlation is equally robust regardless of the distinctiveness of B− items at the time of presentation? First, there is a set of explanations for illusory correlation with group targets that do not depend at all on the distinctiveness of the B− items to account for the effect. Fiedler (1991; Fiedler, Russer, & Gramm, 1993) has proposed an information-loss account for illusory correlation. According to this account (Fiedler, 1991, p. 25), "part of the (encoded) information will be lost, and such information loss will be characterized by an overestimation of the lowest frequency." Fiedler's information-loss account claims that illusory correlation is essentially the product of regression effects, which are especially strong for low-frequency categories. Within the illusory correlation paradigm, of course, the B− category is the most infrequent group and would thus be most subject to information loss and large regression effects. Because memory for Group B behaviors should be poorer than memory for Group A behaviors according to this explanation, there should be greater inaccuracy for reports concerning Group B than Group A. Some of the data in Experiment 1 (e.g., overestimates of Group B behaviors in frequency estimates) are consistent with Fiedler's proposal. However, inconsistent with this position is the finding of better free recall for Group B behaviors than Group A behaviors.

Fiedler's information-loss account depends only on the ultimate relative frequencies of positive and negative behaviors engaged in by Group A and Group B. The order in which these behaviors appear in the list and the distinctiveness of the behaviors at the time of presentation are irrelevant to the size of the illusory correlation effect, according to this proposal. Fiedler would thus predict strong and equal illusory correlations across the three experimental conditions.

Smith (1991) has also proposed a model of illusory correlation that does not require the distinctiveness of B− items at the time of encoding. Smith's model includes no role for biases in attention, encoding, or retrieval. It is, instead, a memory model that depends on the storage and retrieval of specific exemplars. Predictions are based on the ultimate arithmetic difference between positive and negative behaviors that are engaged in by a target group. With 16 desirable and 8 undesirable behaviors, Group A ends up 8 units to the good. With 8 positive and 4 negative behaviors, Group B ends up only 4 units to the good. Predictions of a bias in group evaluations depend only on the final arithmetic differences and not on the sequencing of behaviors in the list or the distinctiveness of behaviors at the time of presentation. Thus, as did Fiedler (1991) in his information-loss model, Smith (1991) would predict strong and equal illusory correlations in the three experimental conditions of Experiment 1.

Thus, the Fiedler (1991) and Smith (1991) explanations can account for the fact that illusory correlation was insensitive to the distinctiveness of behavioral information at the time of encoding. Neither proposal assumes any role of differential attention to the distinctive minority group–minority behavior stimuli, and they do not depend on any bias in encoding. Both models involve processes that occur only at the time of retrieval or judgment. They are thus in stark opposition to the DBE, which proposes that differential attention to and differential encoding of distinctive stimuli underlie illusory correlations.

Although the Fiedler and Smith explanations can account for the findings of Experiment 1, they do not explicitly predict the results from studies, cited earlier, that provided evidence of special attention to and accessibility of the distinctive B− items (Johnson & Mullen, 1994; Stroessner et al., 1992). Furthermore, they offer no predictions or mechanisms for free-recall data. Thus, despite their compatibility with the results of Experiment 1, these models are not completely adequate to account for the accumulated findings on illusory correlation.

A second possibility for the results of Experiment 1 is a process that involves the distinctiveness of the B− items, but one that is not restricted to their distinctiveness at the time of encoding. This approach is consistent with the von Restorff findings described earlier and suggests that items can be distinctive at the time of encounter (as in the balanced and recency-loaded conditions) or can become distinctive after initial presentation (as in the primacy-loaded condition) but before judgment. This extended distinctiveness-based explanation (EDBE) posits that subjects process and rehearse both old and new information throughout the stimulus presentation. Even though information may not be distinctive at the time of encounter (as in our primacy-loaded condition), it can become distinctive as new in-

[4] To ensure that there was no illusory correlation after the first four partitions in the primacy-loaded condition, auxiliary data were collected from 24 additional subjects. Each subject was seated at an individual computer workstation like those used in Experiment 1, provided with the same instructions, and exposed to only the first 16 items of the primacy-loaded condition (as described in Experiment 1). Following these 16 items, subjects were presented with the same filler task, then responded to the same group assignment, frequency estimate, and likability judgments described in Experiment 1. All dependent measures were analyzed using two-group t tests. No significant differences were observed between reports for Groups A and B. For instance, likability ratings were equivalent for Groups A (M = 5.88) and B (M = 5.71), $t(23)$ = .28, *ns*. Also, subjects estimated the same number of undesirable behaviors for both Groups A (M = 4.08) and B (M = 4.38), $t(23)$ = .74, *ns*. Finally, group assignments did not vary in terms of desirable behaviors (Group A, M = 3.71; Group B, M = 4.29), $t(23)$ = .94, *ns*, or undesirable behaviors (Group A M = 4.08, Group B M = 3.92), $t(23)$ = .26, *ns*. Clearly, the first 16 items of the primacy condition did not possess any qualities to create differential judgments of Groups A and B. Because the primacy-loaded condition in the main study produced standard illusory correlation results, the processes that produced the differential group evaluations must have occurred after the fourth critical partition—after all the B− items had already been presented.

formation is processed and thus it can receive further encoding long after it has been presented. As new items are encountered, this information is processed in relation to previously encountered information. This theory thus maintains the importance of distinctive stimuli but extends the conditions under which stimuli will become psychologically distinctive. Again, the EDBE is consistent with the von Restorff findings discussed by Hunt and McDaniel (1993). Despite the fact that items displayed early during presentation are not distinctive at the time of encounter (i.e., the B− items in the primacy-loaded condition), subsequently received information can provide the additional context necessary for their distinctiveness (i.e., infrequency) to become apparent after the fact.

Thus, the pattern of results from Experiment 1 is compatible with both the models that do not require elaborative encoding of distinctive items (Fiedler, 1991; Smith, 1991) and with the EDBE. Notice, however, that the Fiedler and Smith explanations do not predict differential accessibility for B− items (relative to other stimulus categories) at the time of judgment. These explanations are based on regression effects or on retrieval of exemplars from memory. On the other hand, the EDBE specifically predicts that B− items would be more accessible than other stimulus categories at the time of judgment, even when the B− items were not distinctive at the time of presentation. Thus, in all presentation sequences, B− items should be more accessible than the other categories of information at the time of judgment.

Table 4
Means for Likability Ratings, Frequency Estimates, and Group Assignments as a Function of Experimental Conditions in Experiment 2

Variable	Experimental condition: Balanced	Primacy	Recency
Likability ratings			
Group A	6.67	6.30	6.47
Group B	5.07	4.60	5.87
difference = *0*	1.60	1.70	0.60
Frequency estimates			
$p(-\|B) = .33$	0.54	0.61	0.45
$p(-\|A) = .33$	0.46	0.50	0.38
difference = *0*	0.08	0.11	0.07
Group assignments			
$p(A\|+) = .67$	0.56	0.54	0.59
$p(A\|-) = .67$	0.51	0.49	0.57
difference = *0*	0.05	0.05	0.02

Note. Italicized figures indicate correct baserate values. Likability scale ranges from 1 to 10, with higher values indicating more favorable impressions of group members.

