

UC Davis

UC Davis Previously Published Works

Title

VARIATIONS ON A THEME Attentional Processes in Group and Individual Perception

Permalink

<https://escholarship.org/uc/item/22f696gv>

ISBN

978-1-138-83790-4

Authors

Sherman, Jeffrey W

Huang, Lisa M

Sacchi, Dario LM

Publication Date

2015

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

- van Rijswijk, W., & Ellemers, N. (2002). Context effects on the application of stereotype content to multiple categorizable targets. *Personality and Social Psychology Bulletin*, 28, 90-101.
- Vonk, R. (1998). The slime effect: Suspicion and dislike of likeable behavior toward superiors. *Journal of Personality and Social Psychology*, 74, 849-864.
- Waggoner, A.S., & Smith, E.R. (2012). Information processing differences in active versus passive person perceivers. *Social Psychological and Personality Science*, 3, 208-215. doi:10.1177/1948550611415696
- Waggoner, A.S., Smith, E.R., & Collins, E.C. (2009). Person perception by active versus passive perceivers. *Journal of Experimental Social Psychology*, 45, 1028-1031.
- Watson, D., Hubbard, B., & Wiese, D. (2000). Self-other agreement in personality and affectivity: Effects of acquaintanceship, trait visibility, and assumed similarity. *Journal of Personality and Social Psychology*, 78, 546-558.
- Wegner, D.M. (1989). *White bears and other unwanted thoughts: Suppression, obsession, and the psychology of mental control*. London, England: Guilford Press.
- Wilder, D.A. (1984). Intergroup contact: The typical member and the exception to the rule. *Journal of Experimental Social Psychology*, 20, 177-194.
- Winter, L., & Uleman, J.S. (1984). When are social judgments made? Evidence for the spontaneity of trait inferences. *Journal of Personality and Social Psychology*, 47, 237-252. doi:10.1037/h0090437
- Wittenbrink, B., Judd, C.M., & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology*, 72, 262-274. doi:10.1037/0022-3514.72.2.262
- Wyer, R.S., & Carlston, D.E. (1979). *Social cognition, inference, and attribution*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Wyer, N.A., & Hamilton, D.L. (1998). The balance between excitation and inhibition in stereotype use. In R. Wyer (Ed.), *Stereotype activation and inhibition* (pp. 227-242). Mahwah, NJ: Lawrence Erlbaum Associates.
- Wyer, R.S., & Srull, T.K. (1989). *Memory and cognition in its social context*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Yzerbyt, V.Y., Leyens, J., & Corneille, O. (1998). Social judgeability and the bogus pipeline: The role of naive theories of judgment in impression formation. *Social Cognition*, 16, 56-77. doi:10.1521/soco.1998.16.1.56
- Yzerbyt, V.Y., Schadron, G., Leyens, J., & Rocher, S. (1994). Social judgeability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, 66, 48-55. doi:10.1037/0022-3514.66.1.48
- Zajonc, R.B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175. doi:10.1037/0003-066X.35.2.151
- Zebrowitz, L.A., & Montepare, J.M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, 2, 1497-1517. doi:10.1111/j.1751-9004.2008.00109.x
- Zuckerman, M., & Feldman, L.S. (1984). Actions and occurrences in attribution theory. *Journal of Personality and Social Psychology*, 46, 541-550. doi:10.1037/0022-3514.46.3.541

7

VARIATIONS ON A THEME

Attentional Processes in Group and Individual Perception

Jeffrey W. Sherman, Lisa M. Huang,
and Dario L. M. Sacchi

In the spring of 1989, I (Sherman) was deciding where to pursue my graduate studies. One of my options was the University of California, Santa Barbara, where I could work with Dave Hamilton, whose research I admired. In addition to offering an excellent advisor in Dave, UCSB offered a superb secondary advisor in Diane Mackie, with whom I might (and eventually did) also collaborate. On my visit there, I had quickly made friends with Steve Stroessner, my graduate student host, based on our shared affinity for The Replacements and Hüsker Dü, among other things. The beachfront campus didn't suck. Given all this, the decision was easy, and I eagerly signed on.

Despite all of these wonderful features of UCSB, at the time, it was probably a really bad decision. You see, at that time, my father and Dave were in the midst of a three decade-long collaboration that produced many significant ideas and publications, not to mention many years of grant support. It didn't seem to occur to anyone that working with my father's closest collaborator might raise questions about nepotism and doubts about the nature of my contributions to my research with Dave. It didn't seem to occur to anyone that this would contribute to years of being called "Jimmy Sherman's little boy" or "Little Jeffrey Sherman." I'm not complaining. I loved my time at UCSB, and my career has turned out well enough. Still, I have to wonder what the adults were thinking; I was certainly too naïve to realize the potential problems.

