Journal of Personality and Social Psychology
2003, Vol. 84, No. 3, 470–484

0022-3514/03/$12.00 DOI: 10.1037/0022-3514.84.3.470

When Stereotypes Get in the Way: Stereotypes Obstruct Stereotype-Inconsistent Trait Inferences

Daniël H. J. Wigboldus and Ap Dijksterhuis
University of Amsterdam

Ad van Knippenberg
University of Nijmegen

There is a growing body of evidence indicating that people spontaneously make trait inferences while observing the behavior of others. The present article reports a series of 5 experiments that examined the influence of stereotypes on the spontaneous inference of traits. Results consistently showed weaker spontaneous trait inferences for stereotype-inconsistent behavioral information than for stereotype-consistent and stereotype-neutral information. Taken together, the current results suggest that specific spontaneous trait inferences become obstructed by inhibitory processes when behavior is inconsistent with an already activated stereotype. These findings are discussed in relation to stereotype maintenance processes and recent models of attribution in social judgment.

Empirical evidence from research on impression formation (see Gilbert, 1998) and the spontaneous inference of traits (see Uleman, Newman, & Moskowitz, 1996) consistently shows that people infer traits from behaviors of others. Thus, when one learns that the secretary solves the mystery halfway through the book, one spontaneously infers that she must be clever (e.g., Winter & Uleman, 1984). However, although people may often readily infer specific intentions or traits of actors on observing their behaviors, these inference processes may not occur unconditionally. Some actors may, on the basis of their category membership, be more readily ascribed intentions or traits that correspond to their behavior than are others. For instance, would a garbage man who solves the mystery halfway through the book also spontaneously be regarded as clever?

As yet, research into spontaneous trait inferences (STIs) has focused mainly on behaviors, whereas the actors of these behaviors have remained nondescript (e.g., John, Bill) or are described by an irrelevant category label (see, for an exception, Sherman, Lee, Bessenoff, & Frost, 1998). In real life, however, in most cases people have considerably more knowledge of actors than only their names. Gender, age, and skin color, for instance, are salient features that perceivers use to rapidly categorize a person (see Fiske, 1998). On the basis of such categorizations, people automatically activate stereotypical traits (e.g., Devine, 1989; Dijksterhuis & van Knippenberg, 1996; von Hippel, Sekaquaptewa, & Vargas, 1995). For example, the category label *garbage man* may automatically activate the trait *unintelligent*.

The main goal of the present article is to combine insights from research on automatic stereotype activation with research on STIs. That is, to develop a more comprehensive approach to STIs, we aim to take into account not only individuating but also categorical information processing. We propose that the STI process is impeded in the case of unexpected or stereotype-inconsistent behavior. Thus, when a garbage man solves the mystery halfway through the book, people are less likely to spontaneously infer the trait *clever* than when a secretary or a professor shows the same behavior. The result is that people's stereotype of garbage men may conveniently stay intact. In the following, we first review recent evidence for the spontaneous inference of traits from behaviors. Then we further unfold our argument concerning the influence of stereotype activation on the spontaneous encoding of behaviors of others.

STIs

Although most research within person perception has focused on dispositional inferences as a result of intentional causal reasoning (for an overview, see Gilbert, 1998; Gilbert & Malone, 1995), it has always been acknowledged that impressions of others are formed with "remarkable rapidity and great ease" (Asch, 1946, p. 258) and may also be invoked spontaneously and unintentionally (e.g., Asch, 1946; Gilbert, Pelham, & Krull, 1988; Heider, 1958; Uleman, Newman, & Moskowitz, 1996; Winter & Uleman, 1984). STIs are said to occur "when attending to another person's behavior produces a trait inference in the absence of our explicit intention to infer traits or form an impression of that person" (Uleman, Newman, & Moskowitz, 1996, p. 211). Over the past 2 decades, ample empirical evidence for the occurrence of STIs has been obtained (for an overview, see Uleman, Newman, & Moskowitz, 1996), through research paradigms such as the cued recall task (e.g., Winter & Uleman, 1984; Winter, Uleman, & Cunniff, 1985)

Daniël H. J. Wigboldus and Ap Dijksterhuis, Department of Social Psychology, University of Amsterdam, Amsterdam, the Netherlands; Ad van Knippenberg, Department of Social Psychology, University of Nijmegen, Nijmegen, the Netherlands.

This research was supported by the Netherlands Organization for Scientific Research Grant PPS 98-031. We thank Jeff Sherman and Kerry Kawakami for their helpful comments at various stages of this research program and Clemens Wenneker and Daniël Haveman for their help in collecting the data for Study 5.

Correspondence concerning this article should be addressed to Daniël H. J. Wigboldus, Department of Social Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, the Netherlands. E-mail: dwigboldus@fmg.uva.nl

and the relearning paradigm (e.g., Carlston & Skowronski, 1994; Carlston, Skowronski, & Sparks, 1995).

