



# Attentional Processes in Social Perception

Lisa M. Huang<sup>1</sup>, Jeffrey W. Sherman

University of California, Davis, CA, United States

<sup>1</sup>Corresponding author: e-mail address: lmhuang@ucdavis.edu

## Contents

1. What Is Attention Theory?	201
2. Stereotyping and Group Categorization	203
2.1 Category Accentuation and Illusory Correlation	204
2.2 Attention Theory as a Common Model of Stereotype Formation	205
2.3 Empirical Evidence	206
2.4 Further Implications for Stereotype Formation	212
2.5 Stereotype Strength	213
2.6 Face Perception and Group Categorization	215
3. Context-Based Impression Formation	217
3.1 Context Frequency and Impression Strength	217
3.2 Expectancies Shape Context-Based Impression Formation	222
3.3 Changing Impressions	229
4. Conclusion	232
Acknowledgment	236
References	236

## Abstract

In this chapter, we describe how a simple attentional mechanism can account for a wide variety of phenomena in social perception. According to Attention Theory (Kruschke, 1996, 2003), people preferentially attend to differentiating information in order to maximize category learning. When learning multiple social categories, people attend to all features that characterize the first-learned category but shift their attention to features that uniquely distinguish a later-learned category from the first. As a result, they form a stronger impression of the later-learned social category. First, we review research on attentional processes in stereotype formation and group categorization. We show how Attention Theory can account for both category accentuation and illusory correlation in the formation of majority and minority group stereotypes. We then explain how attention shifting influences face perception and racial categorization. Second, we describe attentional processes as they relate to context-based impression formation and the influence of individual- and group-based expectancies on context-based impressions. Last, we discuss implications for impression change.

The environment is packed with more information than one person can mentally process at any given moment. With limited capacity, people must select which pieces of information to attend to at any given point in time. For example, your attention at this moment is focused on reading this chapter, but other events are occurring in your surroundings of which you might not be aware. Perhaps there are sounds of laughter outside your door or cars driving by outside the window or the fluorescent light is flickering above your head, but you are unlikely to notice these peripheral details in your environment if you are engaged with a different task.

Fortunately, our attention is adaptive, so we tend to focus on the most important pieces of information, such as those which meet our current goals (like reading this chapter) or deviate from the usual state of our environment (such as a bee wandering into the room) that demand immediate attention. Indeed, research shows that people selectively direct their attention in ways that optimize learning (Kruschke, Kappenman, & Hetrick, 2005; Matsuka & Corter, 2008; Nosofsky, 1984; Rehder & Hoffman, 2005). In addition, people naturally attend to novel information in order to reduce uncertainty (Gottlieb, Oudeyer, Lopes, & Baranes, 2013; Luque, Vellido Nistal, Le Pelley, & Beesley, 2017) or adapt to a changing environment (Ranganath & Rainer, 2003). Indeed, the brain and nervous system are wired to automatically orient people's attention toward novel stimuli in the environment (Gottlieb et al., 2013; Ranganath & Rainer, 2003; Sokolov, 1963). Even infants as young as 3–6 months old spend more time fixating on novel rather than familiar stimuli, suggesting that attention to novel information is innate (Fantz, 1964).

Attentional processes are adaptive for forming impressions of social targets as well. For example, behaviors that disconfirm perceivers' beliefs about another person attract more attention than confirming behaviors, allowing perceivers to update their impressions or make sense of behavioral discrepancies (Fiske, Lin, & Neuberg, 1999; Hilton, Klein, & von Hippel, 1991). In general, behaviors that are unexpected (Hilton et al., 1991; Roese & Sherman, 2007; Stangor & McMillan, 1992), negative (Fiske, 1980; Pratto & John, 1991; Skowronski & Carlston, 1989), or extreme (Fiske, 1980; Skowronski & Carlston, 1989) attract greater attention. These attentional biases have interesting implications for impressions and judgments of social targets. For example, perceivers tend to weight negative information more heavily than positive information during social judgment (Fiske, 1980). As another example, perceivers tend to remember expectancy-incongruent behaviors better than expectancy-congruent behaviors because they pay more attention to

incongruent behaviors, but their overall judgment of the social target is more in line with the congruent behaviors (e.g., [Sherman & Frost, 2000](#); [Stangor & McMillan, 1992](#); for reviews, see [Sherman, Allen, & Sacchi, 2012](#); [Von Hippel, Sekaquaptewa, & Vargas, 1995](#)).

If attentional processes are adaptive, it should follow that perceivers pay more attention to information that differentiates one social category from another. Differentiating information is informative and allows perceivers to maximally distinguish between two categories. This attentional preference should shape the way that perceivers form subsequent impressions of the social categories. In this chapter, we describe a simple attention-shifting mechanism, as posited by the Attention Theory (AT) of category learning, that explains how perceivers differentially attend to categorical information ([Kruschke, 1996, 2003](#)). We then present research showing how this attentional mechanism relates to social perception processes.

In [Section 1](#), we describe the AT mechanism of learning. In [Section 2](#), we describe attention shifting as it relates to stereotype formation and group categorization. We present research showing how AT can account for both category accentuation and illusory correlation in the formation of majority and minority group stereotypes. We then explain how attention shifting influences face perception and racial categorization. Specifically, we explain how AT can account for hypodescent, the categorization of mixed-race targets as a member of the lower status group. In [Section 3](#), we discuss how the attention-shifting mechanism influences the formation of context-based impressions of individuals. We discuss research showing how this mechanism leads to the formation of stronger impressions of individuals who are encountered in rare contexts relative to common contexts. We then present research that examines how perceivers' expectations influence context-based impression formation of individuals and group members. We describe the implications of these processes for stereotype maintenance and explore whether these impressions can be changed.



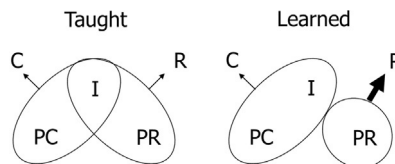
---

## 1. WHAT IS ATTENTION THEORY?

Attention theory was developed, in part, to account for the *inverse base-rate effect* in human learning. 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), 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 (see Fig. 1). Subsequently, when learning about the 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 toward 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 (Kruschke, 1996, 2003). The shift in attention toward the unique predictive



**Fig. 1** *Left:* The core design of the inverse base-rate effect. C and R represent the common and rare disease, respectively. The symptom PC is a perfect predictor of C, the symptom PR is a perfect predictor of R, and the symptom I is an imperfect predictor of both C and R. *Right:* Depiction of what is learned according to Attention Theory.

cue of the rare disease facilitates learning of both diseases (Matsuka & Corter, 2008; Rehder & Hoffman, 2005).

If people preferentially attend to distinctive information when performing a category learning task, then the same principles should apply when they learn about social categories. When one social category is learned first, perceivers should attend equally to all attributes that are associated with the first social category. Initially, no one trait is more predictive of that category than any other trait. However, when perceivers subsequently learn about the second social category, they should attend to traits that differentiate the second category from the first, rather than to traits that apply to both categories. The second category, then, becomes more strongly associated with its differentiating traits.



## 2. STEREOTYPING AND GROUP CATEGORIZATION

In our first line of research, we examined how attentional processes influence the formation of group stereotypes. According to AT, majority group traits are learned before minority group traits because majority group members are encountered with greater frequency. When perceivers subsequently learn about the minority group, they shift their attention toward the traits that most distinguish them from the majority group. The distinguishing traits then become more strongly associated with the minority group compared to traits shared by both groups. Because perceivers attend more to the distinctive traits of the minority group when learning about that group, they should form a stronger association between the minority group and its unique traits than between the majority group and its unique traits, resulting in stronger minority group stereotypes. Indeed, this process could partially explain why minority group members are more often the target of stereotyping than majority group members.

In our research, we showed that AT provides a unifying framework for understanding two prominent effects found in the stereotype formation literature—category accentuation (Tajfel & Wilkes, 1963) and illusory correlation (Hamilton & Gifford, 1976). Whereas category accentuation effects highlight the exaggeration of real intergroup differences as the basis for stereotype formation, illusory correlation shows that stereotypes may be formed in the absence of real group differences. Research on the two effects has proceeded independently, and they have been explained by different mechanisms.

## 2.1 Category Accentuation and Illusory Correlation

Work on category accentuation shows that the division of people into separate categories exaggerates both perceived differences between categories and perceived 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 intergroup 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 ([Tajfel & Wilkes, 1963](#)). Moreover, category members who 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, & Sriram, 1989](#)). An important feature of each of these mechanisms is that they assume real differences between the categories in question and that these processes serve to accentuate those differences.

The distinctiveness-based illusory correlation describes a phenomenon whereby observers perceive an association between distinctive groups and distinctive behaviors when, in fact, no such relation exists ([Hamilton & Gifford, 1976](#); for a review, see [Stroessner & Plaks, 2001](#)). In the model 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 as Group B, and there are more positive than negative behaviors. For example, in the original experiment, members of Group A performed 18 positive and 8 negative behaviors, whereas members of Group B performed 9 positive and 4 negative behaviors. Thus, Group B and negative behaviors both are distinct because of their infrequency. 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 relation 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 rate Group A more favorably than Group B on trait-rating tasks. Second, participants often overestimate the numbers of negative vs positive behaviors performed by Group B. Finally, participants misattribute Group A's negative behaviors to Group B. These

effects cannot be explained by the negativity of the behaviors per se because the effects are reversed when the majority of behaviors are negative rather than positive (Hamilton & Gifford, 1976).

## 2.2 Attention Theory as a Common Model of Stereotype Formation

A number of processes have been posited to account for stereotype formation that is based on the exaggeration of real group differences (Corneille & Judd, 1999; Krueger & Clement, 1994; Krueger & Rothbart, 1990; McGarty & Penny, 1988; McGarty & Turner, 1992; Queller et al., 2006) and on perceptions of illusory group differences (Fiedler, 1991; Hamilton & Gifford, 1976; McGarty, Haslam, Turner, & Oakes, 1993; Rothbart, 1981; Smith, 1991). The mechanisms proposed to account for category accentuation are different from those posited to account for illusory correlation. However, our research showed that Kruschke's (1996, 2001, 2003) attentional mechanism can account for both category accentuation and illusory correlation in stereotype formation.