We did, however, get one thing right: My collaborations with Dave would never involve work on the *distinctiveness-based illusory correlation* (Hamilton & Gifford, 1976), which was the main topic of his collaboration with my dad at the time. I can't remember any explicit discussion about this choice; thankfully, it just seemed to naturally flow that way. As time passed, I began to take ironic pride in the fact that, despite my connections, I didn't really know much about

illusory correlation or the never-ending debates about what processes produce it. In fact, I actively avoided trying to learn too much about it. One's identity can only handle so much assimilation. Nevertheless, given the associations between me and Dave and between me and my dad, it's not surprising that I would frequently be sent papers on illusory correlation to review when I was an assistant professor. Sometimes, I would pointedly turn them down, suggesting that they be sent to the senior Sherman—the one who actually knew about this stuff. Other times, I would take the assignment and end up calling my dad: What is the stupid phi coefficient, again?

This chapter tells the story of how, despite my best efforts, I was eventually hunted down and subdued by the illusory correlation and its charms. I could not escape my fate. However, in another twist of irony, it was my dad and not Dave with whom I have explored the illusory correlation. By the time this exploration began, I had achieved tenure and was no longer as concerned about the appearances of working with my dad. Doing research with him has proved to be a highlight of my career.

Still, I had to get past my principled aversion to the phenomenon. This process was aided greatly by examining it from a completely novel perspective suggested by John Kruschke, a cognitive psychologist in my dad's department at Indiana University. I knew John a little from hanging around the department at Indiana when I was in town, and it was a chance encounter at a conference that got the ball rolling on the project. In fact, by the time I ran into John at that conference, he had already floated the idea to my dad of applying his own work on classification learning to the illusory correlation. For whatever reason, nothing had come of that effort, and so John attempted to interest a different Sherman in this idea, despite my protestations that I neither knew nor cared much about illusory correlation. Nevertheless, by the end of our conversation, we had outlined the first few studies for our new line of research. The first part of this chapter will describe those efforts. Then, we will describe how this idea has been expanded to understanding the perception of the physical and psychological features of individuals, both across groups and contexts.

The Attention Theory of Classification Learning

Attention Theory (AT) was developed, in part, to account for the *inverse base rate effect* in human learning (Kruschke, 1996, 2001, 2003). In the original demonstration of the effect (Medin & Edelson, 1988), participants were asked to diagnose different diseases from patterns of symptoms. On each trial of the learning sequence, a list of symptoms was presented, and participants were asked to diagnose the hypothetical patient as having one of several possible fictitious diseases. After each trial, participants were told the correct diagnosis. The basic design involved a pair of diseases, designated C (for common) and R (for rare),

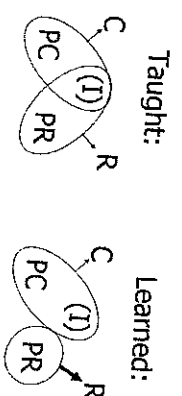


FIGURE 7.1 (left) Core design of the inverse base rate effect. C and R denote common and rare diseases, respectively. PC is a perfect predictor of the common disease; PR is a perfect predictor of the rare disease. I is an imperfect predictor. (right) What is learned according to AT.

which occurred with a 3:1 ratio. During training, every instance of disease C had two symptoms, labeled I (e.g., headache) and PC (e.g., fever), and every instance of disease R had two symptoms, labeled I and PR (e.g., stomach ache). PC and PR were perfect predictors of diseases C and R—PC always predicted C and never R; PR always predicted R and never C. Symptom I was an imperfect predictor of the two diseases, in that all cases of both C and R were associated with that symptom. Following training, participants were tested with combinations of symptoms not shown during training. When tested with ambiguous symptom I (headache) alone, people tended to choose the *common* disease, consistent with the base rates (during training, I appeared with C 75% of the time). However, when presented with the conflicting symptoms PC + PR (fever + stomach ache), participants tended to choose the *rare* disease, contrary (or inverse) to base rates.

AT explains the effect as follows: During training, people first learn that symptoms I and PC are typical of disease C because that case occurs with high frequency. Subsequently, when learning about rare disease R, they realize that the shared symptom I is a misleading predictor, because it already is associated with disease C. As a result, attention shifts away from I, and is focused on the distinct symptom of R, PR. As a result, when learning about disease R, attention is focused primarily on a single, distinctive symptom (PR), whereas, when learning about disease C, attention is divided between the symptoms PC and I. More generally, the theory suggests that, as categories develop, greater attention is devoted to features that distinguish new categories from old ones than is devoted to features that define old ones. For this reason, PR becomes more strongly associated with disease R than PC is associated with disease C, thereby producing the inverse base rate effect (see Figure 7.1).

Attention Theory and Stereotype Formation

There's work on illusory correlation and Tajfel's research (Tajfel & Wilkes, 1963) on category accentuation represent the two seminal approaches to understanding the social cognition of stereotype formation. Whereas illusory correlation shows

that stereotypes may be formed in the absence of real group differences, category accentuation effects highlight the exaggeration of real inter-group differences as the basis for stereotype formation. Research on the two effects has proceeded independently, and they have been explained by different mechanisms. However, when we started working with AT, we realized that it could explain both illusory correlation and category accentuation with a single mechanism.