Experiment 2

In Experiment 2 we collected process data to differentiate these alternative accounts. The latency of subjects' responses on the group-assignment task can be used as evidence for differential accessibility of the different stimulus categories (Johnson & Mullen, 1994). If B− items are more accessible than other stimulus categories at the time of judgment, this finding would support the EDBE but would not be predicted by the information-loss explanation (Fiedler, 1991; Fiedler et al., 1993) or the exemplar-based model (Smith, 1991).

Method

Subjects

Ninety Indiana University undergraduates participated in return for research experience credit in introductory psychology courses. They were randomly assigned (30 per condition) to the three experimental conditions.

Procedure

Stimuli, experimental conditions, instructions, procedures, and dependent measures were exactly the same as those used in Experiment 1. The only difference between the two experiments was the inclusion of a latency measure for responses on the group-assignment task.

Results

Illusory Correlation Analyses

Likability ratings. A 3 × 2 (Experimental Condition × Group Membership, the latter variable being a within-subjects factor) mixed-design ANOVA was conducted on liking for Group A and Group B. As Table 4 illustrates, subjects preferred members of Group A to members of Group B (M = 6.47 vs. M = 5.18), $F(1, 87) = 11.64$, $p < .001$, replicating Experiment 1 and previous illusory correlation studies. Similar to Experiment 1, there was a marginal effect of condition, $F(2, 87) = 2.72$, $p < .08$, indicating that subjects expressed more liking for targets in the recency condition (M = 6.17) than in the primacy condition (M = 5.45). Importantly, replicating Experiment 1, there was no Group × Condition interaction, $F(2, 87) = 0.85$, *ns.*

Frequency estimates. Proportion of undesirable behaviors performed by members of Groups A and B was examined in a 3 (conditions) × 2 (groups, a within-subjects factor) ANOVA. As seen in Table 4, the same pattern of results as in Experiment 1 emerged. Subjects in all three presentation conditions reported that Group B engaged in proportionately more undesirable behaviors (M = .53) than Group A (M = .38), $F(1, 87) = 8.47$, $p < .01$. In addition, there was an effect of condition, $F(2, 87) = 7.89$, $p < .001$, indicating that subjects estimated fewer negative behaviors in the recency condition (M = .41) than the primacy condition (M = .56). The Group × Condition interaction again was not significant, $F(2, 87) = 0.14$, *ns.*

Group assignments. Responses on the group-assignment task were analyzed in a 3 (conditions) × 2 (positive vs. negative valence of behaviors assigned to Group A, a repeated measure) ANOVA. Although subjects demonstrated the same trend of assigning more positive behaviors (M = .57) than negative behaviors (M = .53) to Group A, the effect did not achieve significance, $F(1, 87) = .91$, *ns.* There was, however, an effect of condition, $F(2, 87) = 4.64$, $p < .02$, indicating that subjects assigned more behaviors to Group A in the recency condition (M = .59) than in the primacy condition (M = .52). Again, no Condition × Valence interaction emerged, $F(2, 87) = 0.04$, *ns.*

Response latency analyses. The principal focus of Experiment 2 was on the assessment of stimulus category accessibility. To examine category accessibility, mean response latencies were

Table 5
Group Assignment Response Latencies in Experiment 2

	Type of analysis		
Stimulus category	Overall	Correct	Incorrect
A+	3.19	3.14	2.86
A−	2.80	2.91	2.80
B+	2.93	3.02	3.32
B−	2.68	2.75	3.11

Note. Latencies are in seconds. Incorrect analysis conveys category to which a subject *assigned* a behavior. Thus, incorrect B− reflects p(B− | A−). Because experimental condition did not have a main or interactive effect, means are collapsed across conditions.

computed for each subject for each of the four stimulus categories (A+, A−, B+, and B−) in the group-assignment task. Individual responses that deviated beyond three standard deviations from the overall latency mean (M = 2.99 s, SD = 2.30) were excluded from the computation of the category means. We first analyzed our data using a Presentation Condition × Stimulus Category mixed-design ANOVA to assess whether the stimulus classes differed in terms of response latency and to assess whether presentation condition was a contributor to these differences. Then, we conducted a planned contrast comparing B− responses with the other three categories (following Johnson & Mullen, 1994).

Three analyses of latency data were conducted. First, overall latencies were analyzed regardless of whether subjects selected the correct group. Second, latencies for correct group assignments were analyzed. Third and finally, latencies for incorrectly assigned responses were assessed. If B− items are encoded more extensively, subjects should show faster response latencies to them in the overall and correct response analyses, but not in the incorrect response analysis (which would be indicative of a response bias). Faster B− responses for both correct and incorrect responses would be evidence of a response bias (i.e., fast guessing) because subjects would be indiscriminately and quickly assigning Group B to all negative behaviors. Latencies were analyzed in a 3 (conditions: balanced, primacy, or recency) × 4 (categories: A+, A−, B+, or B−, a repeated measure) mixed-design ANOVA.

For overall latencies, as displayed in Table 5, we found a significant effect for category, $F(3, 261) = 8.56$, $p < .01$, indicating that subjects' latencies varied across the four stimulus categories. Neither condition nor the interaction of condition and stimulus category produced significant effects. Next, we conducted the planned comparison to compare the latencies of B− items to the other three stimulus categories. This contrast was highly significant, $F(1, 261) = 11.44$, $p < .001$, indicating that subjects did respond more quickly to B− items than to members of the three other stimulus categories. Thus, we replicated the findings of Johnson and Mullen (1994) regarding the greater accessibility of B− items, but found this advantage was strong across all three presentation conditions.

We also investigated the latencies for items where there was correct group assignment. A mean was computed for each of the four stimulus categories on the basis of the latencies for behaviors that were correctly assigned. Data from subjects who did not respond correctly to one or more of the four stimulus categories were omitted from analyses.[5] As with the overall latencies, we conducted both the mixed-design ANOVA and the planned contrast. The Presentation Condition × Stimulus Category mixed-design ANOVA revealed only a very weak effect, $F(3, 207) = 1.83$, $p < .15$, for stimulus category. However, the key analysis is the planned comparison of B− items with the other three categories. As the means in Table 5 illustrate, this planned comparison revealed that subjects responded more quickly when correctly assigning B− items than the other three categories, $F(1, 207) = 3.66$, $p < .06$. These results replicated those of Johnson and Mullen, but show that the accessibility advantage for correct assignment of B− items is not limited to infrequency at time of presentation.

Finally, we analyzed the latencies for items with incorrect group assignments. The Condition × Stimulus Category mixed-design ANOVA did not detect any significant main or interactive effects. The planned comparison (contrasting B− items to the other three stimulus categories) also was not significant, $F(1, 198) = .32$, *ns.*

These results support the EDBE because B− information was more accessible than other information, and this was true regardless of the distinctiveness of B− items at the time of encoding. The general accessibility advantage of B− behaviors is not predicted by either the information-loss or the exemplar-based models of illusory correlation formation.

