Journal of Personality and Social Psychology
1997, Vol. 73, No. 3, 459-471

0022-3514/97/$3.00

Stereotypes and Tacit Inference

David Dunning
Cornell University

David A. Sherman
Stanford University

To judge another person's behavior, one often has to come to an understanding of what that behavior was in its detail. Five studies demonstrated that stereotypes influence the tacit inferences people make about the unspecified details and ambiguities of social behavior (e.g., what the behavior specifically was, what stimulus the individual reacted to, what caused the individual to act) and that these inferences occur when people encode the relevant information. One study found that participants who scored low on a measure of modern sexism were just as likely to make tacit inferences based on gender stereotypes as were those who scored high. Discussion centers on the implications of these findings for identification processes in social judgment, as well as whether stereotypes influence tacit inferences at an implicit level.

To what extent are stereotypes inferential prisons? To what degree do people think they "know" a person once they discover his or her ethnic group, gender, social class, or occupation? Are stereotypes maximum security prisons, with people's inferences and impressions of the person never escaping far from the confines of the stereotype? Or are the prisons not so secure, with people escaping the influence of their stereotypes as they learn more about the individual?

Social psychological research gives conflicting evidence on the persistence of stereotypes in the face of individuating information. On the one hand, people have been shown to abandon their stereotypes just as soon as they garner individuating information about a person (Ashmore, 1981; Glick, Zion, & Nelson, 1988; Heilman, 1984). As an example, Locksley and colleagues discovered that stereotypes about gender had no impact on judgments of aggressiveness once participants received concrete information about an individual (Locksley, Borgida, Brekke, & Hepburn, 1980; Locksley, Hepburn, & Ortiz, 1982). On the other hand, researchers have also shown that people hold onto stereotypes as they judge others. For instance, for any given act of aggressiveness, people see a man performing it as more aggressive than a woman (Futoran & Wyer, 1986; Krueger & Rothbart, 1988). People take individuating information into account, but simply add it to a baseline of aggressiveness they associate with men and women.

In this article we portray the cognitive prison of stereotypes to be rather secure, for the impact of the stereotype may not be reduced by specific information about the individual. We argue that stereotypes often lead people to make *tacit inferences* about individuating information. These inferences alter the meaning of the information to affirm the stereotype people possess. Indeed, such inferences can alter the behavior or information about which people believe they have been informed. As an example, consider the sentence *Some felt that the politician's statements were untrue.* Although not specified, we propose that the stereotype of politicians would lead people to make tacit inferences about why people have such feelings in that passage. In particular, it would lead people to believe that the politician was lying, although that was not specified in the passage. Such an inference, however, would not be made if the character belonged to another stereotyped group, such as physicists. In this case the stereotype of physicists might lead people to assume that the physicist was only mistaken in his or her assertions.

We propose that social information often leaves room for these types of tacit inferences. People must often specify the exact nature of the action taken, to identify exactly the behavior under consideration or the exact situation confronting the actor. To the extent that the information is ambiguous about the behavior or situation, stereotypes may guide people in their inferences about that information. Indeed, it is easy to see how such tacit inferences may lead people to confirm their stereotypes and apply them in making social judgments. For example, if the information given above leads people to believe that the politician is lying, but that the physicist is only mistaken, then people may conclude that the politician is more dishonest than the physicist, even though people have received ostensibly the same information about both.

In this regard, our logic is consistent with the notions of Trope (1986; Trope, Cohen, & Alfieri, 1991) concerning the specific processes people follow in making social judgments. According to Trope, before people make judgments about another individual, they must complete a preliminary but important step: They must *identify* the behavior that is to be judged. Many

David Dunning, Department of Psychology, Cornell University; David A. Sherman, Department of Psychology, Stanford University.

David A. Sherman was supported by a National Science Foundation predoctoral fellowship during the conduct of some of this research. We thank Nenshad Bardoliwalla and Jeannine Delwiche for coding the data from Study 3 and Joy Geng for running the participants of Study 5. We also thank Paul Whitmore as well as the participants of the Social/Personality Proseminar at Cornell University for their comments on a previous version of this article. Finally, we extend our gratitude to Karin Mogg for generously supplying us with stimulus materials from Eysenck, Mogg, May, Richards, and Mathew's (1991) study, which served as an inspiration for many aspects of our own materials.

Correspondence concerning this article should be addressed to David Dunning, Department of Psychology, Uris Hall, Cornell University, Ithaca, New York 14853-7601. Electronic mail may be sent via the Internet to dad6@cornell.edu.

stimuli in the social realm are ambiguous, open to many interpretations. As such, people must form an interpretation of the stimulus before making judgments. For example, suppose one turns the corner in the hallway and sees that a person is weeping. Is that behavior an example of sadness, or one of happiness? It can be either, depending on the context surrounding the act. If the person has just failed an exam, the behavior should properly be identified as sadness. However, if the person has just gotten into her first-choice law school, then the behavior should be identified as unbridled joy. Only after the behavior has been identified can people then proceed to make dispositional judgments about the individual.[1]

In this article we propose that stereotypes alter the tacit inferences people make when comprehending descriptions of social behavior. In Trope's (1986) terms, stereotypes influence identification processes in social judgment, leading people to different conclusions about the exact behavior about which they have been informed. We propose that such inferences are made spontaneously, that is, without prompting from any external agent. Moreover, we propose that these inferences can be so strong that people often mistakenly believe that those inferences were not inferences at all, but rather information presented in the original description.

We assessed these proposals by observing people's memory for passages about individuals in differing stereotyped groups. In five studies, we presented people with passages that invited them to make tacit inferences. In subsequent memory tests, we presented them with altered passages that contained the tacit inferences suggested by the relevant stereotype. We predicted that people would be more likely to falsely remember these altered passages as previously presented when the inferences contained in the sentences were consistent with the stereotype contained in them than when they were inconsistent with the stereotype.

Do People Make Tacit Inferences?

There is a long history of work in cognitive psychology, coming mostly from research on text comprehension, demonstrating that people make tacit inferences about social stimuli with which they are confronted. More than 20 years ago, Johnson, Bransford, and Solomon (1973; see also Bransford, Barclay, & Franks, 1972) demonstrated how tacit inferences appeared in memory for text. They presented participants with passages such as *John was trying to fix the bird house. He was pounding the nail when his father came out to watch him and to help him do the work.* Subsequent memory tests revealed that participants had inferred that John was using a hammer, even though no such instrument had been mentioned in the original passage. That is, they misremembered the sentence as *John was using the hammer to fix the bird house when his father came out to watch him and to help him do the work.*

Several years of text comprehension work has confirmed that people readily make tacit inferences about what they read. They have been shown to make inferences regarding the meaning of ambiguous words (Small, Cottrell, & Tanénhaus, 1988; Woll, Weeks, Fraps, Pendergrass, & Vanderplas, 1980; Woll & Yopp, 1978), to specify how different components of the passage refer to one another (O'Brien, Shank, Myers, & Rayner, 1988), to determine the causes and consequences of actions (Singer & Ferreira, 1983; van den Broek, 1990), to specify the emotions felt by the characters in the passage (Gernsbacher, Goldsmith, & Robertson, 1992), and to arrive at the overall theme of the passage (Till, Mross, & Kintsch, 1988). Related work in social psychology has shown that people make inferences regarding the personality traits of the actors in the passage (Winter & Uleman, 1984).

Some researchers have articulated more formal and computational theories about the cognitive processes involved in text comprehension, to show how these processes necessarily involve tacit inference and to demonstrate how such inferences are incorporated into memory of the passage. In one influential model, Kintsch (1988) described text comprehension as a two-step process. First, people decompose the propositions contained in any utterance. For example, in the sentence *The bosses discussed the inadequate work of their summer sales staff, and decided that Scott was the most responsible,* people must break the sentence down into its information units (e.g., work was discussed, the work was inadequate, it was discussed by the bosses, the work was done by the summer sales staff, etc.). Second, people must add propositions to their representation of the sentence to comprehend it. For example, in the sentence above, Scott is portrayed as responsible, but to what does *responsible* refer? Does it refer to a desirable personality trait, or does it refer to the fact that Scott is the one most responsible for the inadequate work? Kintsch suggested that people add such propositions to form a coherent comprehension of the presented message, adding propositions to the memorial representation of the passage. Thus, when a new sentence containing those inferences is shown to an individual, the new sentences will provide a good match to that individual's memorial representation and will be mistakenly remembered as previously encountered (Kintsch, Welsch, Schmalhofer, & Zimny, 1990).