Work on category accentuation shows that the division of people into separate categories both exaggerates perceived differences between categories and similarities within categories (e.g., Corneille & Judd, 1999; Eiser, 1971; Krueger & Clement, 1994; Krueger & Rothbart, 1990; McGarty & Penny, 1988; McGarty & Turner, 1992; Queller, Schell, & Mason, 2006), contributing to the development of group stereotypes. Indeed, research shows that stereotypes are most likely to be formed around attributes for which inter-group differences are large and intragroup differences are small (e.g., Ford & Stangor, 1992). A variety of mechanisms have been shown to contribute to these accentuation effects. First, as in Tajfel's original research, perceptions of individual category members may be biased by category boundaries. Second, category members that heighten between-category differences and within-category similarities may be attended to more carefully, given greater weight in judgments, and remembered more easily (Krueger & Rothbart, 1990; Krueger, Rothbart, & Strim, 1989).

Note that these processes of category accentuation are similar to the attention shifting processes described by AT; in both cases, people focus on the features and exemplars that most effectively distinguish one category from another category. However, one important difference between AT and category accentuation models is that AT does not require that there be real differences between the categories in question. Two categories may be described identically, but still produce differentiated representations as long as one category is learned before the other. Indeed, a factor that causes one category to be learned prior to another (e.g., frequency of exposure, group size) will lead to different and accentuated impressions of the categories (e.g., group stereotypes). The first category will be associated with its most common attributes, and impressions of the second category will form around those features that most clearly differentiate it from the first category.

These same processes also may produce the illusory correlation, in which observers perceive an association between distinctive groups and distinctive behaviors when, in fact, no such relationship exists (Hamilton & Gifford, 1976; for a review, see Stroessner & Plaks, 2001). In the modal demonstration of the effect, two groups (Group A and Group B) are described by a series of positive and negative behaviors. There are twice as many members of Group A than Group B, and there are more positive than negative behaviors. For example, in the original experiment, members of Group A performed 18 positive and 1 negative behaviors, whereas members of Group B performed 9 positive and 1 negative behaviors. Thus, Group B and negative behaviors both are distinctive to their infrequency (see Figure 7.2).

	Desirable Behaviors	Undesirable Behaviors
Group A	18	8
Group B	9	4

FIGURE 7.2 A standard configuration of presented information in a distinctive feature based illusory correlation paradigm.

Although Group A performs more behaviors than Group B, the ratio of positive to negative behaviors is the same in both groups. Consequently, there is no relationship between group membership and desirability. Nevertheless, participants perceive Group A to be more favorable than Group B. This effect is reflected in a number of findings. First, participants misattribute Group A's positive behaviors to Group A, and misattribute Group A's negative behaviors to Group B. Second, participants overestimate the numbers of positive behaviors performed by Group A and negative behaviors performed by Group B. Finally, participants rate Group A more favorably than Group B on trait rating tasks.¹

These effects have drawn intense interest because—unlike the exaggerations of category features described by Tajfel—they demonstrate the formation of differential group impressions in the absence of real group differences. In addition, there is a direct parallel between the illusory correlation and the perceptions of majority and minority social groups outside of the lab. As in the illusory correlation, members of minority groups are numerically distinct, as are negative behaviors (e.g., criminal acts). Thus, the illusory correlation may serve as a model of how minority groups come to be viewed more negatively than majority groups in societies, even if their behavior does not warrant such evaluations.

A variety of explanations have been offered for the illusory correlation (see Sherman et al., 2009; Stroessner & Plaks, 2001). Each of the explanations has received empirical support, suggesting that the effect is multiply determined. Indeed, nothing in this chapter is meant to suggest that there is only one factor in the development of illusory correlations. The original account offered by Hamilton and Gifford (1976) is based on the notion that distinctive information draws attention. Because Group B members and negative behaviors are both numerically distinct, the combination of Group B members performing negative behaviors will be particularly salient (compared to any other group/behavior combination). The greater attention given to the distinctive negative Group B behaviors creates the misperception that Group B is, in fact, less favorable than Group A.

AT offers a different account. According to AT, because Group A is larger than Group B, people learn about Group A first. Because positive behaviors are

more frequent than negative behaviors, the impression formed of Group A is a positive one. Subsequently, in forming impressions of Group B, it must be the negative behaviors (the only remaining behaviors) that distinguish it from Group A, and receive particularly close attention. Thus, in order to distinguish Group B from Group A, perceivers focus attention on their negative behaviors. In this case, it is the distinctiveness of the negative behaviors vis-à-vis the existing impression of Group A that draws attention rather than their raw numerical distinctiveness as described by Hamilton and Gifford (1976).

Thus, from the perspective of AT, whether or not group differences are real is irrelevant. All that matters is that one of the groups is learned about prior to the other one. At that point, the identical process may produce both category accentuation and illusory correlation.