Moreover, in contrast to the implications of the standard DBE, we found a retrieval speed advantage for B− items in group assignment that was neither attenuated in the primacy condition nor enhanced in the recency condition. In other words, the greater accessibility of B− items was not contingent on infrequency at the time of encoding. Thus, we replicated the findings of Johnson and Mullen (1994) concerning the greater accessibility of B− items in group assignment, but found that infrequency at the time of encoding did not influence this enhanced representation.

Free Recall Analyses

Similar to Experiment 1, free recall was examined on the basis of general and correct recall. Two judges scored free recall using gist criterion (interjudge agreement was 93%). The decision of a third judge was used in cases of disagreement. All judges were unaware of the experimental hypotheses. Proportion recalled was analyzed in a 3 (conditions) × 2 (group, a within-subjects factor) × 2 (valence, also within subjects) mixed-design ANOVA. An effect of condition, $F(2, 87) = 4.43$, $p < .02$, revealed that subjects recalled more behaviors in the balanced condition (M = .31) than in either the primacy (M = .23) or recency (M = .24) conditions. The only other effect was a main effect of valence, $F(1, 87) = 30.30$, $p < .001$, indicating that subjects recalled proportionately more undesirable behaviors (M = .31) than desirable behaviors (M = .21). For correctly

[5] The elimination of subjects for correct latency analyses left 25 subjects in the balanced condition, 26 subjects in the primacy condition, and 21 subjects in the recency condition. For the incorrect latency analysis, there were 23 subjects in each of the three conditions.

recalled items, a 3 (conditions) × 2 (group) × 2 (valence) mixed-design ANOVA found only a main effect of valence, $F(1, 87) = 16.16$, $p < .001$, indicating that subjects correctly recalled proportionately more undesirable ($M = .15$) than desirable ($M = .10$) behaviors.

Finally, the relation between free recall and judgment was assessed by taking the difference of proportion of positive and negative behaviors recalled for each group and using them as predictor variables for likability ratings. Again, replicating the findings from Experiment 1, no evidence of memory-based judgments was found. When using the difference between proportion of positive and negative behaviors recalled for Group A, general free recall did not predict likability for Group A, $F(1, 88) = 0.00$, *ns*. Likewise, for Group B, general free recall (based on the Group B recall difference score) did not predict likability for the minority group, $F(1, 88) = 0.00$, *ns*. Using correct free recall, recall associated with Group A did not predict likability for Group A, $F(1, 88) = .22$, *ns*, and recall associated with Group B did not predict likability for Group B, $F(1, 88) = 1.44$, *ns*.

Discussion

Note that neither the information-loss explanation (Fiedler, 1991; Fiedler et al., 1993) nor the exemplar-based memory model (Smith, 1991) predict any accessibility advantage for B− items. The standard DBE would predict faster recall for B− items in the balanced and recency conditions (where they would be infrequent at the time of encounter), but no B− advantage in the primacy condition (where B− items were not distinctive at the time of exposure). Finally, the EDBE would predict faster recall for B− items in all three experimental conditions. The results of Experiment 2 supported the EDBE predictions.

Experiment 2 went beyond Experiment 1 in collecting process data by examining the latencies of subjects' group assignments. The pattern of findings in the balanced presentation condition replicated the results of Johnson and Mullen (1994) by showing a speed advantage for B− items in recall. More important, the strong and equal accessibility advantage for B− items across all three presentation conditions gave strong support for the EDBE, but was not predicted by either the information-loss hypothesis or the exemplar-based illusory correlation model. It is also important to note that the speed advantage observed for B− responses existed for correct assignments but not incorrect assignments. In other words, these faster responses showed no evidence of a response bias (i.e., quick indiscriminate assignment of Group B to all negative behaviors).

Taken together, these findings support a process of illusory correlation that is based on extra encoding for distinctive (B−) items either at the time of encounter when they are distinctive (as in the balanced and recency-loaded conditions), or with postpresentational processing as their ultimate distinctiveness becomes apparent (i.e., the primacy-loaded condition). Our conclusion of a process involving postexposure, but prejudgment, distinctiveness is based on the greater accessibility of B− items and on the strong illusory correlation observed in the primacy condition. If illusory correlations depend solely on the distinctiveness of B− items at the time of encoding, we would not have observed illusory correlations in the primacy condition. The evolution of distinctiveness in the primacy condition must have occurred subsequent to presentation, but before judgment, because B− items were highly accessible (indicated by faster group-assignment responses) at the time of judgment, even when B− items were not distinctive at the time of presentation. Thus, although distinctiveness at encoding is not a necessary ingredient in illusory correlation formation, enhanced encoding afforded to B− items (whether at the time of item presentation or later) does appear to be an important part of the illusory correlation process. The EDBE seems well supported on the basis of the findings of Experiments 1 and 2. Apparently, subjects continue to process, review, and assimilate information about the groups well after initial exposure. As Hunt and McDaniel (1993) have suggested, distinctiveness need not rely on infrequency at the time of presentation.

We have thus argued that the ultimate infrequency of the B− items is the key ingredient and that the process involved in the primacy-loaded condition is a postencoding process. We contend that items can become unusual and distinctive even though they are not perceived as such when originally encountered. However, an alternative possibility exists that does not involve a postencoding process but would predict elaborative encoding of the B− items even in the primacy-loaded condition. In the primacy-loaded condition—the only case where B− items are not distinctive at the time of presentation—not only do the B− items ultimately gain infrequency, but they also are massed in their presentation. That is, all of the B− items appear within the first 16 presentation items, and no other stimulus category was presented in such a massed fashion. Perhaps items presented in a massed fashion, regardless of current or ultimate distinctiveness, are encoded strongly. The question thus remains as to whether it is the ultimate infrequency of the B− items or the massed nature of their presentation that is responsible for the enhanced encoding of these items, their overestimations in frequency judgments, and ultimately, the illusory correlation that was observed. In Experiments 1 and 2, two of the three presentation conditions (primacy and recency-loaded) confounded overall infrequency with massed presentation. To separate these possibilities, Experiments 3–5 were conducted to determine whether the illusory correlations observed in these conditions (especially the primacy-loaded condition, where infrequency at the time of encounter was not a factor) were the product of the B− category's ultimate infrequency or its massed nature in presentation.

Experiments 3–5

People are aware of the frequency of events and process such information automatically without being directed to do so (Jonides & Jones, 1992; Manis, Shedler, Jonides, & Nelson, 1991; Whitlow & Skaar, 1979). Although this research is based on frequency estimates for particular exemplars, additional research has investigated the nature of frequency estimates for general categories as well. In fact, subjects have been shown to encode the frequency of general categories very well (Alba, Chromiak, Hasher, & Attig, 1980; Barsalou & Ross, 1986). Another consistent finding in the frequency estimation literature is that subjects overestimate the occurrence of items from infrequently presented information categories (relative to items from frequently presented categories; Alba et al., 1980; Barsalou & Ross, 1986). This finding is similar to the frequency estimate reports of subjects in illusory correlation experiments, in which subjects overestimate the occurrence of B− items relative to other stimulus classes.