To be sure, there is much in this work that remains controversial. Researchers differ on the types of inferences people are likely to make when reading a narrative, and they also disagree on whether these inferences are made strategically or automatically (see recent discussions, for example, by Graesser, Singer, & Trabasso, 1994, and McKoon & Ratcliff, 1992). However, there is a consensus among researchers in the field that people do make some types of tacit inferences in narrative passages that describe people's actions.

Would Stereotypes Influence the Tacit Inferences People Make?

Although the work discussed above suggests that people make many tacit inferences while reading narratives about others, it provides no evidence for the specific hypothesis that stereotypes would prompt people to make different tacit inferences about the exact same information about individuals in different stereotypical groups.

---

[1] A careful reader may have noticed that we specified the gender of the person crying in the hall rather late in the example. If this was noticed, we congratulate the reader for reading carefully. If it was not, it may be evidence that the reader made a tacit inference about the person in the example at the time he or she read the passage.

More direct evidence of this hypothesis does exist, although it is scant. Kunda and Sherman-Williams (1993) directly asked whether stereotypes would alter the way people construed the features and details of an ambiguous behavior. They presented participants with sentences such as *X hit someone who annoyed him* and asked them explicitly to report the elaborative inferences they made on reading it, that is, what specific, detailed scenarios came to mind as they considered the sentence. Participants reported different acts based on differing stereotypes. When told that a construction worker had hit another person, they described him as punching a coworker. When told that a housewife had hit another person, participants described the action in more benign terms, such as spanking a naughty child. It is important to note that judgments of the target's aggressiveness were dependent on these interpretations. When these inferences about the details of the event were controlled for, either statistically or experimentally, Kunda and Sherman-Williams found no difference in how participants rated the aggressiveness of target individuals.

In a related vein, Slusher and Anderson (1987) discovered that having people imagine the situation surrounding a person's behavior encouraged stereotypical inferences. They presented participants with sentences involving stereotyped occupations (e.g., *Frank, a lawyer, is trying to reach the check-out counter in a crowded department store*) and asked them to spend some time imagining the scene the sentences depicted. In one study, participants tended to imagine details that tended to be stereotypical (e.g., they imagined that the lawyer had acted in an aggressive manner). In another study, Slusher and Anderson found that asking participants to imagine the scene surrounding the sentence tended to prompt them to overestimate the number of stereotype-relevant trait terms (e.g., *aggressive*) that had been explicitly presented in the those sentences.

However, although these two sets of findings support our assertion that stereotypes influence how people interpret the specifics of behavior, they are not conclusive. It is our view that stereotypes influence how people construe the details at the time they read the passage, without any prompting from external forces. Kunda and Sherman-Williams (1993) explicitly asked participants to provide their interpretations of the passages given them. As such, it is unclear whether the inferences that participants made occurred while they were reading the passage. Indeed, it is unclear whether participants, without prompting, would have made any inferences at all. Instead, they may have taken the passage at face value and made no inferences about the specific behaviors they read about or the circumstances surrounding those behaviors. A similar critique can be applied to Slusher and Anderson (1987). They expressly asked participants to imagine particular situations, and so it is unclear whether participants would have made any stereotypical tacit inferences without the experimenter's prompting.

Goals of the Present Research

In five studies, we tested whether stereotypes influence the tacit inferences people make of individuating information about other people. We tested whether these inferences were made without prompting and whether they occur at the time of comprehension. We also tested whether their strength is reflected in people's memorial representation of the passage—whether these inferences can be held so confidently that they are confused and mistaken for previously presented information.

In all studies, participants were presented with sentences that invited tacit inferences. Each sentence associated an ambiguous action or piece of information with one of two stereotypes, such as *Amy found it hard to disguise her feelings toward the Hollywood actor* or *Amy found it hard to disguise her feelings toward the criminal.* People were soon after presented with sentences containing inferences that were either consistent or inconsistent with the one stereotype that had been presented to them (e.g., *Amy found it hard to disguise her attraction toward the Hollywood actor* versus *Amy found it hard to disguise her repulsion toward the Hollywood actor*). We predicted that people would be more likely to falsely recognize a new sentence when it was consistent with the stereotype contained in it than when it was inconsistent with the stereotype.

In the first study we presented participants with a recognition memory test containing stereotype-consistent and -inconsistent sentences. We predicted that participants would be more likely to false-alarm to the former than to the latter. In the second study we explored whether this effect was due to tacit inferences or to a response bias for plausibility. In the third study we examined whether the effect was influenced by the specific task we asked participants to do while they read the sentences. In the fourth study we examined *when* participants make their stereotype-based inferences: at the time of comprehension (i.e., while reading the passage) or at the time of retrieval (i.e., while completing the recognition memory test). Finally in a fifth study we examined whether this phenomenon would generalize to gender stereotypes, as well as whether even low-sexist individuals would make tacit inferences based on gender.

Study 1: Memory for Stereotype-Consistent and -Inconsistent Inferences

In Study 1 we investigated whether stereotypes influence the tacit inferences made from individuating information and alter memory of that information. Participants were presented with passages about individuals that invited tacit inferences. Different participants had different stereotypes associated with those passages. For example, all participants were given a sentence saying that *Amy found it hard to disguise her feelings toward X.* Half the participants were told that "X" was a Hollywood actor; half were told that "X" was a criminal. Afterward, participants were given a memory test that contained different sentences containing an inference that could plausibly be made from one of the stereotypes (e.g., *Amy found it hard to disguise her attraction toward the Hollywood actor*). For half the sentences like this in the memory test, the sentence included an inference consistent with the stereotype presented to participants. For the other half, the sentence included an inconsistent inference, that is, an inference consistent with the unpresented stereotype (e.g., *Amy found it hard to disguise her repulsion toward the Hollywood actor*). Our prediction was straightforward: Participants would falsely recognize more sentences containing stereotype-consistent inferences than sentences containing stereotype-inconsistent ones.

Method

Participants. Participants were 10 Cornell University undergraduates enrolled in a senior/graduate student level course on research methods. They participated as a class exercise on data analysis.

Procedure. The experiment was run in one session during class time. Participants were told that they would be reading several sentences about different individuals. After they indicated that they understood the task, they signed consent forms and sat in front of Macintosh LC III computers, which presented them with a few preliminary instructions and told participants that they should simply "try to form immediate impressions of the person or people" in the sentences they saw. The computer then presented participants with 50 sentences that either described the behavior of an individual or provided some descriptive information about him or her. The computer presented each sentence for 6 s, with no delay between sentence presentation.

After the computer presentation was complete, participants were introduced to a distracter task. They were asked to list as many prime numbers as they could, starting with the number 2, for 1 min. They were then given a recognition test for the presented sentences. The experimenter gave participants a list of 48 sentences and told them that some of the sentences had been presented previously and others had not. Participants were told to indicate whether they had seen the *exact same* sentence on the computer screen. Participants responded on a 4-point recognition confidence scale ranging from 1 (*definitely there*) to 4 (*definitely not there*).

Participants then filled out a few follow-up questionnaires that included a probe for suspicion. The goals and design of the study were described. Participants were later given the data gathered at the session to complete as a data analysis exercise for the class.

Materials. The 50 sentences shown on the computer consisted of 34 filler and 16 critical items. There were two versions of each of the 16 critical sentences, differing in the occupation of the actor in the sentence. For example, one sentence read *The X was unhappy about the amount of liquor being served at the party.* In one version of the sentence, the main character was described as a nun. In the other, the main character was described as a rock musician. For each version of the sentence, a stereotype-consistent interpretation was generated. For example, for the nun version of the sentence above, the stereotype-consistent interpretation presented on recognition test was *The nun was unhappy about the large amount of liquor being served at the party.* For the rock musician version, the stereotype-consistent interpretation was *The rock musician was unhappy about the small amount of liquor being served at the party.* Table 1 presents the 16 critical sentences with the stereotype-consistent interpretations included in the recognition test. For each sentence, inconsistent interpretations consisted of pairing each interpretation with the alternative stereotype associated with each sentence.