Historically, the processes proposed to account for category accentuation are similar to the attention shifting processes described by AT. According to category accentuation models, one key determinant of accentuation is that people attend more carefully to group members who heighten between-group differences and within-group similarities. Similarly, in AT, people focus on the features that most effectively distinguish one category from another category. However, one important difference between-category accentuation and AT 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, any 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. Because attention is directed toward differentiating features of the second category, the association between the second category and its features should be stronger (e.g., a stronger association between a social group and its differentiating trait). Thus, AT provides an account not only of how categories are differentiated from one another but also of which particular features come to characterize those categories.

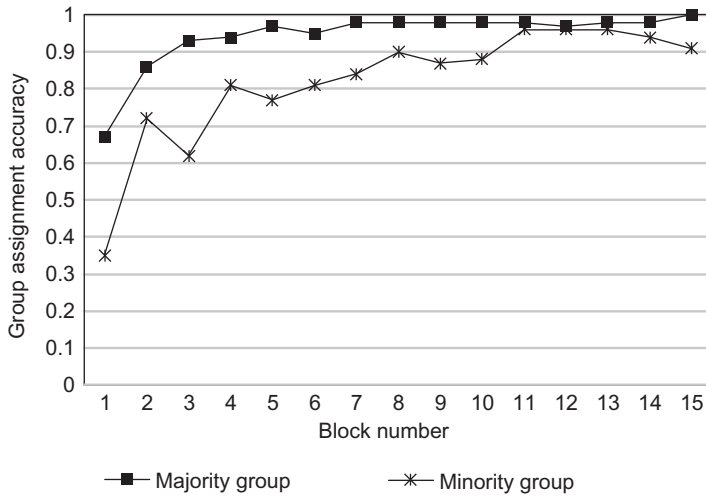
These same processes may produce the illusory correlation. According to AT, because Group A is larger than Group B, people learn about Group A first. Because positive behaviors are more frequent than negative behaviors, the impression formed of Group A is a positive one. Subsequently, in forming impressions 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 proposed 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 attentional process may produce both category accentuation and illusory correlation.

## 2.3 Empirical Evidence

We tested the AT account of stereotype formation across a series of experiments. In our first experiment ([Sherman et al., 2009](#)), we simply replicated the inverse base-rate effect with social category stimuli. During a learning phase, participants learned about two groups of people (e.g., 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., reliable, friendly) 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., stingy). Across 15 blocks of trials, 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 the learning phase, in the test phase, participants were presented with new individuals possessing novel combinations of the traits and were asked to assign these targets to one of the groups.

Responses on the learning phase showed that participants learned the majority group traits faster than the minority group traits (see [Fig. 2](#)). In earlier blocks of trials, participants were more accurate on trials that characterized the majority group, but the difference in accuracy between groups decreased in later blocks as participants subsequently learned the minority





**Fig. 2** Group assignment accuracy per block, separated by group (Sherman et al., 2009, Experiment 1). In earlier blocks of trials, participants were more accurate on trials that characterized the majority group, but the difference in accuracy between groups decreased in later blocks as participants subsequently learned the minority group traits.

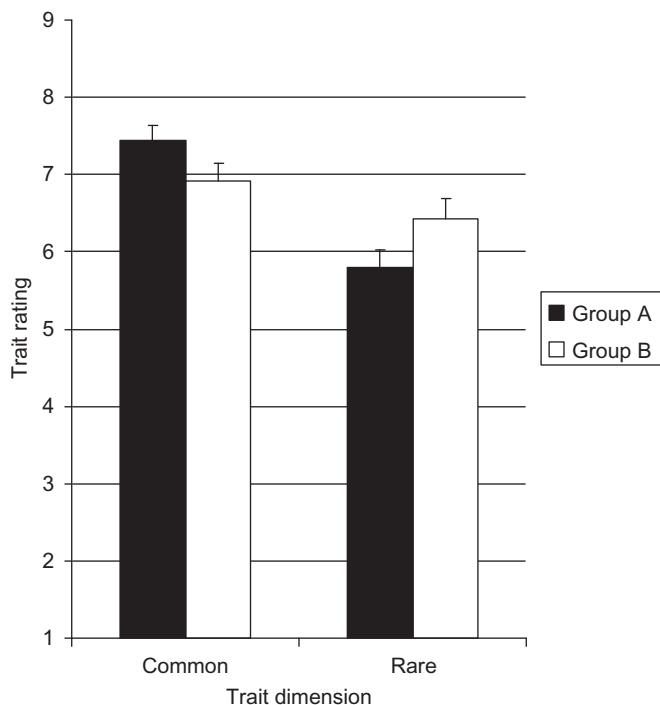
group traits. This finding is consistent with AT; participants learned about the majority group first and the minority group second.

The test phase results reflected the original inverse base-rate effect. First, new individuals who 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, new individuals who 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. As in the original inverse base-rate effect, this finding shows that the association between the minority group and its trait is stronger than the association between the majority group and its trait. When participants were given the competing predictors, the minority group trait outweighed the majority group trait.

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 in the standard demonstration of the effect. 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 in 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 such that the common and rare attributes were 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. Specifically, 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. Participants tended to rate Group A higher on the common trait dimension (e.g., intelligent), whereas they tended to rate Group B higher on the rare trait dimension (e.g., friendly; see Fig. 3). Analogous results were observed for group number estimates: Participants estimated that there were relatively more individuals possessing the rare trait than the common trait belonging to the minority group, even though 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 rare trait were significantly more likely to be assigned to the minority group than those possessing the common trait. Finally, the accuracy of these assignments also supported AT: For the minority group,



**Fig. 3** Trait ratings of Groups A and B (Sherman et al., 2009, Experiment 2). Participants rated Group A higher on the common trait, whereas they rated group B higher on the rare trait. Error bars indicate standard errors.

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 who confirm the majority group stereotype and minority group members who distinguish the minority group from the majority group. No existing model of illusory correlation predicts any of these results (for a full discussion, see Sherman et al., 2009).

Experiment 3 (Sherman et al., 2009) replicated the main results from both Experiments 1 and 2 and, in addition, 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 distinct support for the AT account of category accentuation. 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. In addition, the ratio of common trait descriptions to rare trait descriptions was higher in the majority group (8:1) than in the minority group (2:1; see [Table 1](#)). Even though the common trait occurred more frequently in both the majority and minority groups, there were real differences between the two groups.

According to AT, participants should form stronger associations between the minority group and the rare trait than between the majority group and the common trait. Indeed, the degree of preference for assigning a novel target with the rare trait to the minority group was greater than the degree of preference for assigning a novel target with the common trait to the majority group (see [Table 2](#)). This was true even though the ratio of majority group to minority group members who possessed the common trait was the same as the ratio of minority group to majority group members who possessed the rare trait (2:1). This bias shows that participants accentuated the differences between the two groups. Furthermore, 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 AT.

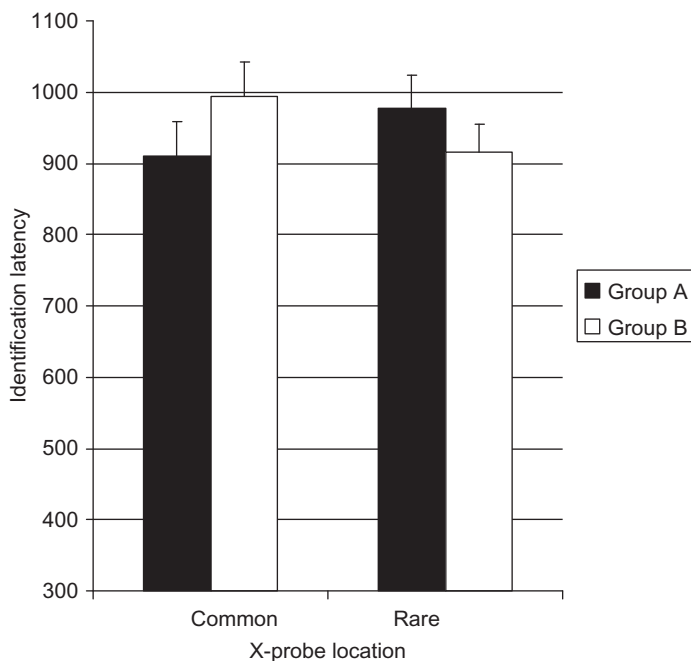
**Table 1** Group–Trait Pairings ([Sherman et al., 2009](#), Experiment 4)

Trait	Number of Pairings	
	Group A	Group B
Common	24	12
Rare	3	6

**Table 2** Group Assignment of Novel Targets ([Sherman et al., 2009](#), Experiment 4)

Trait	Percentage Selected		
	Group A (%)	Group B (%)	Total (%)
Common	67	33	100
Rare	22	78	100

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 appeared on one side of the screen, and participants were required to press a button to indicate the side. This “dot-probe” task can be used to measure the amount of attention being directed at competing items of information by examining response times to identify the location of the probe. Consistent with AT, results showed that, for 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 (see Fig. 4). Thus, these data provided direct



**Fig. 4** X-probe identification latencies (Sherman et al., 2009, Experiment 5). Error bars indicate standard errors.

evidence that people attend to common traits when learning about the majority group and to rare traits when learning about the minority group. Note that these attentional differences occurred even though there were no true differences between the two groups.