Although category infrequency has been shown to lead to overestimation of member items of that category, less is known about whether massed presentation also leads to category member overestimation. Alba et al. (1980) reported that massed presentation (presenting all of the items of a stimulus category consecutively) did not lead to increases in frequency estimation. However, it should be noted that Alba et al. presented all of their information categories in a massed fashion. In contrast, in the primacy-loaded condition of Experiments 1 and 2, only one category (the B− category) was presented in a massed fashion. On the other hand, Greene (1989) has reported that categorical frequency estimates were higher for grouped than distributed categories. Greene's categories consisted of only 2, 3, or 4 exemplars, and half the categories were grouped and half were distributed. Therefore, it remains unclear whether ultimate infrequency, massed presentation, or both, is responsible for the overestimations of B− items observed in the first two experiments in the primacy-loaded condition.

Experiment 3 examines the usefulness of a category presentation paradigm for demonstrating frequency estimation effects that parallel those observed in our illusory correlation studies. Experiment 4 allows the comparison of two equally infrequent classes (relative to two additional categories that occurred more frequently) when only one is presented in a massed fashion. Comparisons between these two infrequent categories allow us to evaluate the effect of massed presentation, while keeping the ultimate infrequency between the two constant. Finally, Experiment 5 equates all four information categories in terms of frequency, but presents only one category in a massed fashion. This allows us to compare frequency estimates for categories of equal frequency that differ only in terms of massed versus distributed presentation.

In all three experiments, four stimulus categories (randomly selected as A, B, C, and D) were used to represent the four stimulus categories presented in Experiments 1 and 2 (i.e., A+, A−, B+, and B−). Of key interest is Category D. Category D, like the B− category in the first two experiments, was an infrequently presented category (with the exception of Experiment 5, where all four categories were presented with equal frequency) and was presented in either a massed (comparable to the primacy and recency conditions of Experiments 1 and 2) or distributed (comparable to the balanced conditions of Experiments 1 and 2) manner.

Method

Stimulus Materials

Experiments 3–5 used the same stimulus materials. Examples from four common and equally familiar categories (animals, clothing, occupations, and states) were chosen, based on Battig and Montague (1969). Examples of these four categories were selected such that the four categories did not differ in terms of frequency of usage in the English language (Francis & Kucera, 1982) or subjective familiarity for subjects in the studied population.[6]

Experimental Conditions

In each of the three experiments, 36 items were presented in accordance with the configurations presented in Table 6. Each experiment featured three experimental presentation conditions (distributed, massed early, and massed late), and the four information categories (A, B, C, and D) were arranged on the basis of the serial presentation orders shown in Table 6. In all three experiments, presentation of the key category, D, was configured in such a fashion that it would either be distributed (in a manner comparable with the balanced conditions of Experiments 1 and 2), massed early (the primacy condition), or massed late (the recency condition) in presentation. The individual experiments differed from each other in terms of how infrequent Category D was relative to the other three categories.

As Table 6 indicates, Experiment 3 was designed to parallel Experiments 1 and 2 using the current paradigm. There were 16 members of Category A (isomorphic to A+ in Experiments 1 and 2), 8 members each of Categories B and C (similar to A− and B+ in Experiments 1 and 2), and 4 members of Category D (similar to the B− category in Experiments 1 and 2). Category D examples were either consistently infrequent at time of encounter (distributed condition), not infrequent at time of encounter (massed early condition), or especially infrequent at time of encounter (massed late condition). If subjects overestimated D items (relative to the other categories) in a frequency-estimation task, results would parallel the findings of Experiments 1 and 2, where subjects overestimated (relative to the other categories) the occurrence of the B− category. This pattern of results would indicate that this paradigm leads to results that parallel illusory correlation results in terms of frequency estimates.

Experiment 4 (see Table 6) featured two infrequent categories (relative to two other more frequently presented categories) that did not vary in terms of ultimate infrequency, but differed in terms of infrequency at the time of encounter. There were 12 members of Categories A and B, and 6 members of Categories C and D. Categories C and D differed from Categories A and B in terms of ultimate infrequency. However, Categories C and D also differed from each other. Category C items were always distributed, whereas Category D items were either distributed (distributed condition) or massed at time of encounter (massed early or massed late conditions). If massed presentations (controlling for ultimate infrequency) can produce overestimations, we would expect that Category D, relative to Category C, would be overestimated in the massed conditions. However, if overestimation depends on the ultimate infrequency of behaviors rather than the massed nature of presentation (as proposed by the EDBE), subjects should estimate that Categories C and D do not differ in terms of presentation frequency, but should overestimate the relative occurrence of both Categories C and D with respect to Categories A and B.

Finally, Experiment 5 (see Table 6) explored the effects of presenting Category D in either a massed (early and late conditions) or distributed (distributed condition) fashion when all four categories possessed the same frequency of occurrence (nine examples). In this experiment, the ultimate frequency for all four stimulus categories was the same. The EDBE, which is based on ultimate infrequency rather than massed presentation, would predict no difference in frequency estimates between Category D and the other three categories across the three presentation conditions. However, if massed presentation leads to overestimation, one would expect subjects to report Category D as occurring more frequently than the other three categories in the massed early and massed late conditions (where Category D was presented in a massed fashion).

[6] We thank David Pisoni for making these norms available to us. Familiarity ratings were based on data collected from undergraduate students at Indiana University.

Table 6
Serial Presentation Configurations for Experiments 3–5

Condition	Partition no. and values								
	Experiment 3—the 16-8-8-4 design								
	1	2	3	4	5	6	7	8	9
Distributed	Rnd	1, 1, 1, 1	Rnd	1, 1, 1, 1	Rnd	1, 1, 1, 1	Rnd	1, 1, 1, 1	Rnd
Massed early	1, 1, 1, 1	1, 1, 1, 1	1, 1, 1, 1	1, 1, 1, 1	Rnd	Rnd	Rnd	Rnd	Rnd
Massed late	Rnd	Rnd	Rnd	Rnd	Rnd	1, 1, 1, 1	1, 1, 1, 1	1, 1, 1, 1	1, 1, 1, 1
Items	1–4	5–8	9–12	13–16	17–20	21–24	25–28	29–32	33–36
	Experiment 4—the 12-12-6-6 design								
	1	2	3						
Distributed	4, 4, 2, 2	4, 4, 2, 2	4, 4, 2, 2						
Massed early	2, 2, 2, 6	5, 5, 2, 0	5, 5, 2, 0						
Massed late	5, 5, 2, 0	5, 5, 2, 0	2, 2, 2, 6						
Items	1–12	13–24	25–36						
	Experiment 5—the 9-9-9-9 design								
	1	2	3						
Distributed	3, 3, 3, 3	3, 3, 3, 3	3, 3, 3, 3						
Massed early	1, 1, 1, 9	4, 4, 4, 0	4, 4, 4, 0						
Massed late	4, 4, 4, 0	4, 4, 4, 0	1, 1, 1, 9						
Items	1–12	13–24	25–36						

Note. Rnd = random assignment of members from Categories A, B, and C. Numbers offset by commas (e.g., *a*, *b*, *c*, *d*) indicate *a* members from Category A, *b* members from Category B, *c* members from Category C, and *d* members from Category D.