In a preliminary study, we asked 25 people to read the four versions of the stimulus sentences (two were stereotype consistent, and two were stereotype inconsistent) and then to judge the plausibility of each version on a 9-point scale. Stereotype-consistent versions for each critical sentence were rated as more plausible than their stereotype-inconsistent counterparts, mean $t = 5.94$ (ts ranged from 2.30 to 9.91), all ps $< .05$.

For the experiment proper we randomly selected one occupation for each sentence and placed it into Occupation Set A. The remaining occupation for each sentence was placed into Occupation Set B. Table 1 indicates which occupation for each sentence was in Occupation Sets A and B. In the experiment, half of the participants saw sentences containing Occupation Set A, and the other half saw Set B. These critical sentences were randomly interspersed among the filler items, with one proviso: To avoid primacy and recency effects in memory, the first five and the last five sentences in the acquisition set were filler items.

Each critical sentence in the acquisition set had a corresponding sentence in the recognition memory task. Prior to the experiment, the 16 critical sentences in the recognition memory task were divided equally into two groups: Interpretation Sets 1 and 2. Table 1 describes which sentences were contained in each set. For some participants, the 8 sentences comprising Interpretation Set 1 were interpretations consistent with the stereotypes that had been shown to participants. For those same participants, the sentences from Interpretation Set 2 were stereotype-inconsistent items. The remaining participants read recognition sentences that contained stereotype-consistent interpretations from Set 2 and stereotype-inconsistent ones from Set 1.

As such, two factors were counterbalanced across participants. The first was the occupation set shown to participants as they read the acquisition sentences. The second was the specific interpretations shown during the recognition test that were stereotype consistent and inconsistent. There were 2 to 3 participants in each individual cell of the 2 × 2 factorial that arose from this counterbalancing.

Of the 32 remaining items on the recognition test, 14 were filler items that had been shown in the acquisition set, and 18 were either altered or completely new sentences.

Results and Discussion

The two counterbalancing factors (occupation and interpretation set) had no impact on any results reported below. They are discussed no further.

To assess memory for the presented sentences, we examined the portion of sentences in the recognition test that participants said had been presented (that is, they circled 1 or 2 on the recognition response scale). Analyses indicated that participants had good memory for presented sentences. Participants correctly recognized an average of 78% of the previously presented filler items and mistakenly recognized only 1% of the filler items that had not been presented, $t(9) = 11.98$, $p < .0001$. However, participants made a greater number of mistakes when dealing with critical items. As predicted, participants mistakenly recognized an average of 35% of stereotype-consistent interpretations, but only 15% of stereotype-inconsistent ones, $t(9) = 4.71$, $p < .002$.[2]

In sum, Study 1 provided initial evidence that stereotypes influence interpretations of another person's behavior. Participants in Study 1 mistakenly recognized sentences that contained tacit inferences when those inferences were consistent with stereotypes of the sentence's protagonist. Indeed, they falsely recognized over a third of those sentences, even though none of these sentences had actually been previously presented. In contrast, when the sentence contained inferences that were inconsistent with the stereotype in question, participants were significantly less likely to claim to have seen the sentence before.

Studies 2 and 3: Alternative Explanations

The results of Study 1 leave open some alternative explanations that were assessed in Studies 2 and 3. For example, perhaps participants did not make the inferences we claimed they were making—perhaps their memory reports were merely the result of response biases. Participants may have had some slim shards of memory about the plethora of sentences that had been presented. As such, when faced with the recognition memory test, they judged whether they had seen a sentence before on the

[2] Analyses in which recognition confidence was the dependent measure produced virtually identical results in all studies reported herein.

Table 1
Critical Acquisition and Recognition Sentences Used in Study 1

Acquisition sentence	Stereotype-consistent interpretation used for occupation set A	Stereotype-consistent interpretation used for occupation set B
1. Some felt that the (physicist's/politician's) statements were untrue.	Some felt that the physicist's statements were mistaken.	Some felt that the politician's statements were lies.
2. The management discussed the poor performance of their summer sales staff and concluded that the (straight-A student/company president's son) was the most responsible.	The management discussed the poor performance of their summer sales staff and concluded that the straight-A student was the most dependable.	The management discussed the poor performance of their summer sales staff and concluded that the company president's son was the most responsible for it.
3. After weighing all the circumstances, the (head of the computer software company/drug dealer) decided that he would have to terminate a few of his employees.	After weighing all the circumstances, the head of the computer software company decided that he would have to fire a few of his employees.	After weighing all the circumstances, the drug dealer decided he would have to kill a few of his employees.
4. The (accountant's/Marine drill sergeant's) personality was a little hard to take.	The accountant's obsessive personality was a little hard to take.	The Marine drill sergeant's overbearing personality was a little hard to take.
5. The (nurse/bar bouncer) hurriedly rushed through the people to the check-out counter in the crowded department store.	The nurse gingerly dodged her way through the people to the check-out counter in the crowded department store.	The bar bouncer hurriedly pushed his way through the people to the check-out counter in the crowded department store.
6. The punch that the (truck driver/bartender) gave the guy just knocked him out.	The punch that the truck driver threw at the guy just knocked him out.	The drink that the bartender gave the guy just knocked him out.
7. Amy found it hard to disguise her feelings toward the (criminal/Hollywood actor).	Amy found it hard to disguise her repulsion toward the criminal.	Amy found it hard to disguise her attraction toward the Hollywood actor.
8. The (fashion model/triathlete) had to be concerned every day about her physical condition.	The fashion model had to be concerned every day about her physical appearance.	The triathlete had to be concerned every day about her physical fitness.
9. After a few drinks, the two (marriage counselors/lumberjacks) had a fight in the restaurant.	After a few drinks, the two marriage counselors had a quarrel in the restaurant.	After a few drinks, the two lumberjacks had a fist fight in the restaurant.
10. The (nun/rock musician) was unhappy about the amount of liquor being served at the party.	The nun was unhappy about the large amount of liquor being served at the party.	The rock musician was unhappy about the small amount of liquor being served at the party.
11. The two men questioned the (priest/gambling casino operator) about his convictions.	The two men questioned the priest about his beliefs.	The two men questioned the gambling casino operator about his prison record.
12. The student couldn't make sense of the (psychiatric patient's/neurobiology professor's) language.	The student couldn't make sense of the psychiatric patient's incoherent language.	The student couldn't make sense of the neurobiology professor's jargon.
13. The (librarian/investment banker) purchased a brand new car.	The librarian purchased a brand new compact car.	The investment banker purchased a brand new sports car.
14. The (used car salesman/computer hacker) was not known for being the most socially skilled person.	The used car salesman was not known for being the most courteous, socially skilled person.	The computer hacker was not known for being the most outgoing, socially skilled person.
15. The (gangster/police officer) felt he had his reputation to uphold.	The gangster felt he had his tough reputation to uphold.	The police officer felt he had his honest reputation to uphold.
16. Everyone giggled when the (TV talk show host/foreign exchange student) spoke.	Everyone giggled in delight when the TV talk show host spoke.	Everyone giggled in embarrassment when the foreign exchange student spoke.

Note. In the far left column, the first occupation listed in parentheses is the occupation used in Set A, and the second is the one used in Set B. Sentences 1–6 and 13–14 comprise Interpretation Set 1, and Sentences 7–12 and 15–16 comprise Interpretation Set 2. The interpretations listed in Columns 2 and 3 include the occupation that render them stereotype consistent. These sentences were made stereotype inconsistent by replacing the stereotype included with the alternative (e.g., exchanging *politician* and *physicist* in Sentence 1).

basis of its plausibility. When they saw a stereotype-consistent sentence that seemed plausible and matched some of their shards of memory, they endorsed it as previously seen. When the sentence was implausible, such as was the case with stereotype-inconsistent items, they may have been biased to perceive the item as not previously seen.

Study 2

Method

We designed Study 2 to examine this response bias interpretation. The procedure was identical to that used in Study 1, except for two changes. The first was that we included only 12 critical sentences as opposed to the 16 used in Study 1 (essentially, we dropped Sentences 13–16 from Table 1). We replaced the 4 omitted sentences with 4 additional filler items.