Review of Research

Stereotype Formation and Perceptions of Group Members

In our first experiment (Sherman et al., 2009), we simply replicated the inverse base effect with social category stimuli. Participants learned about two groups of people (referred to as Group A and Group B) that differed in size (e.g., 75% of the targets encountered belonged to the "majority" group and 25% of the targets encountered belonged to the "minority" group). For each group, there was a perfect predictor trait (e.g., friendly, intelligent) that characterized all members of the group and never characterized members of the other group. In addition, all members of both groups were described with an imperfect predictor (e.g., lazy). Participants were required to guess which group each individual belonged to based on the traits presented about the individual, and they were provided with the correct answer following each response. Following this learning phase, in the test phase, participants were presented with individuals possessing novel combinations of traits that had not been presented during learning and were asked to assign these targets to the two groups.

Two findings are of primary significance. First, novel targets that possessed only the imperfect predictor (I) were much more likely to be assigned to the majority than the minority group. Given that this trait did not differentiate members of the majority and minority groups and that 75% of the targets encountered belonged to the majority group, this demonstrates that participants had, in fact, learned the base rates and were logically assigning ambiguous targets to the more common group. Second, novel targets that possessed both the perfect predictor of the majority group (PC) and the perfect predictor of the minority group (PR) were more likely to be assigned to the minority group, demonstrating an inverse base rate effect. That is, rather than relying on base rates to assign these ambiguous targets to the majority group, participants responded counter to the base rates and assigned the targets to the minority group.

In our second experiment (Sherman et al., 2009), we sought evidence for the AT model of stereotype formation using an illusory correlation paradigm. However, it is not possible to provide a clear test of the AT account in the standard illusory correlation paradigm because there are not two distinct dimensions. Rather, the group descriptions differ in terms of a single global evaluative dimension (positive-negative). The standard illusory correlation results show that Group A is judged more favorably along this evaluative dimension than is Group B, but cannot show that Groups A and B are associated differentially with different dimensions. That is, it is impossible to identify independent positive and negative impressions of the two groups. Thus, one cannot distinguish whether Group A is more positive than Group B, Group A is less negative than Group B, or Group A is both more positive and less negative than Group B.

To test the AT account, we modified the standard illusory correlation design to make common and rare attributes independent trait dimensions. For example, for some participants, Group A consisted of 16 intelligent and 8 friendly members, whereas Group B consisted of 8 intelligent and 4 friendly members. This design maintains the essential features of the illusory correlation paradigm: The majority group is twice the size of the minority group, one trait is more frequent than the other, and the ratios of the two traits are identical both between and within the two groups. However, in this case, we were able to examine differences in perceptions of the two groups independently for the common and rare trait attributes, permitting tests of the AT explanation for the illusory correlation. In this example, AT predicts both that Group A will be judged as more intelligent than Group B, and that Group B will be judged as more friendly than Group A. Results supported both predictions. Analogous results were observed for estimates of the numbers of group members possessing different traits that belonged to each group. Participants estimated that there were relatively more individuals possessing the uncommon than the common trait belonging to the minority group, despite the fact that both traits had the same base rates (.33) among the minority group. The same pattern was observed in group assignments for specific targets: Those possessing the uncommon trait were significantly more likely to be assigned membership to the minority group than those possessing the common trait. Finally, the accuracy of these assignments also supported AT: For the minority group, assignments of rare trait targets were more accurate than assignments of common trait targets, whereas, for the majority group, assignments of common trait targets were more accurate than assignments of rare trait targets. This result is consistent with the AT proposal that perceivers attend more carefully to majority group members that confirm the majority group stereotype and more carefully to minority group members that distinguish the minority group from the majority group, supporting the formation of a distinct minority group stereotype. No existing model of illusory correlation predicts any of these results (for a full discussion, see Sherman et al., 2009).

	Common Trait	Rare Trait
Group A	24	3
Group B	12	6

FIGURE 7.3 Modified configuration of illusory correlation information used in Experiment 4.

Experiment 3 (Sherman et al., 2009) replicated the main results from both Experiments 1 and 2 and also showed that, in the inverse base rate design of Experiment 1, the minority group was rated more strongly on its perfect predictor (PR) than the majority group was rated on its perfect predictor (PC). This provides another piece of evidence that the associations between minority groups and their stereotypes are stronger than are the associations between majority groups and their stereotypes.

Experiment 4 used a modified version of the inverse base rate design to provide further distinct support for the AI account. In this case, rather than the majority and minority groups each possessing a perfect predictor, both groups possessed the common and rare traits probabilistically. Specifically, whereas the common trait described the majority group twice as frequently as the minority group, the rare trait described the minority group twice as frequently as the majority group (see Figure 7.3). Thus, although there were real differences between the two groups, the groups were associated equally strongly with their respective traits (2:1 ratio between groups). However, according to AI, participants should form stronger associations between the minority group and the rare trait than between the majority group and the common trait. Indeed, at test, when participants were presented with novel targets possessing both the common and rare trait, they were more likely to be assigned to the minority group, consistent with the predictions of AI. These results are problematic for all other models of illusory correlation (see Sherman et al., 2009).