Procedure

For each experiment, the four stimulus categories (animals, clothing, occupations, and states) were randomly selected to serve as Categories A, B, C, and D. Once assigned, members of each category were randomly selected to serve as examples on the basis of the number required for that category's presentation. Within each presentation partition (Table 6), the presentation order of individual members of the relevant categories was randomly determined.

Subjects were tested at individual computer workstations located in individual rooms, and they were randomly assigned to conditions (distributed, massed early, and massed late). For all three experiments, subjects were told, "You are participating in an experiment about how people process and retain information. A series of words will be presented here on the monitor. Please read each word as it appears on the screen." Instructions were chosen to parallel the typical "process and retain information" instructions given in the illusory correlation paradigm (e.g., Hamilton & Gifford, 1976).

After receiving instructions, subjects were presented with the 36 examples on the computer monitor at a pace of one item every 4 s. After the presentation, subjects were given a 4-min filler task to eliminate short-term memory effects.

After the filler task, subjects were asked to provide frequency estimates for the four categories. The four categories were presented in a random order, and subjects were asked to estimate "how many examples of this group you think were presented." Subjects were given 15 s to enter their response on the keyboard. Subjects who did not respond in the 15-s period were omitted from the analyses.

Results

Experiment 3

Seventy-five Indiana University undergraduates participated in return for research experience credit and were randomly assigned (25 per condition) to the three experimental conditions. Three subjects (1 in each condition) failed to respond to frequency estimates within the 15-s limit, and their data were omitted from the analyses.

Frequency estimates were examined in two ways. First, we analyzed the actual frequency estimates in a 3 (distributed, massed early, and massed late conditions) × 4 (Categories A, B, C, and D, a within-subjects factor) mixed-design ANOVA to see if subjects were aware of increases in presentations for the four information categories. A main effect for categories, $F(3, 207) = 65.35$, $p < .001$, indicated that subjects were indeed sensitive to changes in frequency of presentation within each category. Subjects saw Category A ($M = 10.42$) as occurring more frequently than either Category B or Category C ($M = 6.56$ and $M = 6.32$, respectively), which were perceived as occurring more frequently than Category D ($M = 4.26$). This finding replicates those of Alba et al. (1980) and Barsalou and Ross (1986) concerning subjects' sensitivity to increases in category size. Experimental condition, however, had no main or interactive effects.

In addition to assessing frequency sensitivity, a proportion was computed based on the subjects' estimated category frequency divided by the actual category frequency. Thus, a ratio of 1.0 would reflect an accurate estimate for a particular stimulus category. Another 3 (conditions) × 4 (categories, a repeated measure) mixed-design ANOVA was conducted on these ratios. Consistent with Alba et al. (1980), subjects overestimated the most infrequent category relative to the frequent categories. A main effect of categories, $F(3, 207) = 13.97, p < .001$, indicated that subjects overestimated Category D ($M = 1.07$) relative to Categories B and C ($M = 0.82$ and $M = 0.79$, respectively), which were overestimated relative to Category A ($M = 0.65$). This pattern of results mirrors the illusory correlation effects obtained in Experiments 1 and 2. Importantly, overestimations of Category D were constant across the distributed, massed early, and massed late presentation conditions. Thus, the overestimation of the infrequent category, as in Experiments 1 and 2, was unaffected by that category's distinctiveness at the time of encounter. These findings are consistent with Alba et al. (1980), who found that category blocking did not affect frequency estimates.

These results are interesting in extending our understanding of illusory correlation. In Experiments 1 and 2, as well as in all previous research concerning illusory correlation with group targets, the paradigm involves a 2 × 2 presentation of category information (Groups A and B; desirable and undesirable behaviors). In fact, accounts of illusory correlation stress the importance of the cooccurrence of distinctive events (the B− items). In Experiment 3 (as well as Experiment 4), the categories are not of a 2 × 2 nature, so that there are no cooccurrences of distinctive features. Rather, these are four independent categories. Still, the very same effects were observed in terms of the frequency estimations. Low frequency categories were relatively overestimated. Thus, the illusory correlation effect (which involves perceptions of the degree of association between features) appears to be part of a more general effect involving biases in the perception of low-frequency categories.

As indicated earlier, overestimations of Category D in the massed early condition could be the product of either of two of the stimulus category's properties: (a) massed presentation or (b) ultimate infrequency. To differentiate between these two competing explanations, Experiment 4 was conducted to provide Category D with a comparison category (C) that differed in terms of massed presentation, but not ultimate infrequency. In this experiment, Categories C and D possessed the same ultimate infrequency. However, Category D was presented in a massed fashion (either early or late), whereas Category C was always presented in a distributed manner. The EDBE, unlike the massed presentation explanation, predicts that subjects will estimate similar presentation frequencies for Categories C and D. However, if massed presentation is an important factor in the overestimation of categories, we would expect subjects to overestimate the occurrence of Category D (relative to Category C) in the two massed conditions.

Experiment 4

Seventy-two Indiana University undergraduates participated in return for research experience credit and were randomly assigned (24 per condition) to the three experimental conditions. All subjects responded to frequency estimates within the 15-s limit.

A 3 (presentation conditions) × 4 (categories, a within-subjects factor) mixed-design ANOVA was conducted on the raw frequency estimates offered by subjects. Frequency estimates illustrated the same pattern as Experiment 3. Again, in line with the findings of Alba et al. (1980) and Barsalou and Ross (1986), the main effect for categories, $F(3, 207) = 26.18, p < .001$, indicated that subjects reported that Categories A and B occurred more frequently ($M = 8.90$ and $M = 8.89$, respectively) than Categories C and D ($M = 5.79$ and $M = 5.62$, respectively). Experimental condition did not influence these estimates. Also, the ratio of estimated-to-actual was analyzed in a Conditions × Categories mixed-design ANOVA. Once again, subjects overestimated the occurrence of the infrequent categories ($M = .97$ for Category C, $M = .94$ for Category D) relative to the frequent categories ($M = .74$ for Category A, $M = .74$ for Category B), $F(3, 207) = 7.75, p < .001$. Again, this finding replicates those of Alba et al. and shows the same pattern found in Experiment 3 with Category D. Furthermore, these results support the EDBE rather than the massed presentation explanation because frequency estimates for Categories C and D did not differ, even though Category D was presented in a massed fashion.