The second change was more substantive. Participants in one condition ($n = 18$) were run in a close replication of Study 1, viewing an acquisition set of 50 sentences. The remaining participants ($n = 16$) also viewed the same set of 50 sentences at acquisition, except that the 12 critical items were replaced by 12 control sentences. Each control sentence closely resembled the critical sentence on which it was modeled, except that the key ambiguous phrase in the critical sentences was replaced by another phrase. For example, the sentence *The nun was unhappy about the amount of liquor being served at the party* was altered to read *The nun was unhappy about the decorations at the party.* Table 2 contains a few examples of control sentences, along with the critical counterparts on which they were modeled. We then gave all participants a recognition memory test containing the stereotype-consistent and -inconsistent interpretations used in Study 1. These interpretations could plausibly follow from our original sentences, but not from the altered ones.

We reasoned that if mistaken recognition of consistent interpretations is a product of response biases, then participants should be more likely to false-alarm to all consistent interpretations regardless of whether they saw our original sentences or the altered, control versions. However, to the extent that false alarms on these sentences were a product of tacit

Table 2
Examples of Critical and Control Acquisition Sentences Used in Study 2

Critical sentence	Control sentence
After weighing all the circumstances, the (head of the computer software company/drug dealer) decided that he would have to terminate a few of his employees.	After weighing all the circumstances, the (head of computer software company/drug dealer) decided that he would have to talk to a few of his employees.
The management discussed the poor performance of their summer sales staff and concluded that the (straight-A student/company president's son) was the most responsible.	The management discussed the poor performance of their summer sales staff and concluded that the (straight-A student/company president's son) was the most handsome.
Amy found it hard to disguise her feelings toward the (criminal/Hollywood actor).	Amy found it hard to disguise her limp in front of the (criminal/Hollywood actor).
The (fashion model/triathlete) had to be concerned every day about her physical condition.	The (fashion model/triathlete) had to be concerned every day about her finances.
The (nun/rock musician) was unhappy about the amount of liquor being served at the party.	The (nun/rock musician) was unhappy about the decorations being used at the party.
The two men questioned the (priest/gambling casino operator) about his convictions.	The two men questioned the (priest/gambling casino operator) about his family.

inference, then participants should be more likely to false-alarm to consistent interpretations only when they were preceded by our original, interpretable stimuli. Thus, for each participant we calculated the percentage of stereotype-consistent interpretations they incorrectly recognized and did the same for the percentage of stereotype-inconsistent sentences. We then submitted these percentages to a 2 (original vs. control sentences at acquisition) × 2 (percentage of stereotype-consistent vs.-inconsistent interpretations falsely recognized) mixed model analysis of variance (ANOVA), with the last factor serving as a within-subject variable.

Results

Three effects of interest emerged. First, participants who viewed critical items overall made a greater percentage of false alarms (M = 37%) than did participants who viewed control items (M = 6%), $F(1, 32) = 66.71$, $p < .0001$. Second, participants made more false alarms when presented with stereotype-consistent interpretations (M = 27%) than they did when presented with stereotype-inconsistent ones (M = 18%), $F(1, 34) = 8.22$, $p < .01$.

However, these main effects were qualified by a predicted two-way interaction, $F(1, 34) = 7.31$, $p < .02$. When participants viewed ambiguous critical sentences at acquisition, they incorrectly recognized a greater percentage of stereotype-consistent interpretations (M = 46%) at recognition than they did stereotype-inconsistent ones (M = 28%), $t(17) = 3.25$, $p < .005$. No such effect was observed for participants who viewed control sentences (Ms = 6% for both stereotype-consistent and -inconsistent items, $t < 1$, ns). In sum, Study 2 served to rule out an alternative explanation based on response biases due to plausibility.

Study 3

Study 3 was designed to test another artifactual alternative explanation for the above findings. In Studies 1 and 2 we asked participants to form "immediate impressions" of the person or people described in the various sentences. Perhaps this explicit instruction to make social judgments prompted people to make tacit inferences, inferences they would not have made under other circumstances. More important, perhaps this explicit instruction prompted people to make tacit inferences that relied heavily on their stereotypes—because how else could people form immediate impressions of the individuals about whom they read? After all, it has been shown that goals at the acquisition of social information have significant effects on the types of inferences that people make (Bassili & Smith, 1986; Uleman & Moskowitz, 1994). Thus, in Study 3 we examined whether tacit inferences were an artifact of our instructions to form impressions of the individuals about whom participants read.

Method

The materials and procedures used were identical to those used in Study 2, with two variations. First, all participants saw ambiguous sentences; no control sentences were used. Second, we varied the task that participants were asked to complete as they saw the acquisition set of sentences. Participants (n = 12) in one condition were presented with ambiguous sentences with instructions to form impressions of the people portrayed in them, as in Studies 1 and 2. The remaining participants (n = 12) were given instructions that did not call for such tacit inferences; they were simply told to assess how "readable" the sentences were. If the tacit inferences observed in Studies 1 and 2 are an artifact of the specific instructions we gave to participants, then participants' performance on the recognition test should interact with the type of instruction (impression or readability) we gave them in Study 3. If these tacit inferences are not artifactually prompted by the particular task, then we should observe the same effects on recognition memory regardless of which task we give to participants.[3]

[3] A careful reader may be surprised by our choice of an alternative task. Usually, researchers contrast the effects of social judgment tasks with the effects of asking participants to remember what they had seen (e.g., Bassili & Smith, 1986). However, for Study 3 we decided not to include a memory task condition, because asking participants to remember what they saw may prompt them to try to elaborate on, or to visualize, the sentences they saw. Such strategies are effective mnemonic devices not unknown among undergraduates (Anderson & Reder, 1979). Asking participants to judge the readability of the sentences they saw, however, was not so prone to causing undergraduates to elaborate on what they had seen.

Results

The only effect to emerge from a 2 (participant given impression or readability instructions at acquisition) × 2 (percentage of stereotype-consistent vs. -inconsistent interpretations falsely recognized) mixed model ANOVA was that false alarms were greater for stereotype-consistent interpretations ($M = 41\%$) than for stereotype-inconsistent ones ($M = 21\%$), $F(1, 22) = 18.54$, $p < .005$. This main effect failed to interact with the task participants pursued at acquisition ($F < 1$). The effect was significant for participants in the impression task condition (Ms = 43% and 25% for stereotype-consistent and -inconsistent interpretations, respectively), $t(11) = 2.86$, $p < .02$, as well as for those in the readability task condition (Ms = 39% and 17% for stereotype-consistent and -inconsistent interpretations, respectively), $t(11) = 3.22$, $p < .01$.

In sum, Study 3 ruled out an alternative explanation that the propensity for participants to make stereotype-consistent inferences was an artifact of our specific instructions. Even when given instructions that did not call for any sort of inferential work, participants were still more likely to give false-alarm responses to stereotype-consistent interpretations of the acquisition sentences than they were to stereotype-inconsistent ones.

Study 4: Do Tacit Inferences Occur at Encoding or Retrieval?

Although Studies 1–3 provide evidence that differing stereotypes can impel people to make divergent tacit inferences about the same piece of information, they still leave one important question unanswered: When do people make these inferences—at the time they read the relevant passage or at the time they confront the recognition test? We contend that people make these inferences at encoding (see von Hippel, Sekaquaptewa, & Vargas, 1995, for arguments that virtually all stereotype processes occur at encoding), when people initially read the passages. However, people may instead remember stimulus sentences veridically, and reinterpret those sentences only at retrieval, when they see similar and plausible sentences on the recognition test.

We designed Study 4 to test whether people make tacit inferences at the time of encoding or at retrieval. We did so by borrowing a procedure used by Winter and Uleman (1984; see also Winter, Uleman, & Cunniff, 1985) to test whether people make inferences about the personality traits of others as they read about their behavior. In these studies, participants were presented with a series of behaviors and then were asked to recall those behaviors using a cued-recall task. The cues used to promote recall were the personality traits people had presumably inferred as they read the original sentence. If people had spontaneously inferred the traits as they read the original sentences, then the traits should have facilitated recall for the original sentences. If the trait had not been inferred, then no facilitation of recall should have occurred. Consistent with their arguments, Winter and Uleman found that people had indeed inferred personality traits concerning the people about whom they had read. When cued with relevant personality traits, they recalled a greater number of sentences than they did when given no cues at all.