Finally, Experiment 5 directly measured the attention directed toward different group/trait pairings. In the learning phase of the experiment, participants were presented with information about members of majority and minority groups, as in the illusory correlation paradigm in Experiment 2. Subsequently, they were shown novel targets along with two descriptions of their behavior. One behavioral description reflected the common trait and the other reflected the rare trait. After one of four randomly determined delays, both behaviors disappeared, an X was displayed on one side of the screen, and participants were required to press a button to indicate the side where the X had appeared. This "dot-probe"

task uses response times to measure the amount of attention being directed at competing items of information. Consistent with AI, results showed that the majority group targets, participants responded more quickly to the probe when it appeared on the side of the screen with the behavior reflecting the common trait. In contrast, for minority group members, participants responded more quickly to the probe when it appeared on the side of the screen with the behavior reflecting the rare trait. Thus, these data provide direct evidence that greater attention is paid to traits that first define the majority group and those that subsequently help to differentiate the minority group from the majority group.

Further Evidence for the Strength of Minority Versus Majority Group Stereotypes

The initial results demonstrated, among other things, that minority group stereotypes carry more weight than do majority group stereotypes. When a single target is described as possessing the attributes of both majority and minority group stereotypes, that person is more likely to be perceived as a minority group member. In our research (and in traditional research on illusory correlation), the groups are previously unknown. As such, this effect can be attributed to the fact that the minority group is unusual and its features are learned only after those of the majority group. In his dissertation research, Dario Sacchi has further examined the implications of AI for the differential strength of majority and minority group stereotypes.

In two experiments, participants first completed the initial learning phase of the social group inverse base rate effect in Experiment 1 described above. After completing learning of the majority and minority groups and their predictors, participants completed different dependent variables that indicated the strength of the stereotypes that were formed. In the first experiment, after learning, participants were asked to complete a scale measuring the extent to which each group was seen as having a stable, underlying essence (Bastian & Haslam, 2008). Results showed that (a) the minority group was perceived as having more *discrete* boundaries than the majority group (e.g., everyone is either a certain type of person or they are not), (b) the traits describing members of the minority group were perceived as more *informative* than the traits describing members of the majority group (e.g., when getting to know a person it is possible to get a picture of the kind of person they are very quickly), and (c) the traits describing the minority group were perceived as having a *biological* basis to a greater extent than the traits describing the majority group (e.g., the kind of person someone is can be largely attributed to their genetic inheritance). These results indicate that one reason for the greater weight given to minority versus majority stereotypes in judgments is that the minority stereotypes are perceived as possessing greater essence.

In the second experiment, participants were asked to make trait ratings about the groups after learning about them and were subsequently presented with

information about 10 new members of each group. Of the 10 new members of each group, 5 behaved in a way that confirmed the newly formed stereotype of their group and 5 behaved in a way that disconfirmed that stereotype. Subsequently, participants were asked to again make trait ratings of the groups. Results showed that minority group stereotypes changed less in response to the five disconfirming group members than did majority group stereotypes, another indication of the relative strength of minority versus majority group stereotypes. Moreover, disconfirming minority group members were judged to be less typical of their groups than were disconfirming majority group members, indicating greater expectancies of homogeneity among minority than majority group members.

These results demonstrate important additional consequences of having to distinguish a novel minority group from an already-known majority group. As a result of this process, the stereotypes applied to novel groups are weighted more heavily in judgments of group members. Moreover, stereotypes about these groups are stronger in that the groups are perceived as having a more stable underlying essence, the stereotypes are more difficult to change, and group members who disconfirm the stereotype are perceived as poor exemplars of the group. Because minority groups are numerically rare, by definition, for many people, knowledge of these groups will form only after larger groups have already been encountered and characterized. This provides one possible explanation (among others) as to why the most prevalent stereotypes in most (all?) cultures are those that pertain to minority groups.

Stereotypes Based on Physical Features: The Case of Hypodescent

To this point, we have described research showing that trait stereotypes of minority groups and their members are weighted more heavily in judgments and are held more strongly than are trait stereotypes of majority groups. The AT explanation for these phenomena suggests that such effects should not be limited to psychological attributes such as traits. For example, physical features that are typical of minority groups might also be expected to be more strongly tied to their groups than are physical features that are typical of majority groups. In fact, the heavily-researched phenomenon of *hypodescent* may be one important example of just such an effect. Hypodescent is the tendency for individuals of mixed-race ancestry to be associated with the minority or more socially subordinate group in their ancestry (Banks & Eberhardt, 1998), the most famous example being that Barack Obama is almost universally perceived as Black and not White.