## 2.4 Further Implications for Stereotype Formation

Beyond the implications for category accentuation and illusory correlation, the AT approach suggests a number of other important novel hypotheses about stereotype formation and impression formation. Perhaps the most basic message of AT is that learning sequence matters, and that what we learn about a group depends on what we already know about other groups. This observation has a number of important implications for stereotype formation. First, as demonstrated in Experiments 1 and 3, when a trait is highly descriptive of both a majority and a minority group, it is likely to be associated primarily with the majority group. Specifically, in Experiments 1 and 3, even when all members of both the majority and minority groups possessed an imperfect predictor (I), a target possessing only this imperfect predictor (I) was most commonly categorized as a member of the majority group, in line with the base-rate frequencies of the groups. The fact that trait ratings of the I attribute were higher for the majority than the minority group in Experiment 3 indicates that participants not only learned that, probabilistically, a person with Trait I likely belonged to the majority group but also formed an impression that the majority group possessed this trait. These findings suggest that attributes that occur with high frequency in both majority and minority groups are unlikely to become associated with minority groups, regardless of how prevalent the attributes may be among those groups. This suggests a possible basis for in-group bias. Because we learn about in-groups prior to learning about out-groups and because the majority of others' behavior is relatively benign, we are likely to form positive impressions of our in-groups. When we subsequently encounter an out-group, we may be limited in the attributes available for differentiating that group from our own. As such, the best way to differentiate an out-group from an in-group may often be by ascribing negative attributes to the out-group. The same process can lead majority groups to be perceived as more human than minority groups (Prazienkova, Paladino, & Sherman, 2017).

There also are implications for how impressions are formed of individual group members. Consider a case in which a perceiver first meets either a White woman or a Black man and then meets a Black woman. When the first

person encountered is a White woman, the feature of the Black woman that distinguishes her from the first target will be her race. In this case, racial stereotypes may play a relatively large role in the impressions formed of the second woman. Knowledge about her race may bias attention, comprehension, memory, and judgment processes toward information that is consistent with whatever racial stereotypes the perceiver holds. When the first person encountered is a Black man, by contrast, the feature of the Black woman that distinguishes her will be her gender. In this case, gender stereotypes may play a larger role in the perceiver's impression of her. In short, the perceiver will attend to the social category that maximally distinguishes the second individual from the first.

## 2.5 Stereotype Strength

The results from the AT and stereotyping studies demonstrated, among other things, that minority group stereotypes carry more weight than majority group stereotypes. The attributes that characterized the minority groups were given greater weight in categorizing ambiguous targets than were the attributes that characterized the majority groups. Specifically, a target possessing the perfect predictors of both the majority and minority groups (PC + PR) was most commonly categorized as a member of the minority group despite the greater frequency of majority group members. Thus, minority stereotypes had greater predictive power than majority stereotypes. It is commonly observed that the most prominent stereotypes in a given society are those that describe minority rather than majority groups. The current results suggest one possible basis for that observation: To distinguish minority groups from majority groups, particular attention is paid to those attributes that permit differentiation of the minority group. In effect, a stronger link may be formed between the minority group and its typical features than between a majority group and its typical features.

The differential strength of majority and minority group stereotypes has additional implications for stereotyping, which [Sacchi \(2015\)](#) explored in his research. For example, if minority group stereotypes are stronger than majority group stereotypes, then it makes sense that minority stereotypes would be more resistant to change than majority group stereotypes. Extensive research shows that people engage in strategies to maintain group stereotypes, even in the face of disconfirming evidence. For example, people may attempt to explain away stereotype-disconfirming behaviors by making situational attributions for those behaviors. If a group member performs a

stereotype-inconsistent behavior, perceivers may attribute that behavior to external, situational factors rather than to stable, dispositional characteristics of the group (Bodenhausen & Wyer, 1985). In addition, group members displaying stereotype-incongruent behavior may be subtyped as unrepresentative of the group (Crocker, Hannah, & Weber, 1983). Subtyping, or “fencing off” atypical group members from the rest of the group, limits the generalization of stereotype-disconfirming information to perceptions of the entire group (Kunda & Oleson, 1995; Maurer, Park, & Rothbart, 1995; Park, Wolsko, & Judd, 2001; Weber & Crocker, 1983).

The research conducted by Sacchi (2015) provides evidence that minority group stereotypes are more resistant to change than majority group stereotypes. After completing the standard learning phase, participants were asked to make trait ratings about the groups and then were 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 vs majority group stereotypes. Moreover, participants rated stereotype-disconfirming minority group members as less typical of their groups than disconfirming majority group members. This finding is consistent with the idea that perceivers subtype disconfirming minority group members more than disconfirming majority group members. Another measure in the study assessed the degree to which participants made situational rather than dispositional attributions for incongruent behaviors, but no differences were observed based on group size. However, this null finding may be explained by the order in which the measures were presented. The items that assessed behavior attribution always came after the items assessing group typicality. Given that participants displayed the expected effect of subtyping on the typicality ratings, it is possible that they were no longer differentially motivated to explain away the stereotype-incongruent behavior of minority targets.

Another implication of stereotype strength is the development of essentialist beliefs about social groups. Essentialism is the idea that an entity such as a social group possesses innate, biologically rooted qualities. People who hold essentialist beliefs about social groups believe that differences between groups are relatively fixed, that social categories are clearly separated and mutually exclusive, and that a person’s characteristics can be inferred based



on their group membership (Rothbart & Taylor, 1992). Research shows that essentialist beliefs are related to greater stereotyping and prejudice (Bastian & Haslam, 2006; Haslam, Rothschild, & Ernst, 2002; Keller, 2005; Levy, Stroessner, & Dweck, 1998), and that those who hold such beliefs are less likely to change their stereotypic beliefs (Bastian & Haslam, 2007). Thus, there is a relation between essentialist beliefs and stereotype strength. Accordingly, people should be more likely to develop essentialist beliefs about minority groups compared to majority groups.

Sacchi (2015) tested the relation between group size and essentialist beliefs. After learning about a majority and minority group, 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: (1) 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); (2) 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 (3) 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). However, participants perceived majority group traits to be just as fixed (i.e., resistant to change) as minority group traits. Overall, the results indicate that minority groups, even novel groups about which we have minimal information, are perceived as having greater essence, and this may be one reason why minority group stereotypes are stronger than majority group stereotypes.

## 2.6 Face Perception and Group Categorization

To this point, we have described research showing that the trait stereotypes of minority groups and their members are weighted more heavily in judgments and are held more strongly than are the trait stereotypes of majority groups. The same 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. AT would suggest that the facial features of racial majorities should be learned before the facial features of racial minorities because racial majorities are more numerous. As a result, minority features are more distinctive, and people should attend more to them, leading to a stronger association between the distinctive facial features and the minority group. Thus, when individuals encounter racially ambiguous individuals who exhibit features of both groups, the features of the minority group should be overweighted and therefore bias classification toward the minority. Obviously, hypodescent has a number of critical political and historical antecedents, and we would not suggest otherwise. However, we were interested in examining the possible role of attentional processes in category learning in producing the phenomenon.

To this end, we conducted two experiments (Halberstadt, Sherman, & Sherman, 2011). In the first experiment, we presented native Chinese and native Caucasian participants in New Zealand with images of Chinese and Caucasian faces that varied systematically in their racial ambiguity. The images ranged from completely unambiguous (100% Chinese or Caucasian) to completely ambiguous (50% Chinese/50% Caucasian; see Fig. 5). The morphed faces were created with a computer program that mathematically averaged the facial contours and grayscale levels in corresponding facial regions of the 100% Chinese and 100% Caucasian faces (Rhodes, Sumich, & Byatt, 1999). 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 in-groups. According to AT, later in life,



**Fig. 5** Example of a Chinese–Caucasian face pair and 4 of its 24 morphed blends (Halberstadt et al., 2011, Experiment 1).

when encountering members of out-groups, participants would have focused attention on those facial features that best distinguished out-group from in-group members, in the same way that people focus on minority group attributes that distinguish them from majority groups. It follows that the facial features of out-group members would be more strongly associated with the out-group than the facial features of in-group members would be associated with the in-group. As such, judgments of ambiguous faces would be influenced more strongly by out-group features, and such targets would be more likely to be assigned to the out-group than the in-group. Consistent with this prediction, we found that Chinese participants were more likely to judge ambiguous faces as Caucasian than were Caucasian participants who, in turn, 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 *in-group overexclusion* (Leyens & Yzerbyt, 1992), which describes the tendency for people to exclude ambiguous individuals from their in-groups in order to protect the distinctiveness of the in-group. In order to rule out such motivational factors, we replicated the first experiment using numerical majority and minority groups of the same race. Because both the majority and minority groups were comprised of monoracial Caucasian faces, we eliminated any motivational factors that could influence group categorization.

Essentially, we replicated the inverse base-rate effect studies we had done with majority and minority groups, swapping facial features for trait descriptions. During a learning phase, participants were presented with the face of a majority group member in 75% of all instances and the face of a minority group member in 25% of all instances. After the learning phase, participants were asked to classify facial morphs of majority and minority group members who 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 study, 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.



### **3. CONTEXT-BASED IMPRESSION FORMATION**

#### **3.1 Context Frequency and Impression Strength**

Up to this point, we have described research showing that AT helps to account for group categorization and stereotype formation. However, AT processes should not be limited to the domain of intergroup perception.

Theoretically, in the same way that perceivers learn how to differentiate people based on the groups that they belong to, they also may learn how to differentiate the behaviors of individuals based on the contexts in which those behaviors occur. Consider a person who behaves differently when he is with his friends vs when he is with his family. If you first meet this person in the context of his friends and only later encounter him in the context of his family, AT suggests that you would 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 would be held more strongly than your impression of him in the friend context. The logic of forming context-based impressions is exactly the same as when forming impressions of majority and minority group members.

We conducted a series of experiments to test this application of AT (Huang, Sacchi, & Sherman, 2017). In the first experiment, participants learned about a target person named Dave and how his traits differed according to context. Just as in the stereotype formation experiments, there was a common trait (PC; presented 75% of the time), a rare trait (PR; presented 25% of the time), and an imperfect predictor trait (I) that was present for all descriptions of the target. Dave was in a different context, in this case, a colored room, depending on the traits he possessed. During learning, participants were presented with different trait descriptions of Dave and had to guess which colored room he was in. For example, participants might have learned that Dave was intelligent and reliable in the blue room, but creative and reliable in the yellow room. After learning the traits associated with each context, participants completed a test phase in which they were presented with different configurations of traits and were asked to indicate which room Dave was in.