In Study 4 we adopted this technique to see if participants made their inferences at the time they read stimulus passages. We presented them with the ambiguous sentences used in the previous studies (e.g., *Amy found it difficult to disguise her feelings toward the Hollywood actor*) and then asked them to recall those sentences a few minutes later. For some of the sentences, participants were given stereotype-consistent cues. That is, they were given short phrases describing the inferences we presumed that participants were making (e.g., *romantically attracted to him*). For other sentences, they were given stereotype-inconsistent cues (e.g., *disgusted and repulsed by him*). For other sentences, they were given no cues at all. If participants made tacit inferences at encoding, then they should recall a greater number of sentences when given stereotype-consistent cues than when given stereotype-inconsistent cues or no cues.

Method

Participants. Participants were 14 Cornell University undergraduates enrolled in an introductory design and environmental analysis class. Participants were given extra credit toward their course grades for taking part.

Procedure. Participants were tested in groups of up to 8. The procedure closely mirrored that of the previous three studies, with all participants given instructions to "form immediate impressions" of the people about whom they read, except that participants viewed only 16 sentences during acquisition. Those sentences consisted of the 12 critical items used in the first two studies, plus 4 filler items. These filler items were included to buffer our results against primacy and recency effects in memory. That is, 2 of the filler items appeared first, and 2 appeared last, in the series of sentences we presented to participants. The 12 critical sentences were shown in random order in the middle of the series.

After viewing the sentences, participants completed a 5-min distracter task in which they wrote down the names of their past jobs and good friends. They then were given a cued-recall task. They were asked to recall as many as they could of the 16 sentences they had seen earlier, but to be as accurate as possible. For 8 of the sentences, they were given cues to aid them in their recall. Participants were asked to write down the sentence associated with each cue. Four of the cues were stereotype consistent, in that they contained the inference we presumed participants would make when exposed to the stereotype mentioned. For example, for the sentence *The nun was unhappy about the amount of liquor being served at the party,* the relevant cue was *didn't like all the alcohol.* The remaining four cues were stereotype inconsistent, in that they presented an inference that contradicted the stereotype in the sentence (and were consistent with the alternative stereotype also connected to the sentence). For example, for the nun sentence above, the stereotype-inconsistent cue was *wanted more alcohol.* A list of all the stereotype-consistent and -inconsistent cues is provided in Table 3. For the remaining 8 sentences, participants were given eight empty slots to recall as many of the noncued sentences as they could remember.

As in the previous three studies, participants viewed sentences containing occupations listed in Set A or B. The specific sentences cued by stereotype-consistent or -inconsistent phrases, or by no cue, were counterbalanced across participants. We randomly grouped the 12 critical sentences into three groups of 4. By means of Latin square, one of these groups of sentences was cued by stereotype-consistent phrases, another by stereotype-inconsistent phrases, and the third by no phrase. In this way, each sentence fell into one of the cue conditions approximately one third of the time.

After completing the recall test, participants filled out a few follow-up questionnaires that included a probe for suspicion. They were then debriefed.

Table 3
Acquisition Sentences and Cues Used in Study 3

Acquisition sentence	Stereotype-consistent cue for occupation set	
	A	B
1. Some felt that the (physicist's/politician's) statements were untrue.	Thought comments were mistakenly in error	Thought comments were lies
2. The management discussed the poor performance of their summer sales staff and concluded that the (straight-A student/company president's son) was the most responsible.	The most dependable worker	Caused a business slump
3. After weighing all the circumstances, the (head of the computer software company/drug dealer) decided that he would have to terminate a few of his employees.	Fired a few people	Killed a few people
4. The (accountant's/Marine drill sergeant's) personality was a little hard to take.	Was too obsessive and compulsive	Was too mean and strict
5. The (nurse/bar bouncer) hurriedly rushed through the people to the check-out counter in the crowded department store.	Gingerly weaved through the people	Pushed way through masses
6. The punch that the (truck driver/bartender) gave the guy just knocked him out.	The jab that was thrown floored him	The drink that was served floored him
7. Amy found it hard to disguise her feelings toward the (criminal/Hollywood actor).	Was disgusted and repulsed by him	Was romantically attracted to him
8. The (fashion model/triathlete) had to be concerned every day about her physical condition.	Was worried about looks and appearance	Was worried about fitness and stamina
9. After a few drinks, the two (marriage counselors/ lumberjacks) had a fight in the restaurant.	Had a verbal spat	Had a fist fight
10. The (nun/rock musician) was unhappy about the amount of liquor being served at the party.	Didn't like all the alcohol	Wanted more alcohol
11. The two men questioned the (priest/gambling casino operator) about his convictions.	Was queried about beliefs	Was queried about criminal record
12. The student couldn't make sense of the (psychiatric patient's/neurobiology professor's) language.	Incoherent babble was unintelligible	Technical jargon was unintelligible

Note. In the far left column, the first occupation listed in parentheses is the occupation used in Set A, and the second is the one used in Set B. These cues listed in Columns 2 and 3 were made stereotype inconsistent by pairing them with the alternative stereotype (e.g., for Sentence 1, pairing *thought comments were mistakenly in error* with *politician*).

To determine which sentences participants had accurately recalled, we gave participant protocols to two coders who were blind to the hypothesis of the study. These coders classified recall of a particular sentence as correct if the participant accurately conveyed the "gist" of the sentence. Interrater reliability was 91%. Disagreements were resolved by means of discussion between the coders.

Results and Discussion

Both counterbalancing factors (occupation and cue set) failed to influence any results reported below. They are discussed no further.

An analysis of participants' recall performance suggested that they made their tacit inferences at the time of encoding. For each participant, we calculated the percentage of sentences accurately recalled when given stereotype-consistent cues, stereotype-inconsistent cues, or no cue. A one-way within-subject ANOVA revealed that there were significant differences depending on the cues that participants received, $F(2, 26) = 18.63$, $p < .0001$. As predicted, participants recalled a higher percentage of sentences cued by stereotype-consistent phrases ($M = 64\%$) than they did sentences cued by inconsistent phrases ($M = 28\%$), $t(13) = 5.04$, $p < .0002$, or by no phrase ($M = 21\%$), $t(13) = 4.64$, $p < .0005$. The difference in recall for stereotype-inconsistent cues and no cues was not significant, $t(13) = 1.44$.

In sum, Study 4 provided evidence that stereotypical tacit inferences are made at encoding and not at retrieval. Stereotype-consistent cues at the time of recall facilitated memory for the original sentences. Such a pattern would occur only if participants made these inferences at encoding, before they confronted the memory test, so that the presence of the cue reminded participants of the sentence associated with it. The data also fail to support an alternative account of how participants approached the memory task. One could argue that participants first recalled sentences and then hunted for the cues associated with those sentences. Such a process could have produced the difference we observed between stereotype-consistent and -inconsistent cues. However, if people recalled sentences first, prior to consulting the cues we provided them, then participants would have recalled just as many no-cue sentences as stereotype-consistent ones.

Study 5: Tacit Inferences Based on Gender

We designed Study 5 with two goals in mind. The first was to generalize our findings to groups that have been of traditional interest to social psychologists. Thus, we created a number of ambiguous sentences that could potentially be interpreted differently depending on whether the protagonist was a man or a

woman. For example, if a person saw the sentence *Elizabeth was not very surprised upon seeing her quantitative SAT score,* would that person be more likely to infer that Elizabeth's score was low than if the protagonist had been named "Bob"?

The second goal was to assess whether participants' attitudes toward gender and gender roles moderated any of our effects. Thus, we selected participants who scored high or low on a scale of modern sexism (Swim, Aikin, Hall, & Hunter, 1995). We wanted to see if participants who scored low would make gender-based tacit inferences at the same rate, or at a lower rate, than their high-scoring counterparts. To the extent that they did, we would have evidence that the prisons that stereotypes create are secure ones, for people who score low on a sexism scale presumably constitute a group that is most likely to be motivated to escape the prison and *not* to make inferences about other people based on their gender.