To provide a more extreme test of the EDBE and the massed presentation explanations, Experiment 5 was conducted using a 9-9-9-9 presentation scheme (Table 6). In this situation, frequency was controlled across all four stimulus classes, and thus, the only difference between Category D and the other three categories was in the massed or distributed nature of presentation. With ultimate frequency of the categories equivalent, the massed presentation explanation would predict that subjects should overestimate the frequency of Category D items in the massed conditions, but not in the distributed condition. On the other hand, the EDBE would predict no difference in the frequency estimates of the categories across any of the three presentation conditions.

Experiment 5

Seventy-eight Indiana University undergraduates participated in return for research-experience credit and were randomly assigned (26 per condition) to the three experimental conditions. All subjects responded to frequency estimates within the 15-s limit.

Frequency estimates were examined in a 3 (presentation conditions) × 4 (categories, a within-subjects factor) mixed-design ANOVA. No significant effects were obtained. Regardless of presentation scheme, subjects estimated the same actual frequency for the four categories ($M = 7.56$, $M = 7.46$, $M = 7.54$, and $M = 7.94$ for Categories A, B, C, and D, respectively). Because all four stimulus categories possessed the same frequency, no analysis was conducted on the estimated-to-actual ratios.

Thus, the findings of Experiments 3–5 provide support for the EDBE but are inconsistent with the massed presentation hypothesis. These results suggest that the ultimate infrequency of stimulus categories, and not massed presentation, leads to the overestimation of infrequent categories. This reinforces the conclusion from Experiments 1 and 2 that illusory correlation can be based on a postencoding mechanism and that the ultimate distinctiveness of a stimulus class determines its extent of encoding and its accessibility.

General Discussion

Illusory correlation with group targets is a robust and well-documented finding. It is based on overestimations of distinctive behaviors by distinctive groups and can contribute to the formation of group stereotypes. Two groups about whom identical evaluative information is presented are perceived and evaluated differently as a result of the illusory correlation process.

The predominant explanation for illusory correlation has been the DBE. This theory proposes that infrequent categories of behavior (e.g., the B− items in Experiments 1 and 2), because of their distinctiveness, receive more extensive encoding than other categories and therefore are more accessible in memory. This enhanced accessibility leads to greater weight given to this category when making group judgments as well as to errors in frequency estimates and biases in recall.

The DBE is a theory that clearly involves processes that occur during the encoding of stimuli. A key component of the standard DBE is the distinctiveness of items at the time of encoding. On the basis of the experimental conditions of Experiments 1 and 2, the standard DBE would predict an attenuated illusory correlation in the primacy-loaded condition (where B− items are not distinctive at the time of encounter) and an enhanced illusory correlation in the recency-loaded condition (where B− items are especially distinctive at the time of encounter). In other words, the size of the illusory correlation should depend on the distinctiveness of the B− items at the time of presentation.

However, other research suggests the possibility that distinctiveness of the B− items at the time of their presentation is not a necessary precondition for illusory correlation formation. As Hunt and McDaniel (1993) pointed out, enhanced memory for distinctive items had been observed by von Restorff even when the distinctiveness of these items did not become apparent in the stimulus set until long after their presentation. According to this possibility, illusory correlation formation could involve postencoding processes, where the key element was the ultimate distinctiveness of B− items rather than their distinctiveness at the time of presentation. In this case, strong illusory correlations would be observed in all three presentation conditions: balanced, primacy-loaded, and recency-loaded.

The results of Experiment 1 provided evidence for the involvement of postencoding processes. The levels of illusory correlation were strong and constant across the three experimental conditions. Regardless of the sequencing of the behaviors in the list and regardless of the distinctiveness of the B− items at the time of presentation, the illusory correlation was equal in its strength. The primacy-loaded condition results are especially troublesome for the standard DBE. At the time of presentation in this condition, the B− items were not at all distinctive, and yet a strong illusory correlation was obtained. It thus appears that the ultimate difference in frequency between the B− items and the other stimulus classes is important, not simply the difference in frequency at the time of encoding. These findings left us with the task of accounting for these results while at the same time maintaining an explanation compatible with previous process data that supported the standard DBE (Hamilton et al., 1985; Johnson & Mullen, 1994; Stroessner et al., 1992).

Two possibilities appeared viable. The first approach was represented by a set of explanations that did not rely on item distinctiveness as the basis for the illusory correlation effect. Both Fiedler's (1991) information-loss approach and Smith's (1991) exemplar-based model depend on the overall differences in frequencies among the stimulus categories rather than on item distinctiveness at the time of presentation. Both explanations propose processes that occur at the time of either retrieval or judgment. Thus, these interpretations are generally compatible with the findings of Experiment 1. More important, these explanations do not predict differences in the processing of B− items as opposed to the other types of items. Biases in attention, encoding, or retrieval of the item types play no role in these explanations. The second possibility, which we term the *EDBE*, does propose a role for the distinctiveness of the B− items and does predict that these items will be processed differently from the other items. Thus, this proposal is consistent with previously reported process data that indicate biases in attention to and encoding of B− items (Hamilton et al., 1985; Johnson & Mullen, 1994; Stroessner et al., 1992). However, the EDBE further develops the standard DBE in suggesting that items can be distinctive (and processed differently) not only at the time of encounter, but also when their distinctiveness becomes apparent subsequent to their initial presentation, as additional information is presented and processed.

These two possibilities were examined in Experiment 2. Because the Fiedler and Smith models do not predict additional processing for B− behaviors, neither theory would predict an advantage for B− items in group-assignment response latencies. In contrast, the EDBE predicts that B− behaviors should be recalled quickly in all three conditions. The latency results indicated that the B− items were indeed more accessible, suggesting that they had indeed been processed more thoroughly than the other stimulus types. This result occurred even in the primacy-loaded condition where the B− items were not at all distinctive at the time they were presented. Thus, these results provide support for the EDBE and are problematic for Fiedler's (1991) and Smith's (1991) explanations. Also, contrary to the information-loss model, free recall from the first two experiments did not indicate any evidence of superior memory for Group A, as predicted by Fiedler (1991). Rather, free recall data from Experiment 1 actually showed better recall for Group B than Group A.

Finally, Experiments 3–5 were conducted to test an interpretation of the illusory correlation and latency results of Experiments 1 and 2 that did not depend on postencoding processes. These studies did support the involvement of postencoding processes in that the ultimate infrequency of items, rather than the massed nature of their presentation, led to biases in frequency estimates.

We thus propose a theory of illusory correlation formation with group targets that is based on subjects' processing the behavioral information about group members in a continuous manner. Items of information are not simply read, stored, and left unaltered in memory as new items are considered. Rather, as each new piece of information is processed, old items of information may be reconsidered, reviewed, and assimilated to the new information. Through this process, information that was not distinctive at the time of encounter (as in our primacy-loaded condition) may become distinctive after the fact, as new information is acquired. This now-distinctive information can receive additional processing and consideration, long after it was first received. Thus, as new items are encountered, they are considered in relation to old information. When this happens, group representations and the representations of the individual behaviors will thus reflect the entire stream of behaviors, regardless of how distinctive any category of items may have been at the time of presentation.