Hypodescent has a number of critical political and historical antecedents, but we were interested in examining the possible role of attentional processes in category learning in contributing to the phenomenon. To this end, we conducted two experiments (Halberstadt, Sherman, & Sherman, 2011). In the first experiment,

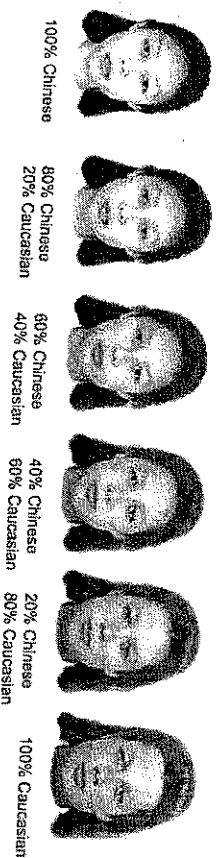


FIGURE 7.4 Example of a Chinese-Caucasian face pair used in Study 1 of Halberstadt et al., 2011 and 4 of its 24 morphed blends.

we presented native Chinese and native Caucasian participants in New Zealand with pictures that morphed images of Chinese and Caucasian targets. The morphs ranged from completely unambiguous (100% Chinese or Caucasian) to completely ambiguous (50% Chinese/50% Caucasian; see Figure 7.4). Participants' task was to judge whether each face was that of a Chinese or Caucasian person.

Given their backgrounds, we assumed that the Chinese participants would have been first exposed to Chinese faces, whereas the Caucasian participants would have been first exposed to Caucasian faces. As such, we assumed that each group of participants would have first learned the facial features that were typical of their ingroups. According to AT, later in life, when encountering members of outgroups, participants would have focused attention on those facial features that best distinguished outgroup from ingroup members. Indeed, AT predicts that the facial features of outgroup members would be more strongly associated with the outgroup than the facial features of ingroup members would be associated with the ingroup. As such, judgments of ambiguous faces should be influenced more strongly by outgroup features, and such targets should be more likely to be assigned to the outgroup than the ingroup. This is analogous to the inverse base rate effect observed with other features in category learning. Consistent with this prediction, we found that Chinese participants were more likely to judge ambiguous faces as Caucasian than were Caucasian participants (who were more likely to judge those same faces as Chinese than were Chinese participants).

An alternative account of these findings can be found in the phenomenon of *ingroup overexclusion* (Leyens & Yzerbyt, 1992), which describes the tendency for people to exclude ambiguous individuals from their ingroups in order to protect the distinctiveness of the ingroup. In order to rule out such motivational factors, we conducted a second experiment, in which we directly manipulated the facial features of members of novel majority and minority groups to which the participants did not belong. Essentially, we replicated the inverse base rate effect studies we had done with majority and minority groups, swapping facial features for trait descriptions. After the learning phase, participants were asked to classify facial morphs of majority and minority group members that varied

in ambiguity. Consistent with the predictions of AT, results showed that ambiguous faces were more likely to be assigned to the minority than the majority group. Just as in the inverse base rate studies, when the typical features of majority and minority groups were both present, the features of the minority group were weighted more heavily in judgments of group membership.

Impressions of Individual Targets in Different Contexts

Most recently, we have begun to explore the implications of AT for how people may form context-dependent impressions of individual targets. Theoretically, in the same way that perceivers learn how to differentiate people based on the groups that they belong to, they may also learn how to differentiate the behavior of single individuals based on the particular contexts in which different behavior occurs. Consider a person who behaves very differently when he is with his friends versus when he is with his family. If you first meet this person in the context of his friendships and only later encounter him in the context of his family, AT suggests that you will pay particularly close attention to whatever behaviors distinguish his family behavior from his friend behavior. Subsequently, because of this shift of attention, your impression of him in the family context will be held more strongly than your impression of him in the friend context. The logic is exactly the same as in the research already described.

We have conducted a number of experiments to test this application of AT. In one variant, participants learned about a target person named Dave (naturally!) and how his behavior varied according to context. Just as in the original stereotype formation experiments, there was a common trait (C; presented 75% of the time), a rare trait (R; presented 25% of the time), and an imperfect predictor trait that was present for all descriptions of Dave. In this case, the contexts were different colored rooms in which the behavior occurred. During learning, participants were presented with different trait descriptions of Dave and had to guess which colored room he was in. So, for example, participants might learn that Dave is friendly and honest in the blue room (C) but intelligent and honest in the red room (R). After accurately learning the traits associated with each context, participants were presented with different configurations of traits and were asked to guess which room Dave is in.

When Dave was described only with the imperfect predictor (I), participants were more likely to guess that he was in the common context than the rare context. This shows that they had accurately learned the base rate frequencies of the different contexts. Of most interest, when Dave was presented as having both the common context trait and the rare context trait, participants were more likely to guess that he was in the rare context than the common context. That is, they demonstrated an inverse base rate effect. Thus, impressions of a person in a novel context may be stronger than impressions first formed in another context.