Method

Participants and selection. Participants were 40 Cornell University undergraduates enrolled in intermediate-level psychology courses. They received extra credit toward their course grade for participating.

Participants were selected according to their scores on the Modern Sexism Scale (Swim et al., 1995). The scale had been administered to several hundred students in intermediate-level psychology courses at the beginning of the semester. Of the participants taking part in Study 5, 18 (16 women) had scored in the bottom 40% of the distribution of the pretest and were termed the *low-sexism group* (M = 15.2, SD = 3.7). The 22 participants (18 women) who were included in the *high-sexism group* had scored in the top 40% (M = 26.5, SD = 3.0).[4]

Procedure. The procedure was identical to the one used in Study 1. Presentation of the 12 critical and 38 filler sentences was done on PowerComputing Power Center 132 personal computers.

Materials. The 50 sentences shown on the computer consisted of 38 filler and 12 critical items. There were two versions of each of the 12 critical items, differing in the gender of the actor in the sentence. For each version of the sentence, a stereotype-consistent interpretation was written. Table 4 presents the 12 critical sentences with the stereotype-consistent interpretations included in the recognition test. For each sentence, inconsistent interpretations consisted of pairing each interpretation with the other gender associated with each sentence.

In a preliminary study we asked 20 people to read the four versions of the stimulus sentences (two were stereotype-consistent, and two were stereotype-inconsistent) and then to judge the plausibility of each version on a 9-point scale. For each critical sentence, the stereotype-consistent versions were rated as more plausible than their stereotype-inconsistent counterparts, mean t = 3.28 (ts ranged from 1.80 to 6.70), all ps < .05, one-tailed.

For the experiment proper we randomly selected one gender for each sentence and placed it into Gender Set A. The remaining gender for each sentence was placed into Gender Set B. Table 4 indicates which gender for each sentence was in Gender Sets A and B. In the experiment, half of the participants saw sentences containing Gender Set A, and the other half saw sentences containing Gender Set B.

Each critical sentence in the acquisition set had a corresponding sentence in the recognition memory task. Prior to the experiment, we divided the 12 critical sentences in the recognition memory task equally into two groups: Interpretation Sets 1 and 2 (see Table 4). For some participants, the 6 sentences comprising Interpretation Set 1 were interpretations consistent with the stereotypes shown to participants. For those same participants, the sentences from Interpretation Set 2 were stereotype-inconsistent items. The remaining participants read recognition sentences that contained stereotype-consistent interpretations from Set 2 and stereotype-inconsistent ones from Set 1. Thus, gender set and interpretation set were counterbalanced across participants.

Results and Discussion

The two counterbalancing factors (gender and interpretation set) had no impact on any results reported below. They are discussed no further. Two participants were dropped from the analyses presented below. One (from the high-sexism group) showed no ability to discriminate between old and new filler items on the recognition tests. Indeed, the number of false positives this participant exhibited for new filler items was 13 SD higher than the grand mean. The second participant (from the low-sexism group) exceeded the grand mean of false alarms to stereotype-inconsistent items by 3.4 SD (1.2 SD from the nearest neighbor).

To assess memory for the presented sentences, we again examined the proportion of sentences in the recognition test that participants labeled as "old" (they circled 1 or 2 on the recognition response scale). Analyses indicated that participants had good memory for presented sentences. Participants correctly recognized an average of 77% of the previously presented filler items and mistakenly recognized only 9% of the filler items that had not been presented, $t(37) = 28.5$, $p < .0001$.

Participants again falsely recognized a greater number of stereotype-consistent interpretations (29%) than they did stereotype-inconsistent ones (18%). A 2 (participant had high or low modern sexism score) × 2 (percentage of stereotype-consistent vs.-inconsistent interpretations falsely recognized) mixed model ANOVA indicated that this tendency was significant, $F(1, 36) = 8.48$, $p < .01$, and did not interact with levels of modern sexism, interaction $F(1, 36) = 0.01$, *ns*. Simple effects tests revealed that participants who scored low on modern sexism were almost as likely to display evidence of stereotype-consistent memory errors (Ms = 27% and 17% for stereotype-consistent and -inconsistent interpretations, respectively, $t(36) = 1.88$, $p < .08$, two-tailed) as were their peers who scored high on modern sexism (Ms = 30% and 20% for stereotype-consistent and -inconsistent interpretations, respectively, $t(36) = 2.22$, $p < .05$).

In sum, Study 5 provided evidence of the generality of stereotype-driven tacit inferences. The tacit inferences that participants made about a brief description of a person's behavior were different depending on whether that person was male or female. This tendency was not qualified by level of sexism. Even participants who scored low on modern sexism, and who presumably were the most committed to gender egalitarianism, still made different inferences of protagonists based on their gender. Indeed, the rate at which they made such inferences was statistically equivalent to that of participants who expressed a greater degree of sexist thought on the modern sexism scale.

General Discussion

How do people deal with specific, concrete information about others in stereotyped groups? Do they abandon their stereotypes

[4] In contrast to Swim et al. (1995), we scored the Modern Sexism Scale so that high scores indicated more evidence of modern sexism. Swim et al. scored the scale in the opposite direction.

Table 4
Acquisition and Recognition Sentences Used in Study 5

Acquisition sentence	Stereotype-consistent interpretation used for gender set	
	A	B
1. Jane (Bill) administered the medicine to the patient.	Jane, the nurse, administered the medicine to the patient.	Bill, the doctor, administered the medicine to the patient.
2. Melanie (Don) got into a fight with her boyfriend (his girlfriend) about how much she (he) wanted to have sex.	Melanie got into a fight with her boyfriend about how little she wanted to have sex.	Don got into a fight with his girlfriend because he wanted to have sex more often.
3. Carol (Bob) didn't like Bob's (Carol's) attitude toward sports.	Carol didn't like Bob's enthusiastic attitude toward sports.	Bob didn't like Carol's negative attitude toward sports.
4. Gloria (Mike) argued with Mike (Gloria) about how much he (she) was committed to the relationship.	Gloria argued with Mike about how little he was committed to the relationship.	Mike argued with Gloria about her being committed to the relationship too much.
5. Cindy (Richard) was concerned every day about her (his) physical condition.	Cindy was concerned every day about her physical appearance.	Richard was concerned every day about his physical fitness.
6. Archie's (Edith's) friends were amazed at the shape he (she) kept his (her) room in.	Archie's friends were amazed at how messy he kept his room.	Edith's friends were amazed at how neat she kept her room.
7. The women (men) at the office liked to talk around the water cooler.	The women at the office liked to gossip around the water cooler.	The men at the office liked to talk sports around the water cooler.
8. Dick (Jane) was unhappy about the amount of liquor being served at the party.	Dick was unhappy about the small amount of liquor being served at the party.	Jane was unhappy about the large amount of liquor being served at the party.
9. When Jack (Jill) found out that his (her) friend had been murdered, he (she) became very upset.	When Jack found out that his friend had been murdered, he became very angry.	When Jill found out that her friend had been murdered, she became very sad.
10. Elizabeth (Bob) was not very surprised upon receiving her (his) math SAT scores.	Elizabeth was not very surprised upon receiving her low math SAT scores.	Bob was not very surprised upon receiving his high math SAT scores.
11. Laura (Luke) had a problem with expressing her (his) emotions.	Laura had a problem with expressing her emotions too much.	Luke had a problem with not expressing his emotions.
12. Paul (Linda), the lawyer, made a plea to the jury.	Paul, the lawyer, made a logical plea to the jury.	Linda, the lawyer, made an emotional plea to the jury.

Note. In the far left column, the first gender listed in parentheses is the gender used in Set A, and the second is the one used in Set B. Sentences 1–6 comprise Interpretation Set 1, and Sentences 7–12 comprise Interpretation Set 2. The interpretations listed in Columns 2 and 3 include the gender that render them stereotype consistent. These sentences were made stereotype inconsistent by replacing the stereotype included with the alternative (e.g., exchanging *Bill* and *Jane* in Sentence 1).

once they are given such individuating information, taking the data at face value, or do they use their stereotypes to alter their impressions of other individuals? In the five studies we conducted, we found evidence that people use their stereotypes as they encounter individuating information about other people. Indeed, we found that stereotypes, in a sense, may render the information given as not "individuating" at all. For example, with the same description of an altercation in a restaurant, stereotypes can lead people to believe that two lumberjacks had a fist fight but that two marriage counselors had only a verbal spat. When told that a person is worried about her physical condition, stereotypes can lead people to believe that a fashion model is vain but that a triathlete is health conscious. After these tacit inferences, the information that people are given is altered so that it is no longer the "same" information. People extract a specific meaning from the information provided that confirms their stereotype.