Results of the test phase showed that participants successfully learned the perfect predictors associated with each context; when Dave possessed only the common trait, they said that he was in the common context, and when he possessed only the rare trait, they said that he was in the rare context. Moreover, when Dave was described only with the imperfect predictor, participants were more likely to say that he was in the common context than the rare context. This finding shows that they associated the imperfect predictor trait with the common context, consistent with AT. Of most interest, when Dave was presented as having both the common trait and the rare trait, participants were more likely to say that he was in the rare context than the common context. That is, they demonstrated an inverse base-rate effect

**Table 3** Proportion of Each Context Selected for Each Given Predictor Trait(s) (Huang et al., 2017, Experiment 1)

Predictor(s)	Context Chosen	
	Common	Rare
Common trait (PC)*	0.93	0.04
Rare trait (PR)*	0.04	0.92
Imperfect predictor trait (I)*	0.54	0.34
PC + PR*	0.43	0.53

*Note:* Note that the responses of each trait or trait combination do not sum to 1. Participants had completed two replications of the inverse base-rate task; discrepancies in the sums reflect participants' selection from the incorrect context pair.

\* $P < 0.001$ .

(see Table 3). Thus, impressions of the target in a rarely occurring context were stronger than impressions formed in a commonly occurring context. These effects were replicated in a second experiment in which participants formed both positive and negative trait impressions. These data show that the same processes that influence the formation of distinct group stereotypes also can lead to the formation of distinct context-based impressions of individual targets. As in the case of minority group stereotypes, when people learn about a person in a rare (i.e., “minority”) context, the features of that context are weighted more heavily in subsequent judgments than are the features of the common (i.e., “majority”) context.

A key component of AT is that participants learn one impression prior to another. In the first two experiments, we found that participants had learned the common context impressions prior to the rare context impressions. In earlier blocks of learning trials, participants were more accurate on common context trials, but the difference in accuracy between the two contexts decreased in later blocks as participants subsequently learned the rare context impression. Although this finding provides evidence that participants had learned one impression before another, the learning of the two impressions was interwoven, providing a relatively weak test of the role of learning order on the differential strengths of common and rare impressions. Therefore, in a third experiment, we directly manipulated the learning order of context-based trait impressions to test whether the learning order mechanism does, in fact, account for differential impression strength (Huang et al., 2017, Experiment 3). In this experiment, participants formed context-based trait impressions in sequential order. Unlike the other experiments, impressions

occurred with equal frequency. In addition, participants learned about the traits of four different target people (Steve, Bill, Chuck, and Dave) rather than a single target person. As in the inverse base-rate paradigm, each target always possessed two traits—a perfect predictor and an imperfect predictor. In the first half of the learning phase, participants were presented with each target and the traits that the target possessed, and they guessed which of four color rooms the target was in based on the traits. For example, Steve was friendly and practical, and he appeared in the blue room. In the second half of the learning phase, the targets possessed a new combination of traits and now appeared in different color rooms. Each of the four targets possessed a new perfect predictor trait plus the same imperfect predictor trait from the first half of the learning phase. Again, participants were presented with each target person and his traits and guessed in which of the four new color rooms he appeared. For example, Steve was now greedy and practical, and he now appeared in the orange room.

If learning order accounts for the greater strength of rare context-based impressions, then the second-learned impression should be stronger than the first-learned impression, even if both impressions occurred with equal frequency. Thus, when a target possessed both the first-learned trait and the second-learned trait simultaneously, participants should be more likely to classify them as being in the second-learned context. However, the results of the test phase only weakly supported predictions. When a target possessed both perfect predictor traits, participants were more likely to select the second context (40%) than the first context (35%), as predicted, but this difference was not statistically significant. Furthermore, when a target possessed the imperfect predictor trait, participants selected the first and second contexts about equally, even though AT would suggest that they would have been more likely to select the first context. Instead, participants classified the targets in accordance with the base-rate frequencies (i.e., 50:50 base-rate appearance of first-learned and second-learned trait impressions).

The results of the experiment provide some support that learning order may account for the greater strength of rare context-based impressions, but it leaves open the possibility that nonattentional mechanisms also may account for the inverse base-rate effect (e.g., [Juslin, Wennerholm, & Winman, 2001](#); [Winman, Wennerholm, Juslin, & Shanks, 2005](#)). Another possibility is that the experimental procedure was designed such that participants were not directly comparing the two context-based impressions with each other, as they may have been when common and rare impressions were presented together in the same learning blocks. Rather, they may have focused on

differentiating among the four independent targets rather than between the first-learned and second-learned impressions. The task had included four targets rather than one target so as to make the task sufficiently challenging. However, this minor change in procedure may have inadvertently changed the participants' focus of comparison. Another possible explanation is that attention shifting does occur when perceivers learn rare impressions but is not caused by learning order. For example, perceivers' attention may be drawn toward distinctive information such as minority group traits or rare context-based impressions, but not necessarily because they are learned after more common information.

Despite the weak findings from Experiment 3, additional variations of the experiment underscore the robustness of the inverse base-rate effect and, correspondingly, the strength of rare context impressions. In one variation, we reversed the classification order during learning (Huang et al., 2017, Experiment 4). In the previous experiments, participants had formed impressions of the target by guessing which context he was in, given the traits he possessed. However, in real-life situations, people probably are more likely to infer another person's traits based on the context in which that person is encountered than to infer the context based on the traits. For example, people are more likely to infer a person's behavior depending on whether she is with her friends or her family than to infer who she is with depending on her behavior. In this experiment, participants guessed which trait the target possessed, given the contexts in which he appeared. In this case, the traits represented the common (C) and rare (R) events and the contexts represented the perfect (PC, PR) and imperfect (I) predictors of those events. Also, the contexts were represented by people rather than color rooms. For example, participants might have learned that the target Dave was creative when he was with Bob and Chris but honest when he was with John and Chris. As predicted, when Dave was with the imperfect predictor context (e.g., Chris), participants were more likely to say that he possessed the common trait than the rare trait, supporting the hypothesis that shared attributes are more strongly associated with common trait impressions. When Dave was with common and rare context people simultaneously (e.g., Bob and John), participants were more likely to select the rare trait than the common trait, indicating that they had formed a stronger rare context-based impression than common context-based impression.

In another variation of the study, we examined the formation of evaluative impressions rather than trait impressions (Huang & Sherman, 2016). In this experiment, participants learned about positive and negative behaviors

that Dave performed rather than the traits he possessed. During the learning phase, participants guessed which of two behaviors Dave performed—a positive or a negative behavior—depending on the people he was with. Half of the participants formed a positive common impression and negative rare impression (positive common condition), and the other half formed a negative common impression and positive rare impression (negative common condition). During the test phase, participants guessed whether Dave was good or bad depending on the context people he was with.

Again, we replicated the inverse base-rate effect. When Dave was with the common and rare context people together, they were more likely to select the rare evaluation than the common evaluation, indicating a stronger rare context impression, irrespective of the valence of the common impression. Furthermore, when Dave was with the imperfect predictor person, they were more likely to select the common evaluation, consistent with AT. However, the latter finding was stronger when the common evaluation was positive.

### **3.2 Expectancies Shape Context-Based Impression Formation**

Together, our studies show that people form stronger rare context-based impressions than common context-based impressions, at least in part because rare impressions are learned after common impressions. When learning rare impressions, perceivers shift attention toward information that uniquely distinguishes the rare impression from the common impression, resulting in a stronger rare impression. Other important aspects of behavior may also yield differences in the strengths of context-based impressions. A considerable body of research has shown that people pay more attention to behavior that violates an expectancy than to behavior that is consistent with an expectancy, in part, to try to explain the cause of the unexpected events (for reviews, see [Roese & Sherman, 2007](#); [Sherman et al., 2012](#)). One way that people try to understand unexpected behavior is by looking to the context in which it occurred as a potential explanation ([Gawronski, Ye, Rydell, & De Houwer, 2014](#)). For example, if a perceiver holds an expectation that a target person is intelligent, and they then observe the target acting unintelligently, the perceiver pays especially close attention to the context in which the unintelligent behavior occurred as a way of understanding it. As in the case of greater attention being directed at behavior in rarely occurring contexts, greater attention to contexts in which unexpected behavior occurs may lead to stronger impressions of behavior in those contexts than of behavior in contexts associated with expected behavior. However, although people may form stronger

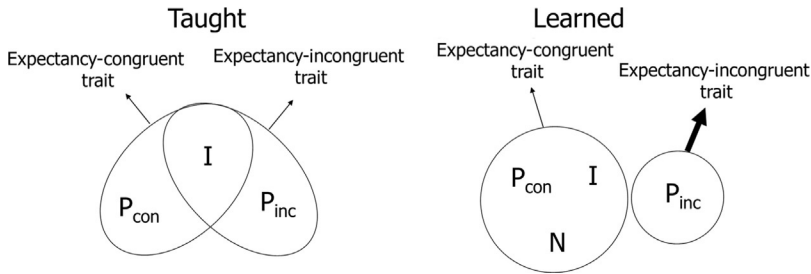


impressions in contexts in which unexpected behavior occurs, this process may limit the generalization of that behavior. Specifically, the unexpected behavior may affect impressions only within the context in which it was observed.

Research conducted by Gawronski and colleagues lends support to this hypothesis (Gawronski, Hu, Rydell, Vervliet, & De Houwer, 2015; Gawronski, Rydell, Vervliet, & De Houwer, 2010; Rydell & Gawronski, 2009). In their studies, they show that the generalization and contextualization of evaluative impressions depend on the order in which they were formed. When a target individual behaves differently across contexts (e.g., positively in one context and negatively in another context), perceivers generalize the first-learned evaluation of the target to new contexts and contextualize the second-learned evaluation so that it is elicited only in its learned context. The researchers suggest that the second-learned (i.e., counterattitudinal) evaluation of the target is contextualized because it violates expectations. Consequently, perceivers search for contextual cues that may explain the discrepancy.