This idea that old items of information can be reconsidered and rerepresented as other information is presented is not unique. Evidence of postencoding information processing has been established in a variety of areas. Much of this work, however, involves memory processes rather than judgment processes (e.g., Hunt & McDaniel's discussion of von Restorff effects, 1993). Loftus's (1979; Loftus, Miller, & Burns, 1978) work on eyewitness testimony demonstrates how subsequently received misleading information can alter one's original representation of an event. Similarly, Snyder and Uranowitz (1978) showed how new information about a target person can affect the interpretation of previously presented information. In other words, old information is reconsidered and can change its meaning in the light of new information. Further evidence of memory reconstruction processes was provided by Hirt (1990; Hirt, Erickson, & McDonald, 1993) and Conway and Ross (1984), who pointed to many situations where new circumstances and facts lead to a reconsideration of old information and a reconstruction in the representation of that old information. New perspectives given subsequent to information presentation can also change the memory for and representation of previous facts. Hasher and Griffin (1978) showed how a new perspective on a story that is given after the facts have been received can alter the amount, the content, and the accuracy of information recalled (see also Anderson & Pichert, 1978; Dooling & Christiaansen, 1977). The present results fit nicely into this literature on postencoding changes in information representation. Our findings, like these others, suggest a dynamic organization of interpretive processes, where past experiences are accessed, reconsidered, and reprocessed in the light of new information and experiences.

Although a postencoding information processing explanation is consistent with the current set of findings, postencoding processing is neither a universally accepted nor a universally supported point of view. McCloskey and Zaragoza (1985) disagreed with Loftus et al. (1978) that misleading questions lead to changes in an earlier memory structure. Rather they provided evidence that new traces are laid down and the old traces are maintained intact in memory rather than being updated. Bransford and Johnson (1972) found that presenting a thematic context after a passage of text had been read did not enhance memory or change the representation of the already encoded information. Similar effects have been reported by Zadny and Gerard (1974). Also, Srull and Wyer (1980) found that postinformation priming of a trait concept did not affect the interpretation of an actor's attributes although preinformation priming had a large and significant effect (see Neely, 1991, however, for a discussion of evidence supporting backward priming).

It is clear that no definitive conclusions can yet be drawn about the generality and importance of postencoding processing of information. Sometimes previously acquired information is reconsidered, reviewed, and altered in the context or extensiveness of its representation as new information is received. At other times, initial memory representations seem impervious to change in response to subsequent information. The current findings simply add one more instance where after-the-fact review of information occurs and alters its representation. It will take a good deal of subsequent research before we understand exactly when, how, and under what conditions these postencoding processes are manifested.

There is yet another class of models that must be considered in the light of our results. These models concern themselves with the memory for and latency of response to items from categories of different sizes. Set size effects or fan effects refer to the finding that the more facts that are associated with a concept, or the larger the category, the poorer the overall recall for those facts and the longer it takes to recognize those facts (Anderson, 1976; Lewis & Anderson, 1976; Sternberg, 1966). The typical explanation for such effects is a spreading activation process such as Anderson's (1976) ACT model. According to this model, source nodes have a fixed capacity for emitting activation. The more paths from a node, the less the activation of any node and the slower the spread of activation. Thus, as set size increases, memory for any concept will diminish and latency for retrieval of or verification of any item will increase (see also, Sternberg, 1966, for a serial exhaustive scanning model with similar predictions). Such effects occur at the time of judgment and are thus dependent only on ultimate frequency differences and not on any processing or reprocessing differences for items from small set sizes. Consistent with these set-size theories, we have reported that estimates of frequencies from smaller set sizes are larger relative to their actual sizes than are estimates of frequencies from larger set sizes.

Although set size effects have not, to date, been used to account for the formation of illusory correlation for the paradigms used in our study, their potential applicability is clear. As B− is the smallest stimulus category, its items should be better remembered and responded to most quickly. These of course are some of the important results of Experiments 1 and 2 and of previous studies using a similar paradigm. Such differential memory based on set size could mediate the differential judgments of Groups A and B. Despite the applicability of a set size explanation of our results (and previous illusory correlation findings as well), we prefer an interpretation that involves the reprocessing of information that becomes distinctive after its presentation to an interpretation based solely on set size for several reasons. First, manipulations of distinctiveness other than set size have been found to produce illusory correlations. Spears, van der Pligt, and Eiser (1986) showed that self-relevance of information could serve as the basis of illusory correlation, independently of differential frequency. Also, Sanbonmatsu, Sherman, and Hamilton (1987) manipulated distinctiveness by attentional instructions rather than infrequency and replicated typical illusory correlation effects. Thus, any factor that calls special attention to a particular group or behavior works to produce differential judgments of groups, and different category sizes are not necessary for these effects. Given these results, an explanation that depends on distinctiveness more generally, rather than set size specifically, appears more appropriate.

Second, and relatedly, the findings of von Restorff cited by Hunt and McDaniel (1993) refer to memory effects that depend on distinctiveness generally and not simply on set size. The fact that von Restorff effects depend on ultimate distinctiveness, and not simply on distinctiveness at encoding, is consistent with our findings. In accounting for these memory effects, Hunt (1993) has proposed mechanisms similar to our reprocessing proposal, mechanisms that involve active processing of the information at the time that its distinctiveness becomes apparent.

At this point it is not possible to specify the exact nature of the reprocessing engaged in by our subjects. However, what is important at this point is that our results expand on the standard DBE and support a process of illusory correlation formation that includes postencoding mechanisms rather than being limited to effects that occur only during encoding. Whether this process involves reprocessing as subsequent items of information are received or whether the process is one that occurs later, either at retrieval or judgment, is not the central issue. Several candidate processes have been proposed, including our reprocessing account and set size effects. At this point, the current findings, in combination with the existing literature, seem to support an active reprocessing interpretation, but it will certainly take further work to decide conclusively among the rather similar postencoding possibilities.

The present results also speak to one other aspect of the DBE. According to the DBE, the judgments of Groups A and B are made in a memory-based fashion (Hastie & Park, 1986). That is, the evaluations of the groups are not formed on-line as the behavioral information is received, but rather are made on the basis of whatever information is accessed at the time of judgment. Direct evidence for the memory-based processing aspect of the DBE has been elusive. One important indicator of memory-based processing is the degree of correlation between memory and judgments. If judgments are made at the time they are requested, based on whatever information is accessed at that time, clearly judgment–recall correlations should be high. Hamilton et al. (1985) reported a significant correlation between the recall of information and judgments of group targets, but only for negative behaviors by Group B members. In the present study, we reported the correlations of liking ratings for Groups A and B with the evaluative content of subjects' free recall, and in no case were these correlations significant. Thus, no evidence was provided for a memory-based judgment of these group targets.