Five studies provided convergent support for these assertions. In Study 1, participants read passages about stereotyped individuals and then were given a recognition test. Participants falsely recognized sentences that were consistent with their stereotypes to a greater degree than they recognized sentences that were inconsistent with their stereotypes. Study 2 replicated those effects and found that they were not the byproduct of response biases. Study 3 further demonstrated that these inferences were not a product of asking participants about their impressions of the individuals depicted in the sentences. Study 5 found similar recognition memory effects to sentences containing tacit inferences inspired by the gender of the protagonist, an effect that occurred to an equal degree for individuals who scored high or low on a test of sexist thought. Study 4 provided convergent evidence for stereotypical inference in a cued-recall procedure.

Study 4 also provided evidence that participants made such inferences while encoding the original information at acquisition. Providing participants with stereotype-consistent inferences at the time of recall facilitated their memory for those sentences. They recalled a greater number of sentences with this type of cue than they did with stereotype-inconsistent cues or no cue at all. Because these stereotype-consistent inferences successfully aided recall, it can be assumed that they had been made at the time when people read the original sentence. If such inferences were made later, such as at the time of the recall test, they would have failed to facilitate accurate recall.

Questions for Future Research

The five studies described in this article leave many open questions to be addressed by further research. For example, what specific kinds of inferences do stereotypes commonly and effectively influence? In the present research we did not attempt to ascertain the types of inferences that stereotypes might alter, or the types of inferences that stereotypes leave unaffected. Instead, we created stimulus sentences that contained a number of different kinds of ambiguities. We presented words that carried different lexical meanings (e.g., does *convictions* stand for beliefs or a prison record?), sentences that left the stimulus to which characters were reacting unclear (e.g., was the nun unhappy about too much or too little alcohol?), or phrases that could fit a variety of specifications (e.g., *feelings* can refer to positive as well as negative emotions).

Future research could profit by creating a typology of differing types of tacit inference that may occur when people encounter information about their social world, and then exploring which types of inferences are influenced by stereotypes. In this regard, it may be helpful to monitor the growing work in cognitive psychology on text comprehension (e.g., Graesser, Singer, & Trabasso, 1994; McKoon & Ratcliff, 1992), which is struggling over similar issues. These researchers are striving to enumerate all the various categories of inferences that people might make, and seeing if people make them naturally as they comprehend text. As such, this research is relevant to the present set of studies, as the type of text material participants confront in the text comprehension work bears a strong resemblance to our methodology.

However, we do not propose that the processes we have described here are constrained to text; they can occur for visual stimuli as well. Is Mona Lisa's smile (or is it a smirk?) one of wryness, discomfort, or coquettishness? Facial expressions, voice tone, gestures, and actions are often ambiguous. They require context in order to be interpretable. Part of that context may be the stereotype associated with someone's group (see Biernat & Manis, 1994; Biernat, Manis, & Nelson, 1991, for related arguments). Keltner (1995), for example, found that people more readily identify facial expressions as embarrassment, shame, or anger when they come from demographic groups traditionally of low status (e.g., African Americans) as opposed to those of high status (e.g., Whites). Such differences in perception occur even though the faces from the different groups were chosen to reflect similar levels of movement and intensity. Similarly, people differ in their perceptions of when push comes to shove, perceiving more aggression in behaviors performed by African Americans than by Whites (Sagar & Schofield, 1980).

On Identification Processes

The present research is also relevant to Trope's (1986) model of social judgment. According to his model, social judgment is a two-step process, involving (1) *identification* of the behavior to be judged and (2) *inferences* about the dispositions of the person doing the behavior. In his model, and in the experiments he has conducted to test the model, Trope described the identification task as one of classification. The social perceiver must take a stimulus, such as seeing someone weeping, and classify it as a happy or sad act. The five studies we report here, however, suggest that there are many other tasks to be completed in the identification stage of social judgment. People may have to specify the exact nature of the behavior (e.g., was the fight a fist fight or merely a verbal spat?), the meaning of vague terms (e.g., does *physical condition* refer to one's appearance or health?), and the exact stimulus to which the individual is reacting (e.g., is the person upset because there is too much or too little liquor?). They may infer the emotional state or attitude of others (e.g., are the person's feelings positive or negative toward the Hollywood actor?), their likely personality traits (e.g., what about the Marine drill sergeant's personality makes it hard to take?), or the causal antecedents of their behavior (e.g., why did people think the politician's statements were untrue?).

This is not to suggest that Trope (1986) is inaccurate in his characterization of the identification process. Far from it, for identification processes obviously require people to classify the behaviors to which they are exposed. Rather, our analysis just expands the types of judgments or inferences people must make during the identification process, before they can begin to reach any conclusions about the dispositions possessed by the person they are judging. In short, our experiments heighten the importance of the identification stage and suggest that it should receive more scrutiny in work on social cognition.

Are Tacit Inferences Implicit?

The five studies we conducted suggest that stereotypes prompt tacit inferences that are spontaneous in nature. That is, they occur at the time that participants encounter information about others. Tacit inferences also occur without prompting by an outside agent. In future work, it would be profitable to address whether tacit inferences inspired by stereotypes are also *implicit,* that is, made without intention or awareness of the perceiver (Greenwald & Banaji, 1995). This consideration is important, for the scope and impact of stereotypes and tacit inferences in social judgment may hinge on whether they are implicit.

For example, consider the work of Devine (1989) on automatic versus controlled components of prejudice. According to Devine, almost all people possess stereotypes that can be automatically activated when the "appropriate" stimulus person appears. What distinguishes high- versus low-prejudiced people is whether they succeed at inhibiting the effects of those stereotypes. Low-prejudiced individuals exert effort to countervail the influence of the unwanted stereotype. However, consider the possibility that tacit inferences occur implicitly, outside of conscious control or awareness. If that is the case, low-prejudiced individuals may never have a chance to counteract the effects of stereotypes on tacit inferences. The "meaning" of the stimulus might be determined before low-prejudiced individuals have a chance to negate it, before they have an opportunity to identify and consider other possible and less stereotypical interpretations. If that is the case, then their judgments and actions may carry a good deal of prejudice even though those individuals have no intention to discriminate. Indeed, with every intention *not* to discriminate, the operation of these inference processes may constrain even those low-prejudiced individuals to consider

stimuli that "look" stereotypical, thus influencing their responses. Results from Study 5 suggest that this process may occur for low-prejudiced people. In that study, participants who scored low on a measure of modern sexism made as many stereotypical tacit inferences about men and women, as indexed by recognition memory errors, as did participants who scored higher on the measure.

The presence of tacit inferences also carries implications for any well-intentioned attempt to rid oneself of stereotypes. One plausible way to rid oneself of a stereotype is to garner individuating information about members of a stereotyped group. If one found out how "they" behave in concrete, specific situations, one would discover whether members of a stereotyped group behaved in a stereotypical way. However, consider the impact that tacit inferences, made automatically and outside awareness, may have on such an enterprise. If one's interpretation of individuating information is shaped by tacit inferences, one might be left with information and an impression that confirms the stereotype. As a consequence, even when well intentioned, one may confirm one's stereotype when trying to disprove it by gathering individuating information.

Although we have no specific data on whether the tacit inferences made by our participants were implicit, two findings from the five studies suggest that they may have been. First, we tested for tacit inferences by examining participants' *memory* for stimulus sentences. If tacit inferences were not implicit, that is, if they were made under conscious control, we can presume that participants would have been aware of this fact. As such, they would have had little trouble recognizing that the stereotype-consistent interpretations in the recognition test were alterations of the original sentences, ones that just happened to contain the same inferences they themselves had mindfully made. Second, in Study 3 we observed tacit inferences even in a condition that did not require participants to make any tacit inferences whatsoever. That is, when participants were merely asked to assess the readability of stimulus sentences, their memory of those sentences was still distorted by stereotype-based tacit inferences.