Based on AT, we similarly predicted that perceivers would generalize expectancy-congruent impressions to new contexts but contextualize expectancy-incongruent impressions, thereby limiting the influence of incongruent behaviors in changing the overall impression of the target. We examined these questions across a series of experiments (Huang & Sherman, 2018). First, we tested the novel prediction that participants would form stronger impressions of others in contexts in which unexpected behavior occurs than in contexts in which expected behavior occurs. Just as people shift attention to contexts that predict rarely occurring impressions, they also should shift their attention to contexts that predict expectancy-incongruent behaviors as a means of differentiating the incongruent impression from the congruent impression. Second, we hypothesized that participants would contextualize incongruent impressions into the unique contexts in which they are formed, but that they would generalize congruent impressions across all other contexts, included new contexts and contexts shared between the two impressions, thereby allowing them to maintain their original expectations of the target individual (see Fig. 6).

In the first experiment, we manipulated trait expectancy by exposing participants to a paragraph describing the target Dave as either intelligent or unintelligent. Participants then completed a learning phase in which they learned about the intelligent and unintelligent behaviors that Dave performed depending on the people he was with (i.e., the context). Thus, the behaviors



**Fig. 6** Conceptualization of expectancy-congruent and expectancy-incongruent impressions, adapted from the inverse base-rate design. *Left:* Depiction of the target's behaviors that participants are taught during the learning phase.  $P_{con}$  is the perfect predictor context of the expectancy-congruent trait,  $P_{inc}$  is the perfect predictor context of the expectancy-incongruent trait, and  $I$  is the imperfect predictor context that occurs for both the expectancy-congruent and expectancy-incongruent traits. *Right:* Depiction of learned impressions.  $N$  represents new contexts in which the target is encountered. We predicted that perceivers form a stronger association between the expectancy-incongruent impression and its unique predictor context (i.e., a stronger expectancy-incongruent impression). However, we predicted that they hold an expectancy-congruent impression in all other contexts, including  $P_{con}$ ,  $I$ , and  $N$ .

that Dave performed were either congruent or incongruent with their prior expectations. Unlike our previous experiments, each behavior type occurred with equal frequency. One context person always predicted the expectancy-congruent behavior (perfect predictor of congruent behavior, or  $P_{con}$ ), another context person always predicted the expectancy-incongruent behavior (perfect predictor of incongruent behavior, or  $P_{inc}$ ), and a third context person was present in all cases (imperfect predictor, or  $I$ ). As in our previous experiments, participants were presented with Dave and guessed which behavior he performed depending on the people he was with. In each trial, Dave was always with a combination of two people—a perfect predictor person plus the imperfect predictor person ( $P_{con} + I$  or  $P_{inc} + I$ ). In a subsequent test phase, Dave was presented with new combinations of context people, and participants indicated whether he was intelligent or unintelligent depending on the context.

As predicted, the results showed that participants formed stronger impressions in contexts associated with unexpected than expected behavior from Dave. When Dave was with  $P_{con}$  and  $P_{inc}$  together, participants were more likely to say that he possessed the incongruent trait than the congruent trait, indicating that they had formed a stronger impression in the context associated with unexpected behavior. This finding also suggests that participants had attended more to the context of the incongruent behaviors when learning about those behaviors. Second, the results showed that participants

had contextualized the expectancy-incongruent impression into its unique perfect predictor context: When Dave was in the imperfect predictor context, participants were more likely to say that he possessed the expectancy-congruent trait than the expectancy-incongruent trait. Even though the shared context person was present in all encounters with Dave, participants had disassociated the incongruent impression from that context, thereby isolating that impression to a single context. These results cannot be attributed to different base-rate frequencies during the learning phase because each type of behavior had occurred with equal frequency. Last, the results showed that participants had generalized the expectancy-congruent impression to new contexts. When Dave was with new people, participants were more likely to say that Dave possessed the congruent trait than the incongruent trait, indicating that their impression of him in new contexts was consistent with their original trait expectation. Overall, the results suggest that participants had formed context-based impressions that limited the generalization of incongruent behaviors and maintained original expectations of the target (see Table 4).

In the next study, we examined whether similar effects would be observed when expectancies were based on group stereotypes rather than individual trait expectations. On the one hand, people engage in similar expectancy maintenance processes when they try to make sense of stereotype-incongruent behaviors as they do with trait-incongruent behaviors (Sherman et al., 2012). For example, people tend to make situational attributions for stereotype-disconfirming behaviors but dispositional attributions for stereotype-confirming behaviors (Crocker et al., 1983; Deaux & Emswiller, 1974;

**Table 4** Proportion of Each Trait Selected for Each Given Predictor Context(s) (Huang & Sherman, 2018, Experiment 1)

Predictor(s)	Trait Chosen	
	Congruent	Incongruent
P <sup>**</sup> <sub>con</sub>	0.97	0.03
P <sup>**</sup> <sub>inc</sub>	0.02	0.98
P <sub>con</sub> + P <sub>inc</sub> <sup>*</sup>	0.45	0.55
I <sup>**</sup>	0.67	0.39
N <sup>**</sup>	0.62	0.38

*Note:* I=imperfect predictor context, N=new contexts, P<sub>con</sub>= context that perfectly predicts the expectancy-congruent behaviors, and P<sub>inc</sub>= context that perfectly predicts the expectancy-incongruent behaviors.

<sup>\*</sup>P<0.01.

<sup>\*\*</sup>P<0.001.

Feldman-Summers & Kiesler, 1974; Jackson, Sullivan, & Hodge, 1993; Sherman, Stroessner, Conrey, & Azam, 2005; Yee & Eccles, 1988), leading them to maintain expectations of future stereotypic behaviors (Bodenhausen & Wyer, 1985). They also tend to be more critical of stereotype-disconfirming information than stereotype-confirming information (Macrae, Shepherd, & Milne, 1992). Highly prejudiced individuals, in particular, may scrutinize stereotype-inconsistent information in an attempt to explain it away (Sherman et al., 2005). On the other hand, stereotypic expectancies may not have the same effects as individual expectancies on impressions because people tend to expect less consistency among behaviors performed by different group members than among behaviors performed by a single individual (Hamilton & Sherman, 1996). Consequently, people may feel less need to resolve behavioral inconsistencies among multiple group members (Stern, Marrs, Millar, & Cole, 1984; Susskind, Maurer, Thakkar, Hamilton, & Sherman, 1999).

We tested these competing hypotheses in a second experiment (Huang & Sherman, 2018). The second experiment was similar to the first, except that expectancies were based on group stereotypes rather than individual trait expectancies. Participants read descriptions of a stereotypically extraverted group (sales professionals) and a stereotypically introverted group (writing professionals) and were then told that they would learn about a person named Steve who was either a member of the sales group (extraverted stereotype condition) or a member of the writers group (introverted stereotype condition). They then completed a learning phase and a test phase that were similar to those in the first experiment.

Participants did not form a stronger impression in the context associated with unexpected behavior. When Steve was with both perfect predictor contexts ( $P_{\text{con}} + P_{\text{inc}}$ ), participants were equally likely to say that he possessed the stereotype-congruent trait and the stereotype-incongruent trait, indicating that they did not attend more to the context that uniquely predicted the stereotype-incongruent behaviors. This finding supports previous research suggesting that people are less inclined to resolve behavioral inconsistencies among multiple group members compared to behavioral inconsistencies within a single individual (Stern et al., 1984; Susskind et al., 1999). However, our results did show that participants had contextualized the stereotype-incongruent impression into its unique predictor context, just as participants in the first experiment had contextualized the expectancy-incongruent impression. When Steve was with the imperfect

**Table 5** Proportion of Each Trait Selected for Each Given Predictor Context(s) (Huang & Sherman, 2018, Experiment 2)

**Trait Chosen**

Predictor(s)	Congruent	Incongruent
$P_{con}^*$	0.91	0.09
$P_{inc}^*$	0.08	0.92
$P_{con} + P_{inc}$	0.49	0.51
$I^*$	0.55	0.45
$N^*$	0.64	0.36

*Note:*  $I$ =imperfect predictor context,  $N$ =new contexts,  $P_{con}$ =context that perfectly predicts the stereotype-congruent behaviors, and  $P_{inc}$ =context that perfectly predicts the stereotype-incongruent behaviors.  
\* $P < 0.001$ .

predictor context ( $I$ ), participants selected the stereotype-congruent trait significantly more often than the stereotype-incongruent trait, thereby isolating the stereotype-incongruent impression to a single context. Furthermore, participants generalized the stereotype-congruent impression to new contexts in which Steve had not been encountered previously. When Steve was with novel context people ( $N$ ), they selected the stereotype-congruent trait significantly more often than the stereotype-incongruent trait (see Table 5). Together, these findings indicate that participants had maintained a stereotype-congruent impression of Steve by contextualizing the incongruent impression into a single context and generalizing the congruent impression across all other contexts.

The processes observed in this experiment are similar to other stereotype maintenance processes in that stereotype-incongruent behaviors have a weaker influence on group impressions than do stereotype-congruent behaviors. Even though people may attend more to incongruent behaviors, they do so because the behaviors are unexpected, not because they are integrating those behaviors into their original impression of the group. For example, people may scrutinize incongruent behaviors and make sense of them by generating situational attributions, but they do not change their original impressions as a result of those behaviors (for a review, see Sherman et al., 2012). People also may resolve stereotype incongruency at the level of the individual by subtyping, or separating, stereotype-disconfirming group members into a subcategory of the group (Richards & Hewstone, 2001; Weber & Crocker, 1983).

Because these individuals are viewed as atypical group members, their behaviors fail to generalize to impressions of the whole group. Although these cognitive processes allow perceivers to maintain coherent impressions of social groups, they promote stereotype maintenance.

We examined the extent to which context-based impressions of a single group member transfer to other individuals, as opposed to being set aside or subtyped (Huang & Sherman, 2018, Experiment 2). After completing the first test phase about Steve, participants completed a second test phase that assessed their impressions of a target who belonged to the same group as Steve (in-group target) and a target who belonged to the other group (out-group target). For example, if Steve was a member of the sales group, then the in-group target also was a member of the sales group, and the out-group target was a member of the writers group. Conversely, if Steve was a member of the writers group, then the in-group target also was a member of the writers group, and the out-group target was a member of the sales group. The in-group and out-group targets were presented individually with the same context people with whom Steve had appeared during the first test phase, and participants selected which trait each target possessed for each context person or context people they were with.