In another variant of this design, we switched which information about Dave was given and which classification needed to be learned. In this case, participants were shown pictures of Dave in different contexts and they had to guess which trait he exhibited in that context. At test, participants were shown different combinations of contexts and were asked to guess which trait Dave exhibited. For example, participants might see Dave in a room that was half red and half blue and predict which trait he exhibited. In a different study, the contexts were different friends of Dave (e.g., Jim and Bob), and participants were tested with different combinations of friends (e.g., Dave is with Jim and Bob). Results demonstrated the predicted effects. When presented with the imperfect predictor (I) context, they were more likely to assign the common trait descriptor to Dave, demonstrating that they had learned the base rates of the different traits. However, when presented with both the common and the rare context (e.g., half red and half blue), participants were more likely to assign Dave the rare trait than the common trait, demonstrating an inverse base rate effect.

These data show that the same processes that influence the formation of distinct group stereotypes also can lead to the formation of distinct contextualized impressions of individual targets. As in the case of minority group stereotypes, when people learn about a person in a novel ("minority") context, the features of that context are weighted more heavily in subsequent judgments than are the features of the common ("majority") context in which the target was first encountered.

Implications for Attitude Change

These findings have important implications for producing attitude change. AT suggests that people may pay particularly close attention to attitude-disconfirming information when it appears in a novel or rare context, in order to differentiate their established attitudes from the conflicting novel context. Further, AT suggests that not only will people form different attitudes in the novel context but those contextualized attitudes may be held more strongly than the original attitudes that they undermine, due to the attention shifting mechanism. One important implication is that, if the novel attitudinal context can be introduced in more common contexts, it may undermine or even overwhelm the prevalent attitude. Consider an alcoholic who has very favorable attitudes toward alcohol in social settings, but has established a new, negative attitude in a novel setting such as a rehab center. AT suggests that, if the rehab center context can somehow be carried into the more prevalent social context, then it may undermine the positive attitude toward alcohol in that context. One possible means to achieve this "contextual conflict" is through the use of super-conditioning, in which an unconditioned stimulus is trained as a conditioned inhibitor (e.g., Williams & McDermott, 2002). For example, a rubber band worn on the wrist during exposure to novel, contextualized counterattitudinal information may effectively be used

to carry that counter-attitude back into more common everyday social settings. Of course, the likelihood of this outcome depends on many other variables, such as the strength with which the original attitude is held, the strength of the counterattitudinal information, and so on. Moreover, the ability to undermine an original attitude may depend on the extent to which that attitude, itself, is already contextualized (e.g., Gawronski, Rydell, Vervliet, & De Houwer, 2010). If the original attitude generalizes across many contexts, the process of distinguishing the new attitudinal context from the old one may be compromised.

Conclusion

Altogether, this line of research demonstrates the potential power of a very simple idea. The fundamental message of AT is that what people learn about something depends on what they already know; this is hardly a controversial claim. The attentional process described by AT reflects a rather elementary mechanism for distinguishing some things from others to accomplish learning. To date, this mechanism has been shown to be able to account for common errors in category learning, such as the inverse base rate effect and base rate neglect, category accentuation processes, the formation of illusory correlations, the relative strength of minority versus majority stereotypes, judgments of hypodescent, and the formation of contextualized impressions of other people. We look forward to exploring further applications of this idea, including to processes of attitude formation and change.

As a personal matter, I (Sherman) am gratified to have finally made a contribution to our understanding of illusory correlation. Perhaps, I picked up more than I thought as a graduate student working with Dave. That there remain interesting questions and avenues of research in this literature is a tribute to the tremendous generativity of the research program Dave established. Personally, I view my work on this topic as a tribute to Dave and as an acknowledgement of his continuing influence on how I think about psychology. Notably, though the AT model departs in some important ways from Dave's original account of the illusory correlation, we have retained attentional processes as the fundamental engine of the effect. Finally, this line of research has provided the opportunity to work closely and extensively with my dad, my other major intellectual role model. Working with him on advancing understanding of the illusory correlation brings the Hamilton-Sherman axis full circle, back to where it could not have started when I was in graduate school.

Note

1. Note that these effects are due to the numerical distinctiveness of the negative behaviors rather than to their negativity per se. Thus, the effects are reversed when the majority of the behaviors are negative rather than positive. In this case, Group B is perceived to be more positive than Group A (Hamilton & Gifford, 1976).