Finally, in Study 5 we found that low-sexist participants made gender-related tacit inferences at virtually the same rate as their high-sexist counterparts. If the production of stereotypical tacit inferences were under the control of the individual, one would assume that this group would have been less likely than their high-sexist counterparts to make such inferences. After all, it is safe to presume that this group would be the most motivated to think about men and women in an egalitarian manner, and not to make inferences that women are gossips, are emotional, score low on math tests, are more committed to relationships than men, are less interested in sex, more likely to be teetotalers, more concerned about their physical appearance, and to assume that a stimulus woman must be a nurse. However, their answers on the recognition memory test suggested that they were just as likely as their high-sexist counterparts to make such inferences.

However, all these observations are only suggestive, not conclusive, evidence of the implicit nature of these tacit inferences. Further, and more rigorous, evidence is necessary before we can conclude that the tacit inferences we observed were made implicitly, outside the control or awareness of the individual making them.

Concluding Remarks

Life is fraught with ambiguities. Although this fact is one all people recognize in the abstract, it is one people may often miss in their day-to-day affairs. When people describe themselves or others, they often fail to recognize that the common and mundane terms they use (e.g., *She is intelligent, I have good leadership skills*) are indeterminate in their meaning. The net result of this lack of recognition is judgmental bias (Dunning, Meyerowitz, & Holzberg, 1989; Gilovich, 1990; Griffin, Dunning, & Ross, 1990) and interpersonal disagreement (Dunning & Cohen, 1992; Dunning & McElwee, 1995; Hayes & Dunning, 1997). The five studies described in this article suggest another way in which ambiguity, unrecognized, may play a role in thought and judgments about others. When information about another person is indeterminate in meaning, people may fill in ambiguities and details based on stereotypical cues about that person.

Thought of in this way, stereotypes may confer both the benefits and the costs that prisons provide for their inmates. On the benefit side, much like prisons guide and constrain the behaviors of prisoners in presumably helpful ways, stereotypes may similarly guide and constrain people in the interpretations they can make about other individuals out of all the infinite number of interpretations that are possible in human life. However, this benefit in interpretation may carry some obvious costs. Real prisons provide a life for their prisoners that hardly resembles life as it looks like in the real world. Similarly, the interpretations of behavior prompted by stereotypes may provide people with impressions of other people's behavior that does not resemble what that behavior looks like in reality.

References

Anderson, J. R., & Reder, L. M. (1979). An elaborative processing explanation of depth of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 385–403). Hillsdale, NJ: Erlbaum.

Ashmore, R. D. (1981). Sex stereotypes and implicit personality theory. In D. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 37–81). Hillsdale, NJ: Erlbaum.

Bassili, J. N., & Smith, M. C. (1986). On the spontaneity of trait attribution: Converging evidence for the role of cognitive strategy. *Journal of Personality and Social Psychology, 50,* 239–245.

Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology, 66,* 5–20.

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology, 60,* 485–499.

Bransford, J. D., Barclay, J. R., & Franks, J. J. (1972). Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology, 3,* 193–209.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5–18.

Dunning, D., & Cohen, G. L. (1992). Egocentric definitions of traits and abilities. *Journal of Personality and Social Psychology, 63,* 341–355.

Dunning, D., & McElwee, R. O. (1995). Idiosyncratic trait definitions: Implications for self-description and social judgment. *Journal of Personality and Social Psychology, 68,* 936–946.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity

and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 59,* 1082–1090.

Eysenck, M. W., Mogg, K., May, J., Richards, A., & Mathews, A. (1991). Bias in interpretation of ambiguous sentences related to threat in anxiety. *Journal of Abnormal Psychology, 100,* 144–150.

Futoran, G. C., & Wyer, R. S., Jr. (1986). The effects of traits and gender stereotypes on occupational suitability judgments and the recall of judgment-relevant information. *Journal of Experimental Social Psychology, 22,* 475–503.

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. (1992). Do readers mentally represent character's emotional states? *Cognition & Emotion, 6,* 89–112.

Gilovich, T. (1990). Differential construal and the false consensus effect. *Journal of Personality and Social Psychology, 59,* 623–634.

Glick, P., Zion, C., & Nelson, C. (1988). What mediates sex discrimination in hiring decisions? *Journal of Personality and Social Psychology, 55,* 178–186.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review, 101,* 371–395.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102,* 4–27.

Griffin, D. W., Dunning, D., & Ross, L. (1990). The role of construal processes in overconfident predictions about the self and others. *Journal of Personality and Social Psychology, 59,* 1128–1139.

Hayes, A. F., & Dunning, D. (1997). Construal processes and trait ambiguity: Implications for self–other agreement in personality judgment. *Journal of Personality and Social Psychology, 72,* 664–678.

Heilman, M. E. (1984). Information as a deterrent against sex discrimination: The effects of applicant sex and information type on preliminary employment decisions. *Organizational Behavior and Human Performance, 33,* 174–186.

Johnson, M. K., Bransford, J. D., & Solomon, S. K. (1973). Memory for tacit implications of sentences. *Journal of Experimental Psychology, 98,* 203–205.

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. *Journal of Personality and Social Psychology, 68,* 441–454.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95,* 163–182.

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29,* 133–159.

Krueger, J., & Rothbart, M. (1988). Use of categorical and individuating information in making inferences about personality. *Journal of Personality and Social Psychology, 55,* 187–195.

Kunda, Z., & Sherman-Williams, B. (1993). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin, 19,* 90–99.

Locksley, A., Borgida, E., Brekke, N., & Hepburn, C. (1980). Sex stereotypes and social judgment. *Journal of Personality and Social Psychology, 39,* 821–831.

Locksley, A., Hepburn, C., & Ortiz, V. (1982). Social stereotypes and judgments of individuals. *Journal of Experimental Social Psychology, 18,* 23–42.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review, 99,* 440–466.

O'Brien, E. J., Shank, D. M., Myers, J. L., & Rayner, K. (1988). Elaborative inferences during reading: Do they occur on-line? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 410–420.

Sagar, H., & Schofield, J. W. (1980). Racial and behavioral cues in black and white children's perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology, 39,* 590–598.

Singer, M., & Ferreira, F. (1983). Inferring consequences in story comprehension. *Journal of Verbal Learning and Verbal Behavior, 22,* 437–448.

Slusher, M. P., & Anderson, C. A. (1987). When reality monitoring fails: The role of imagination in stereotype maintenance. *Journal of Personality and Social Psychology, 52,* 653–662.

Small, S. I., Cottrell, G. W., & Tanenhaus, M. K. (1988). *Lexical ambiguity resolution.* San Mateo, CA: Morgan Kaufman.

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and racism: Old-fashioned and modern prejudices. *Journal of Personality and Social Psychology, 68,* 199–214.

Till, R. E., Mross, E. F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory & Cognition, 16,* 283–298.

Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review, 93,* 239–257.

Trope, Y., Cohen, O., & Alfieri, T. (1991). Behavior identification as a mediator of dispositional inference. *Journal of Personality and Social Psychology, 61,* 873–883.

Uleman, J. S., & Moskowitz, G. B. (1994). Unintended effects of goals on unintended inferences. *Journal of Personality and Social Psychology, 66,* 490–501.

Broek, P. van den (1990). Causal inferences and the comprehension of narrative text. In A. C. Graesser & G. H. Bower (Eds.), *Inferences and text comprehension* (pp. 175–196). San Diego, CA: Academic Press.

Hippel, W. von, Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 117–254). San Diego, CA: Academic Press.

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47,* 237–252.

Winter, L., Uleman, J.S., & Cunniff, C. (1985). How automatic are social judgments? *Journal of Personality and Social Psychology, 49,* 904–917.

Woll, S., Weeks, D. G., Fraps, C. L., Pendergrass, J., & Vanderplas, M. A. (1980). Role of sentence context in the encoding of trait descriptors. *Journal of Personality and Social Psychology, 39,* 59–68.

Woll, S., & Yopp, H. (1978). The role of context and inference in the comprehension of social action. *Journal of Experimental Social Psychology, 14,* 351–362.

Received March 21, 1996
Revision received December 16, 1996
Accepted December 23, 1996 ■