The results did, in fact, show that context-based impressions can transfer to other individuals, but only when those individuals are encountered in a counterstereotypic context. In all other contexts, participants held impressions of the in-group and out-group targets that were consistent with the stereotypes of their respective groups. Impressions of the in-group target were consistent with the in-group stereotype in all contexts except for  $P_{inc}$ , the context in which Steve had performed counterstereotypic behaviors. When the in-group target was with  $P_{con}$ ,  $P_{con} + P_{inc}$ , the imperfect predictor person (I), or new context people (N), participants selected the stereotype-congruent trait significantly more often than the stereotype-incongruent trait. However, when the in-group target was with  $P_{inc}$ , participants selected the stereotype-incongruent trait significantly more often than the stereotype-congruent trait, indicating that contextual associations with Steve's stereotype-incongruent behaviors translated to a stereotype-incongruent impression of a fellow in-group member in the same context.

Impressions of the out-group target were consistent with the out-group stereotype in all contexts except for  $P_{con}$ , the context associated with the opposing in-group's stereotypic behaviors. In this context, participants demonstrated no dominant trait impression. For a comparison of the participants'

**Table 6** Impressions of the Main Target, In-group Target, and Out-group Target Within Each Given Context (Huang & Sherman, 2018, Experiment 2)

Context	Impression Target		
	Main Target (Steve)	In-group Member	Out-group Member
$P_{con}$	■	■	=
$P_{inc}$	◇	◇	◇
$P_{con} + P_{inc}$	=	■	◇
I	■	■	◇
N	■	■	◇

*Note:* I = imperfect predictor context, N = new contexts,  $P_{con}$  = context that perfectly predicts stereotypic in-group behaviors, and  $P_{inc}$  = context that perfectly predicts stereotypic out-group behaviors. The *square* (■) represents greater selection of the in-group stereotype, the *diamond* (◇) represents greater selection of the out-group stereotype, and the *equal sign* (=) represents equal selection of each trait.

impressions of the main target Steve, the in-group target, and the out-group target, see Table 6.

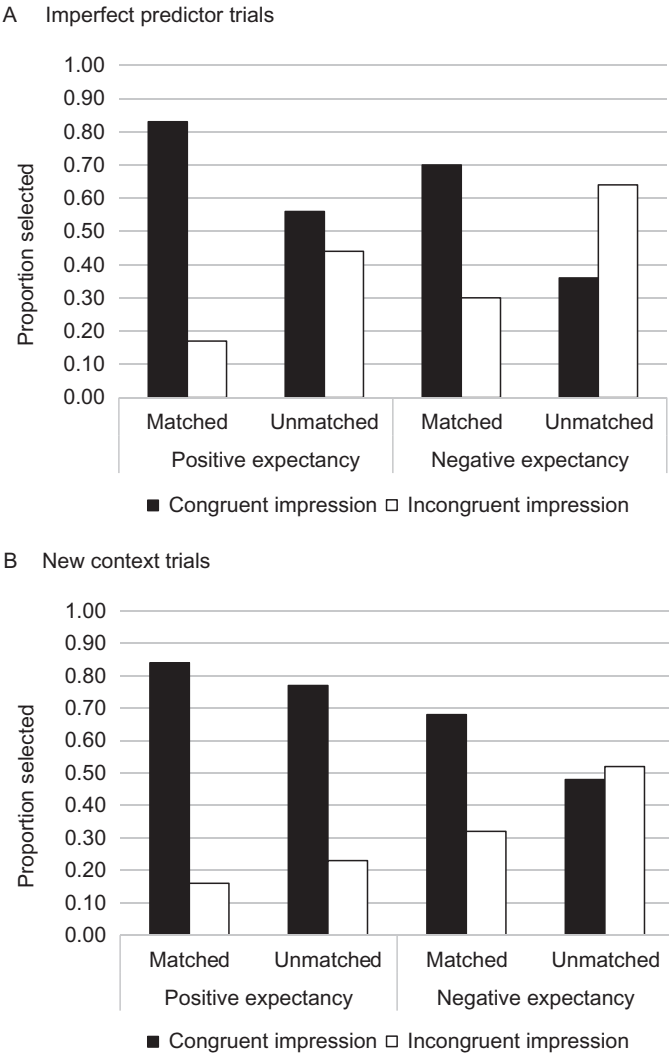
### 3.3 Changing Impressions

The experiments outlined in Section 3.2 show that people’s expectancies can bias context-based impressions toward expectancy maintenance. Although participants had attended more to expectancy-incongruent behaviors (at least when expectancies were of a single individual), they had contextualized those behaviors to their unique predictor context and, instead, generalized the behaviors that confirmed their prior expectations. This was true whether the expectation was based on an individual trait expectancy or on a group-based stereotype. These findings beg the question—can impressions change if a person’s behaviors do not match with initial expectations? Across two experiments, we examined how behavioral patterns change perceivers’ impressions of an individual (Huang & Sherman, 2018, Experiments 3a and 3b). In one experiment, all participants held a positive expectancy of a target person named Dave, and in the other experiment, all participants held a negative expectancy of Dave. After forming the expectancy, participants completed a learning phase in which they learned about the good and bad behaviors that Dave performed, depending on the context people he was with. Dave performed each behavior type with different frequencies. For half of the participants, he performed expectancy-congruent behaviors in 75% of all instances and expectancy-incongruent behaviors in the other 25% of

instances (matched condition). For the other half of participants, he performed expectancy-incongruent behaviors in 75% of all instances and expectancy-congruent behaviors in the other 25% of instances (unmatched condition). Thus, in the matched condition, the target's behavioral pattern matched with participants' expectations because the target performed mostly expectancy-confirming behaviors, whereas in the unmatched condition, the target's behavioral pattern did not match with participants' expectations because the target performed mostly expectancy-disconfirming behaviors. The remaining aspects of the learning phase were the same as the two previously described experiments. In a subsequent test phase, participants were presented with new configurations of the contexts and indicated whether Dave was good or bad for each person or combination of people he was with.

Overall, participants partially updated their impressions when the target's behavioral patterns did not match with their original expectations. First, when the target's behavioral pattern matched with expectations (i.e., the more frequently occurring behaviors were consistent with expectations), participants held an expectancy-congruent impression of him in the imperfect predictor context (i.e., the context shared between the two behavior types) and in new contexts. However, when the target's behavioral pattern did not match with expectations (i.e., the more frequently occurring behaviors were not consistent with expectations), these effects were weakened or reversed to be more in line with the target's behaviors (see [Fig. 7](#)). Interestingly, participants also demonstrated a preference for a positive impression, overall. These experiments provide preliminary evidence that perceivers can update their impressions if their original expectations do not accurately reflect the target's behaviors. However, the participants did not fully adjust their impressions away from original expectancies, so it is not clear if the effects of expectancies can be eliminated entirely. Research on anchoring and adjustment ([Epley & Gilovich, 2006](#); [Kruglanski & Freund, 1983](#)) and group expectancies ([Ottati, Claypool, & Gingrich, 2005](#)) show that people anchor their impressions onto their initial expectancies and may not fully adjust away from those expectancies when learning new information about an impression target. However, in the current experiment, all expectancy-disconfirming behaviors occurred within a single context. Changing impressions may be more effective when a target is observed performing expectancy-incongruent behaviors across multiple contexts because impressions could then be generalized broadly across contexts.





**Fig. 7** (A and B). Impressions of the target with (A) the imperfect predictor context and (B) new context people (Huang & Sherman, 2018, Experiments 3a and 3b). When the target’s behaviors matched expectations, participants selected the expectancy-congruent impression significantly more often than the expectancy-incongruent impression. This impression was weakened or reversed when the target’s behaviors did not match expectations.



## 4. CONCLUSION

In this chapter, we described how a simple attentional mechanism can account for a wide variety of phenomena in social perception. Specifically, we described how AT (Kruschke, 1996, 2001) can explain phenomena in stereotype formation, face perception and group categorization, context-based impression formation, and expectancy maintenance. AT rests on the idea that people preferentially attend to differentiating information; when people learn about multiple social categories, they shift their attention toward the unique features of the category that they learn last, resulting in a stronger association between that category and its unique attributes. These processes have interesting implications for impression formation.

First, we discussed how attention influences the ways that perceivers form stereotypes of majority and minority group members. We showed that AT can account for two prominent findings in the stereotype formation literature—category accentuation and illusory correlation (Sherman et al., 2009). Whereas category accentuation effects highlight the exaggeration of real intergroup differences as the basis for stereotype formation, illusory correlation shows that stereotypes may be formed in the absence of real group differences. In both cases, however, minority group attributes are learned after the majority group attributes because encounters with minority group members occur less frequently. When learning about the minority group, perceivers shift their attention to attributes that best differentiate it from the majority group, regardless of whether these attributes accurately represent the minority group or not. As a result, people form stronger stereotypes of the minority group. Follow-up studies provided additional support for the strength of minority group stereotypes. In these studies, participants held stronger essentialist beliefs about minority groups and were less likely to change minority group stereotypes in the face of disconfirming evidence (Sacchi, 2015).

Next, we explained how attention shifting can influence face perception and group categorization. Specifically, we explained how it can account for hypodescent, the tendency to categorize mixed-race faces as belonging to the minority group than the majority group (Halberstadt et al., 2011). AT would suggest that the facial features of racial majorities are learned before the facial features of racial minorities because racial majorities are encountered more frequently. When learning the facial features of the minority group, perceivers should focus their attention on features that best

distinguish the minority faces from the majority faces. As a result, they should form a stronger association between the differentiating facial features and the minority group. Thus, when individuals encounter racially ambiguous individuals who exhibit features of both groups, the features of the minority group should be weighted more heavily when identifying racial group membership. We provided evidence from our studies supporting this account.