References

- Banks, R. R., & Eberhardt, J. L. (1998). Social psychological processes and the legal bases of racial categorization. In J. L. Eberhardt & S. T. Fiske (Eds.), *Confronting racism: The problem and the response* (pp. 54-75). Thousand Oaks, CA: Sage.
- Bastian, B., & Haslam, N. (2008). Psychological essentialism and social identification: Immigrant from two perspectives. *Asian Journal of Social Psychology*, 11, 127-140.
- Cornille, O., & Judd, C. M. (1999). Accentuation and sensitization effects in the categorization of multi-faceted stimuli. *Journal of Personality and Social Psychology*, 77, 927-941.
- Eiser, J. R. (1971). Enhancement of contrast in the absolute judgment of attitude statements. *Journal of Personality and Social Psychology*, 17, 1-10.
- Ford, T. E., & Stangor, C. (1992). The role of diagnosticity in stereotype formation: Perceiving group means and variance. *Journal of Personality and Social Psychology*, 63, 356-367.
- Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus contextualization in automatic evaluation. *Journal of Experimental Psychology: General*, 139, 683-701.
- Halberstadt, J., Sherman, S. J., & Sherman, J. W. (2011). Why Barack Obama is Black: A cognitive account of hypodescent. *Psychological Science*, 22, 29-33.
- Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in intergroup perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology*, 12, 392-407.
- Krueger, J., & Clement, R. W. (1994). Memory-based judgments about multiple categories: A revision and extension of Tajfel's accentuation theory. *Journal of Personality and Social Psychology*, 67, 35-47.
- Krueger, J., & Rothbart, M. (1990). Contrast and accentuation effects in category learning. *Journal of Personality and Social Psychology*, 59, 651-663.
- Krueger, J., Rothbart, M., & Strim, N. (1989). Category learning and change: Differences in sensitivity to information that enhances or reduces intercategory distinctions. *Journal of Personality and Social Psychology*, 56, 866-875.
- Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 3-26.
- Kruschke, J. K. (2001). The inverse base rate effect is not explained by eliminative inference. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27, 1385-1400.
- Kruschke, J. K. (2003). Attention in learning. *Current Directions in Psychological Science*, 12, 171-175.
- Leyens, J.-P., & Yzerbyt, Y. (1992). The ingroup overexclusion effect: Impact of valence and confirmation on stereotypical information search. *European Journal of Social Psychology*, 22, 549-569.
- McGarty, C., & Penny, R. E. C. (1988). Categorization, accentuation and social judgement. *British Journal of Social Psychology*, 22, 147-157.
- McGarty, C., & Turner, J. C. (1992). The effects of categorization on social judgment. *British Journal of Social Psychology*, 31, 253-268.
- Medin, D. L., & Edelson, S. M. (1988). Problem structure and the use of base-rate information from experience. *Journal of Experimental Psychology: General*, 117, 68-85.
- Queller, S., Schell, T., & Mason, W. (2006). A novel view of between-categories contrast and within-category assimilation. *Journal of Personality and Social Psychology*, 91, 406-422.

- Sherman, J. W., Kruschke, J. K., Sherman, S. J., Percy, E. J., Petrocelli, J. V., & Conrey, F. R. (2009). Attentional processes in stereotype formation: A common model for category accentuation and illusory correlation. *Journal of Personality and Social Psychology*, 96, 305–323.
- Stroessner, S. J., & Plaks, J. E. (2001). Illusory correlation and stereotype formation: Tracing the arc of research over a quarter century. In G. B. Moskowitz (Ed.), *Cognitive social psychology: The Princeton symposium on the legacy and future of social cognition* (pp. 247–259). Mahwah, NJ: Lawrence Erlbaum Associates.
- Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgment. *British Journal of Psychology*, 54, 101–114.
- Williams, B. A., & McDevitt, M. A. (2002). Inhibition and superconditioning. *Psychological Science*, 13, 454–459.

8

TWO (OR MORE?) COGNITIVE APPROACHES TO STEREOTYPE FORMATION

Biased or Reality Based?

Russell Spears and Wolfgang Stroebe

Although the stage for cognitive theories of stereotype formation and stereotyping was set by Allport's (1954) classic volume *The Nature of Prejudice*, two cognitive theories which have dominated research in this area have been Tajfel's accentuation theory of stereotyping (e.g., Tajfel, 1957, 1959, 1969; Tajfel & Wilkes, 1963) and Hamilton's distinctiveness theory of illusory correlation (e.g., Hamilton & Gifford, 1976). Whereas stereotype formation according to accentuation theory is the result of categorization processes, stereotype formation according to Hamilton is (or can be) the result of memory distortions in processing information about different groups. One major difference between these theories is that according to accentuation theory, stereotypes reflect an accentuation of actual differences between social categories, whereas the mechanism proposed by Hamilton does not require real group differences to exist. The focus of this chapter is to address the explanation of stereotype formation by these two theories, particularly in relation to this reality versus bias theme.

Although both of these theories are cognitive in nature, they derive from two very different cognitive approaches to social psychology, namely social judgment and social cognition. Whereas the origins of social judgment theories can ultimately be traced back to Fechner's psychophysics (Eiser, 2012; Eiser & Stroebe, 1971), the social cognition approach derived from the more recent information processing framework relying heavily on cognitive representation of information and memory processes. The central question for social judgment theories is how people select the frame of reference they employ in their judgments. Social identity theory (SIT; Tajfel & Turner, 1979) and self-categorization theory (SCT; Turner, Hogg, Oakes, Reicher, & Wetherell, 1987) continued this tradition. The social cognition tradition, on the other hand, derived from information processing and memory research and focuses on encoding and recall of information.