In a recent study, McConnell, Sherman, and Hamilton (1994) explored the on-line versus memory-based aspects of judgments of individuals and groups within an illusory correlation paradigm. Judgments were clearly more likely to be formed on-line for individual targets, as suggested by greater recall for individual target information, by a primacy effect for recall of individual information, and by shorter group-assignment latencies for individual target information. On the other hand, impressions of group targets were not memory-based in the strong sense of that process. Under memory-set instructions (similar to those used in the present studies), judgment–recall correlations were not significant. Only when group targets were used and when the instructions interfered with on-line processing (by asking subjects to judge the comprehensibility of the sentences describing the behaviors) were significant judgment–recall correlations observed.

Thus, it is likely that some degree of on-line impression formation occurs for group targets. This on-line impression formation may not result in a fully integrated and coherent impression, but may involve some partial extraction of trait information. Later, at the time of judgment, this information, rather than (or in addition to) the memory for specific items of behavioral information, could be used to render a judgment. In this case, judgment–recall correlations might be quite low, as was observed in the current studies. These considerations suggest that the distinction between on-line and memory-based judgments is not a dichotomy, but may better be thought of as a continuum. Impressions of group targets may not be formed in a complete, integrated way as information about group members is processed, but neither are group impressions based solely on the recall of specific items of behavioral information at the time of judgment.

In sum, the current studies offer further insight into the processes involved in group stereotyping in the illusory correlation paradigm. Although the importance of the distinctiveness of minority group–minority behavior items has been supported, the traditional DBE has been expanded. This expansion raises the possibility of distinctiveness developing through a postencoding process and also suggests a reconsideration of strong memory-based aspects of impressions of group targets.

References

Alba, J. W., Chromiak, W., Hasher, L., & Attig, M. S. (1980). Automatic encoding of categorize size information. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 370–378.

Anderson, J. R. (1976). *Language, memory, and thought.* Hillsdale, NJ: Erlbaum.

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior, 17,* 1–12.

Barsalou, L. W., & Ross, B. H. (1986). The roles of automatic and strategic processing in sensitivity to superordinate and property frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 116–134.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology, 80* (Monograph Supplement 3, Pt. 2).

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717–726.

Chapman, L. J. (1967). Illusory correlation in observational report. *Journal of Verbal Learning and Verbal Behavior, 6,* 151–155.

Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. *Journal of Personality and Social Psychology, 47,* 738–748.

Dooling, D. J., & Christiaansen, R. E. (1977). Episodic and semantic aspects of memory for prose. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 428–436.

Fiedler, K. (1991). The tricky nature of skewed frequency tables: An information loss account of distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology, 60,* 24–36.

Fiedler, K., Russer, S., & Gramm, K. (1993). Illusory correlations and memory performance. *Journal of Experimental Social Psychology, 29,* 111–136.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar.* Boston: Houghton Mifflin.

Greene, R. L. (1989). Negative practice effects on frequency discrimination. *American Journal of Psychology, 102,* 225–232.

Hamilton, D. L., Dugan, P. M., & Trolier, T. K. (1985). The formation of stereotypical beliefs: Further evidence for distinctiveness-based illusory correlations. *Journal of Personality and Social Psychology, 48,* 5–17.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12,* 392–407.

Hamilton, D. L., & Sherman, S. J. (1989). Illusory correlations: Implications for stereotype theory and research. In D. Bar-Tal, C. F. Graumann, A. W. Kruglanski, & W. Stroebe (Eds.), *Stereotyping and prejudice: Changing conceptions* (pp. 59–82). New York: Springer-Verlag.

Hasher, L., & Griffin, M. (1978). Reconstructive and reproductive processes in memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 318–330.

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

Hirt, E. R. (1990). Do I see only what I expect? Evidence for an expectancy-guided retrieval model. *Journal of Personality and Social Psychology, 58,* 937–951.

Hirt, E. R., Erickson, G. A., & McDonald, H. E. (1993). The role of expectancy timing and outcome consistency on expectancy-guided retrieval. *Journal of Personality and Social Psychology, 65,* 640–656.

Hunt, R. R. (1993, November). *What von Restorff really did: The subtlety of distinctiveness.* Paper presented at the annual meeting of the Psychonomic Society, Washington, DC.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language, 32,* 421–445.

Johnson, C., & Mullen, B. (1994). Evidence for accessibility of paired distinctiveness in distinctiveness-based illusory correlation in stereotyping. *Personality and Social Psychology Bulletin, 20,* 65–70.

Jonides, J., & Jones, C. M. (1992). Direct coding for frequency of occurrence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 368–378.

Koffka, K. (1935). *Principles of Gestalt psychology.* New York: Harcourt, Brace, and Company.

Lewis, C. H., & Anderson, J. R. (1976). Interference with real world knowledge. *Cognitive Psychology, 8,* 311–335.

Loftus, E. F. (1979). *Eyewitness testimony.* Cambridge, MA: Harvard University Press.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19–31.

Manis, M., Shedler, J., Jonides, J., & Nelson, T. (1991). *The availability heuristic in judgments of set-size and frequency of occurrence.* Unpublished manuscript. University of Michigan, Ann Arbor, Michigan.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114,* 1–16.

McConnell, A. R., Sherman, S. J., & Hamilton, D. L. (1994). On-line and memory-based aspects of individual and group target judgments. *Journal of Personality and Social Psychology, 67,* 173–185.

Mullen, B., & Johnson, C. (1990). Distinctiveness-based illusory correlations and stereotyping: A meta-analytic integration. *British Journal of Social Psychology, 29,* 11–28.

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphries (Eds.), *Basic processes in reading: Visual word recognition.* Hillsdale, NJ: Erlbaum.

Sanbonmatsu, D. M., Sherman, S. J., & Hamilton, D. L. (1987). Illusory correlation in the perception of individuals and groups. *Social Cognition, 5,* 1–25.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition, 19,* 523–542.

Smith, E. R. (1991). Illusory correlation in a simulated exemplar-based memory. *Journal of Experimental and Social Psychology, 27,* 107–123.

Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology, 36,* 941–950.

Spears, R., van der Plight, J., & Eiser, J. R. (1986). Generalizing the illusory correlation effect. *Journal of Personality and Social Psychology, 51,* 1127–1134.

Srull, T. K., & Wyer, R. S. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality and Social Psychology, 38,* 841–856.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652–654.

Stroessner, S. J., Hamilton, D. L., & Mackie, D. M. (1992). Affect and stereotyping: The effect of induced mood on distinctiveness-based illusory correlation. *Journal of Personality and Social Psychology, 62,* 564–576.

Whitlow, J. W., & Skaar, E. (1979). The role of numerosity in judgments of overall frequency. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 409–421.

Zadny, J., & Gerard, H. B. (1974). Attributed intentions and informational selectivity. *Journal of Experimental Social Psychology, 10,* 34–52.

Received August 18, 1993
Revision received March 7, 1994
Accepted March 8, 1994 ■