Third, we provided support for the role of AT in the formation of context-based impressions using the inverse base-rate paradigm (Huang et al., 2017; Huang & Sherman, 2016). In the same way that perceivers learn how to differentiate people based on the groups that they belong to, they also learn how to differentiate their impressions of an individual based on the contexts in which the individual's behaviors are observed. Because they occur more frequently, impressions of a target in a commonly occurring context are learned prior to impressions in a rarely occurring context. When forming an impression in a rare context, perceivers shift their attention toward the features that uniquely distinguish that context from the impression formed in the common context, resulting in a stronger rare context impression, as demonstrated by the inverse base-rate effect.

Last, we described how expectancies shape context-based impression formation (Huang & Sherman, 2018). We showed that perceivers form stronger contextualized impressions from expectancy-incongruent behaviors compared to expectancy-congruent behaviors, indicating that incongruent behaviors draw more attention to the context that uniquely predicts those behaviors. However, this was true only for individual target expectancies and not for stereotypic group expectancies. We also showed that perceivers may use contexts as a means for maintaining their original expectancies. For both an individual target and a group member target, perceivers generalized the expectancy-congruent impression across multiple contexts, whereas they contextualized the expectancy-incongruent impression to the unique context in which it was formed, thereby limiting its influence in changing overall impressions of the target. Furthermore, we showed that impressions of single group members translate only weakly to other group members. Context-based impressions of an in-group and an out-group were based primarily on the stereotypes of their groups rather than individual group members, except in the two unique predictor contexts for which there already was a strong, unambiguous association with a trait impression. Finally, we showed that people can update their impressions of an individual if that individual's pattern of behaviors does not align with original

expectations. However, people may not fully adjust away from the expectancy, so the effects of expectancies may not be eliminated entirely.

The research we have reviewed has important implications for impression change. Throughout, we showed how a simple attentional preference for distinctive information influences our impressions of individuals and social groups. Are there ways to encourage impression change or prevent expectancy-biased impression formation from occurring in the first place? One option may be to focus on commonalities rather than differences. Although focusing on differentiating features of minority groups facilitates learning, it results in exaggerated judgments of minority group traits. Instead, perceivers could switch their focus of attention toward traits that are shared by both groups. Theoretically, this strategy should weaken minority group stereotypes and make them more amenable to change and less subject to essentialist beliefs (Sacchi, 2015). Focusing on commonalities also has been shown to reduce intergroup bias (for a review, see Gaertner & Dovidio, 2009). The downside of this strategy is that perceivers may not learn the social categories as well as when they focus on differentiating features. In some cases, though, this may be desirable. If perceived differences between groups are an exaggeration of reality (as in the case of category accentuation) or if differences between groups simply do not exist (as in the case of illusory correlation), then blurring the boundaries between groups would be a more accurate representation of reality. If perceivers do focus on differentiating features of the minority group rather than on common features shared between groups, then they could take the extra step of recalling how the majority group's traits differ from the minority group's traits. However, this strategy would require more effort and may not be practical if the perceiver is not aware of the ways that they are forming group impressions or if they simply are not motivated to be accurate.

Another option is to change the standard of comparison. As the first-learned group, the majority group is the standard (i.e., the "default") against which the minority group is compared. Differences between groups are exaggerated because perceivers focus on characteristics that make the minority group different from the majority group. Research similarly has shown that when perceivers use the in-group as the standard of comparison for the out-group, judgments of an out-group's trait level are typically contrasted away from the in-group's level of the same trait (Gawronski, Bodenhausen, & Banse, 2005). Instead of comparing the minority group to the majority group, perceivers could compare the minority group against other standards. For example, perceivers could compare individual group

members to other members of the same group. Comparing individual group members to each other may combat perceptions of out-group homogeneity and decrease the likelihood of stereotyping the group, particularly if the group is in the numerical minority (Simon, 1992; Simon & Mummendey, 1990). Another strategy is to compare a minority group against multiple groups rather than a single majority group. Although social psychological research on intergroup perceptions typically focuses on the relationship between two groups (e.g., an in-group vs an out-group or a majority vs minority group), this dichotomy is not always realistic. People are members of multiple racial, ethnic, religious, and occupational social groups, among others. By comparing against multiple groups, perceivers may focus less on individual characteristics that distinguish one group from another and more on multiple characteristics that can be compared against multiple groups.

Our research also has important implications for attitude change. Consider the following case. In our society, attitudes toward racial minorities, particularly African Americans, are predominantly negative. However, research shows that automatically activated attitudes toward African Americans change as a function of the context (Allen, Sherman, & Klauer, 2010; Barden, Maddux, Petty, & Brewer, 2004; Maddux, Barden, Brewer, & Petty, 2005; Wittenbrink, Judd, & Park, 2001). In one study, participants expressed anti-Black bias when targets were presented in a negative context (e.g., a ghetto street corner), but no bias whatsoever when they were presented in a positive context (e.g., a church interior; Wittenbrink et al., 2001). Considering that standard, context-free measures of implicit bias have consistently demonstrated anti-Black bias (e.g., Devine, Plant, Amodio, Harmon-Jones, & Vance, 2002; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; Monteith, Voils, & Ashburn-Nardo, 2001), people seem to hold a generalized negative attitude toward African Americans, whereas they hold a contextualized positive attitude that appears only when Black targets are presented in positive contexts. The challenge here is to determine how people can “decontextualize” the positive attitude (i.e., form a context-free representation of the positive impression) in order to override the predominant, generalized negative attitude (for further review of contextualized attitude change, see Gawronski et al., 2018). AT suggests one means for doing so. In our research, we showed that perceivers may pay particularly close attention to the context in which a target is performing counter-expectational behaviors. As a result, they may contextualize this impression and hold it more strongly than impressions drawn from expected

behaviors. One important implication is that, if the counter-expectational impression (e.g., a positive impression of a Black target) can be introduced across multiple contexts, it may undermine or even overwhelm the expected impression (e.g., a negative impression of a Black target). One of the best ways to form a generalized positive impression is, perhaps, to increase contact with negatively stereotyped group members in multiple contexts but to do so only in those contexts in which positive experiences are most likely to occur (Pettigrew & Tropp, 2006, 2008).

Altogether, our research on attentional processes in social perception demonstrates the potential power of a very simple idea. The fundamental message of AT is that what people learn depends on what they already know, 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 a variety of outcomes, some of which yield inaccurate knowledge: the inverse base-rate effect, category accentuation processes, the formation of illusory correlations, the relative strength of minority vs majority stereotypes, judgments of hypodescent, the relative strength of context-based impressions, and the formation of contextualized impressions of other people. We hope this research provides a foundation for further exploration in social perception.

## ACKNOWLEDGMENT

Preparation of this chapter was partially supported by an Anneliese Maier Research Award from the Alexander von Humboldt Foundation to J.W.S.

## REFERENCES

- Allen, T. J., Sherman, J. W., & Klauer, K. C. (2010). Social context and the self-regulation of implicit bias. *Group Processes and Intergroup Relations*, 13(2), 137–149.
- 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.
- Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, 87(1), 5–22.
- Bastian, B., & Haslam, N. (2006). Psychological essentialism and stereotype endorsement. *Journal of Experimental Social Psychology*, 42, 228–235.
- Bastian, B., & Haslam, N. (2007). Psychological essentialism and attention allocation. *Journal of Social Psychology*, 147, 531–541.
- Bastian, B., & Haslam, N. (2008). Immigration from the perspective of hosts and immigrants: Roles of psychological essentialism and social identity. *Asian Journal of Social Psychology*, 11, 127–140.

- Bodenhausen, G. V., & Wyer, R. S. (1985). Effects of stereotypes on decision making and information-processing strategies. *Journal of Personality and Social Psychology*, 48(2), 267–282.
- Corneille, O., & Judd, C. M. (1999). Accentuation and sensitization effects in the categorization of multifaceted stimuli. *Journal of Personality and Social Psychology*, 77(5), 927–941.
- Crocker, J., Hannah, D. B., & Weber, R. (1983). Person memory and causal attributions. *Journal of Personality and Social Psychology*, 44(1), 55–66.
- Deaux, K., & Emswiller, T. (1974). Explanations of successful performance on sex-linked tasks: What is skill for the male is luck for the female. *Journal of Personality and Social Psychology*, 29(1), 80–85.
- Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L. (2002). The regulation of explicit and implicit race bias: The role of motivations to respond without prejudice. *Journal of Personality and Social Psychology*, 82(5), 835–848.
- Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, 33(5), 510–540.
- Eiser, J. R. (1971). Enhancement of contrast in the absolute judgment of attitude statements. *Journal of Personality and Social Psychology*, 17(1), 1–10.
- Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic: Why the adjustments are insufficient. *Psychological Science*, 17(4), 311–318.
- Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, 146(3644), 668–670.
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69(6), 1013–1027.
- Feldman-Summers, S., & Kiesler, S. B. (1974). Those who are number two try harder: The effect of sex on attributions of causality. *Journal of Personality and Social Psychology*, 30(6), 846–855.
- 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.
- Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38(6), 889–906.
- Fiske, S. T., Lin, M., & Neuberg, S. L. (1999). The continuum model: Ten years later. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 231–254). New York: Guilford Press.
- 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.
- Gaertner, S. L., & Dovidio, J. F. (2009). A common ingroup identity: A categorization-based approach for reducing intergroup bias. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 489–505). New York: Psychology Press.
- Gawronski, B., Bodenhausen, G. V., & Banse, R. (2005). We are, therefore they aren't: Ingroup construal as a standard of comparison for outgroup judgments. *Journal of Experimental Social Psychology*, 41(5), 515–526.
- Gawronski, B., Rydell, R. J., Vervliet, B., & De Houwer, J. (2010). Generalization versus contextualization in automatic evaluation. *Journal of Experimental Psychology: General*, 139(4), 683–701.
- Gawronski, B., Ye, Y., Rydell, R. J., & De Houwer, J. (2014). Formation, representation, and activation of contextualized attitudes. *Journal of Experimental Social Psychology*, 54, 188–203.
- Gawronski, B., Hu, X., Rydell, R. J., Vervliet, B., & De Houwer, J. (2015). Generalization versus contextualization in automatic evaluation revisited: A meta-analysis of successful and failed replications. *Journal of Experimental Psychology: General*, 144(4), e50–e64.

- Gawronski, B., Rydell, R. J., De Houwer, J., Brannon, S. M., Ye, Y., Vervliet, B., & Hu, X. (2018). Contextualized attitude change. *Advances in Experimental Social Psychology*, 57, 1–52.
- Gottlieb, J., Oudeyer, P., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, 17(11), 585–593.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480.
- Halberstadt, J., Sherman, S. J., & Sherman, J. W. (2011). Why Barack Obama is Black: A cognitive account of hypodescent. *Psychological Science*, 22(1), 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.
- Hamilton, D. L., & Sherman, J. W. (1996). Perceiving persons and groups. *Psychological Review*, 103(2), 336–355.
- Haslam, N., Rothschild, L., & Ernst, D. (2002). Are essentialist beliefs associated with prejudice? *British Journal of Social Psychology*, 41, 87–100.
- Hilton, J. L., Klein, J. G., & von Hippel, W. (1991). Attention allocation and impression formation. *Personality and Social Psychology Bulletin*, 17(5), 548–559.
- Huang, L. M., & Sherman, J. W. (2016). *Context-based evaluative impressions* [Unpublished raw data]. Davis, CA: Department of Psychology, University of California.
- Huang, L. M., & Sherman, J. W. (2018). *Person expectancies and the formation of context-based impressions* [Manuscript submitted for publication]. Davis, CA: Department of Psychology, University of California.
- Huang, L. M., Sacchi, D. L. M., & Sherman, J. W. (2017). On the formation of context-based person impressions. *Journal of Experimental Social Psychology*, 68, 146–156.
- Jackson, L. A., Sullivan, L. A., & Hodge, C. N. (1993). Stereotype effects of attributions, predictions, and evaluations: No two social judgments are quite alike. *Journal of Personality and Social Psychology*, 65(1), 69–84.
- Juslin, P., Wennerholm, P., & Winman, A. (2001). High-level reasoning and base-rate use: Do we need cue-competition to explain the inverse base-rate effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(3), 849–871.
- Keller, J. (2005). In genes we trust: The biological component of psychological essentialism and its relationship to mechanisms of motivated social cognition. *Journal of Personality and Social Psychology*, 88, 686–702.
- 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(1), 35–47.
- Krueger, J., & Rothbart, M. (1990). Contrast and accentuation effects in category learning. *Journal of Personality and Social Psychology*, 59(4), 651–663.
- Krueger, J., Rothbart, M., & Sriram, 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.
- Kruglanski, A., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impression primacy, ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology*, 19, 448–468.
- 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(5), 171–175.
- Kruschke, J. K., Kappenman, E. S., & Hetrick, W. P. (2005). Eye gaze and individual differences consistent with learned attention in associative blocking and highlighting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(5), 830–845.
- Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, 68(4), 565–579.
- Levy, S. R., Stroessner, S. J., & Dweck, C. S. (1998). Stereotype formation and endorsement: The role of implicit theories. *Journal of Personality and Social Psychology*, 74, 1421–1436.
- Leyens, J.-P., & Yzerbyt, V. Y. (1992). The ingroup overexclusion effect: Impact of valence and confirmation on stereotypical information search. *European Journal of Social Psychology*, 22(6), 549–569.
- Luque, D., Vadillo Nistal, M. A., Le Pelley, M. E., & Beesley, T. (2017). Prediction and uncertainty in associative learning: Examining controlled and automatic components of learned attentional biases. *Quarterly Journal of Experimental Psychology*, 70(8), 1485–1503.
- Macrae, C. N., Shepherd, J. W., & Milne, A. B. (1992). The effects of source credibility on the dilution of stereotype-based judgments. *Personality and Social Psychology Bulletin*, 18(6), 765–775.
- Maddux, W. W., Barden, J., Brewer, M. B., & Petty, R. E. (2005). Saying no to negativity: The effects of context and motivation to control prejudice on automatic evaluative responses. *Journal of Experimental Social Psychology*, 41, 19–35.
- Matsuka, T., & Corter, J. E. (2008). Observed attention allocation processes in category learning. *The Quarterly Journal of Experimental Psychology*, 61(7), 1067–1097.
- Maurer, K. L., Park, B., & Rothbart, M. (1995). Subtyping versus subgrouping processes in stereotype representation. *Journal of Personality and Social Psychology*, 69(5), 812–824.
- McGarty, C., & Penny, R. E. C. (1988). Categorization, accentuation, and social judgement. *British Journal of Social Psychology*, 27(2), 147–157.
- McGarty, C., & Turner, J. C. (1992). The effects of categorization on social judgment. *British Journal of Social Psychology*, 31, 253–268.
- McGarty, C., Haslam, S. A., Turner, J. C., & Oakes, P. J. (1993). Illusory correlation as accentuation of actual intercategory difference: Evidence for the effect with minimal stimulus information. *European Journal of Social Psychology*, 23, 391–410.
- 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.
- Monteith, M. J., Voils, C. I., & Ashburn-Nardo, L. (2001). Taking a look underground: Detecting, interpreting, and reacting to implicit racial biases. *Social Cognition*, 19(4), 395–417.
- Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(1), 104–114.
- Ottati, V., Claypool, H. M., & Gingrich, B. (2005). Effects of a group stereotype on memory for behaviors performed by a group member. *European Journal of Social Psychology*, 35, 797–808.
- Park, B., Wolsko, C., & Judd, C. M. (2001). Measurement of subtyping in stereotype change. *Journal of Experimental Social Psychology*, 37(4), 325–332.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783.
- Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, 38, 922–934.
- Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, 61(3), 380–391.

- Prazienkova, M., Paladino, M. P., & Sherman, S. J. (2017). On the cognitive determinants of out-group dehumanization: Illusory correlation and the dehumanization of (numerical) group minorities. *Social Cognition*, 35(6), 639–662.
- 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.
- Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews. Neuroscience*, 4, 193–202.
- Rehder, B., & Hoffman, A. B. (2005). Eye tracking and selective attention in category learning. *Cognitive Psychology*, 51(1), 1–41.
- Rhodes, G., Sumich, A., & Byatt, G. (1999). Are average facial configurations attractive only because of their symmetry? *Psychological Science*, 10, 52–58.
- Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, 5, 52–73.
- Roese, N. J., & Sherman, J. W. (2007). Expectancies. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 91–115). New York: Guilford Press.
- Rothbart, M. (1981). Memory and social beliefs. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup relations* (pp. 145–181). Hillsdale, NJ: Erlbaum.
- Rothbart, M., & Taylor, M. (1992). Category labels and social reality: Do we view social categories as natural kinds? In G. R. Semin & K. Fiedler (Eds.), *Language, interaction and social cognition* (pp. 11–36). Thousand Oaks, CA: Sage Publications.
- Rydell, R. J., & Gawronski, B. (2009). I like you, I like you not: Understanding the formation of context-dependent automatic attitudes. *Cognition and Emotion*, 23(6), 1118–1152.
- Sacchi, D. L. M. (2015). *Category learning and group size: A social-cognitive account of the formation of minority stereotypes* [Unpublished doctoral dissertation]. Davis, CA: University of California.
- Sherman, J. W., & Frost, L. A. (2000). On the encoding of stereotype-relevant information under cognitive load. *Personality and Social Psychology Bulletin*, 26(1), 26–34.
- Sherman, J. W., Stroessner, S. J., Conrey, F. R., & Azam, O. A. (2005). Prejudice and stereotype maintenance processes: Attention, attribution, and individuation. *Journal of Personality and Social Psychology*, 89(4), 607–622.
- 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(2), 305–323.
- Sherman, J. W., Allen, T. J., & Sacchi, D. L. M. (2012). Stereotype confirmation and disconfirmation. In B. Gawronski & F. Strack (Eds.), *Cognitive consistency* (pp. 390–423). New York: Guilford.
- Simon, B. (1992). The perception of ingroup and outgroup homogeneity: Reintroducing the intergroup context. *European Review of Social Psychology*, 3(1), 1–30.
- Simon, B., & Mummendey, A. (1990). Perceptions of relative group size and group homogeneity: We are the majority and they are all the same. *European Journal of Social Psychology*, 20(4), 351–356.
- Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, 105(1), 131–142.
- Smith, E. R. (1991). Illusory correlation in a simulated exemplar-based memory. *Journal of Experimental Social Psychology*, 27, 107–123.
- Sokolov, E. N. (1963). Higher nervous functions: The orienting reflex. *Annual Review of Physiology*, 25, 545–580.
- Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social developmental literatures. *Psychological Bulletin*, 111(1), 42–61.

- Stern, L. D., Marrs, S., Millar, M. G., & Cole, E. (1984). Processing time and the recall of inconsistent and consistent behaviors of individuals and groups. *Journal of Personality and Social Psychology*, 47(2), 253–262.
- 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: Erlbaum.
- Susskind, J., Maurer, K., Thakkar, V., Hamilton, D. L., & Sherman, J. W. (1999). Perceiving individuals and groups: Expectancies, dispositional inferences, and causal attributions. *Journal of Personality and Social Psychology*, 76(2), 181–191.
- Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British Journal of Psychology*, 54(2), 101–114.
- Von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. *Advances in Experimental Social Psychology*, 27, 177–254.
- Weber, R., & Crocker, J. (1983). Cognitive processes in the revision of stereotypic beliefs. *Journal of Personality and Social Psychology*, 45(5), 961–977.
- Winman, A., Wennerholm, P., Juslin, P., & Shanks, D. R. (2005). Evidence for rule-based processes in the inverse base-rate effect. *The Quarterly Journal of Experimental Psychology*, 58A(5), 789–815.
- Wittenbrink, B., Judd, C. M., & Park, B. (2001). Spontaneous prejudice in context: Variability in automatically activated attitudes. *Journal of Personality and Social Psychology*, 81(5), 815–827.
- Yee, D. K., & Eccles, J. S. (1988). Parent perceptions and attributions for children's math achievement. *Sex Roles*, 19(5), 317–333.