In a recent set of studies (Uleman, Hon, Roman, & Moskowitz, 1996), a research paradigm was used that specifically aimed to demonstrate the spontaneous occurrence of trait inferences at the encoding of behavioral information. In this so-called recognition probe paradigm (McKoon & Ratcliff, 1986), participants were presented with trait-implying and control sentences on a computer screen. Each sentence was immediately followed by a probe word. The participant's task was to indicate as quickly and accurately as possible whether the probe word was part of the preceding sentence. In the experimental trials, probe words were traits that were implied by the trait-implying sentences (e.g., "She couldn't get herself to greet her new neighbor," followed by *shy*). Following McKoon and Ratcliff's (1986) reasoning concerning inferences about predictable events, Uleman and colleagues suggested that when a trait is spontaneously inferred at the encoding of a sentence, it becomes relatively harder to indicate that this trait was not present in the sentence. Activation of a relevant trait at encoding is thus assumed to interfere with the subsequent trait probe task. This is exactly what was found. Trait-implying sentences resulted in more errors (Uleman, Hon, et al., 1996, Experiment 1) and longer reaction times (RTs; Uleman, Hon, et al., 1996, Experiments 2 and 3) than control sentences did.

Research based on the recognition probe paradigm seems to provide convincing evidence in favor of the assumption that traits are spontaneously inferred at the encoding of behavioral information. These trait inferences are said to be spontaneous, because making them actually hampered participants' performance on the task at hand (i.e., indicating that a word was not part of the preceding sentence). The finding that participants nevertheless made these inferences constitutes a strong argument in favor of the spontaneity of STIs. The inferences are said to occur at the encoding stage, because in this research paradigm trait inferences were measured directly after encoding instead of after a delay (such as is the case in the cued recall task and the savings in relearning paradigm). Still, the current recognition probe paradigm may not be conclusive with respect to the encoding issue because it does not discriminate between inferences that occur while the participant is reading the behavioral sentence and inferences that occur during the testing of the trait probe (for a similar argument, see Keenan, Potts, Golding, & Jennings, 1990). To overcome this problem, researchers have applied response window techniques to the recognition probe paradigm to be able to force participants to respond so quickly that one can be quite confident that most conscious retrieval processes are eliminated (e.g., McKoon & Ratcliff, 1986, Experiment 4; van Overwalle, Drenth, & Marsman, 1999, Experiment 3). Although not conclusive, these experiments provided further evidence that inferences are made spontaneously at encoding of the behavioral information and are not the result of backward integration processes (i.e., processes that take place during the testing of the trait probe and that aim to integrate the trait probe with the preceding behavioral sentence). In summary, there is converging evidence from several research paradigms for the assumption that people spontaneously make a trait inference when confronted with the behavior of others. These STIs seem to occur automatically at the encoding of behavioral information. In general, observers spontaneously seem to encode "doing" as "being" when confronted with the behaviors of others.

Stereotyping and Trait Accessibility

Surprisingly little attention has been paid to the influence of the social context on the occurrence of STIs (see Uleman, Newman, & Moskowitz, 1996). Thus far, STIs have been investigated in what might be called a social vacuum. That is, stimulus actors in most STI research have remained nondescript persons who perform a specific trait-implying behavior. In real life, however, in most cases one knows much more about an actor than just his or her behavior. Social categories such as gender, age, and ethnicity have been shown to be available to perceivers automatically within milliseconds (see Fiske, 1998, for a review). Moreover, stereotypes associated with these categories seem to be activated automatically in the presence of a category member (e.g., Devine, 1989; Dijksterhuis & van Knippenberg, 1996; von Hippel et al., 1995), even when the presentation of the category information is nonconscious (e.g., Perdue & Gurtman, 1990).

In research on stereotype activation, stereotypes are mostly interpreted as mental representations in which a social category is associated with traits that are stereotypical for this category (see Stangor & Lange, 1994). Although recent research has demonstrated that the activation of a category facilitates the access to stereotype-consistent trait terms (e.g., Devine, 1989; Dijksterhuis & van Knippenberg, 1996; Gilbert & Hixon, 1991; Lepore & Brown, 1997; Macrae, Bodenhausen, & Milne, 1995; Macrae, Stangor, & Milne, 1994), it has also shown that category activation inhibits the access to stereotype-inconsistent trait terms (e.g., Dijksterhuis & van Knippenberg, 1995, 1996). For example, on activation of the category label *professor*, consistent trait terms such as *smart* are activated and access to inconsistent trait terms such as *aggressive* is inhibited.

Stereotypes and STIs

Apart from the spontaneous inferences made on the basis of the behavior an actor displays, observers are likely to make stereotypical inferences automatically on the basis of the actor's most salient category memberships. The primary question in the present article is whether stereotypes associated with an actor influence STIs related to the behavior of the actor. As yet, this question has received little theoretical and empirical attention (see Uleman, Newman, & Moskowitz, 1996). The answer to this question, however, is important, because it not only leads to a better and more valid understanding of STIs in a social context but also provides more insight into the way stereotypes exert their self-maintaining influence at the first stage of information processing, during the encoding of social information (von Hippel et al., 1995; see also van Knippenberg & Dijksterhuis, 2000). As yet, most research on the influence of stereotypes on attributions has focused on explicit judgments (e.g., Jackson, Sullivan, & Hodge, 1993). Using a more implicit measure—namely, language use—Maass and others (Maass, Milesi, Zabbini, & Stahlberg, 1995; Wigboldus, Semin, & Spears, 2000) have found that stereotypes influence linguistic attributions. They showed that people describe stereotype-consistent behavior at a more abstract level (i.e., using relatively more traits) than they describe stereotype-inconsistent behavior. Although in this research the focus is on language use and not on encoding processes, these findings are consistent with our present argument.

We propose that by inhibiting or facilitating access to specific traits, stereotypes may respectively impede or foster the STI process. Our reasoning is based in part on research demonstrating that the accessibility of trait terms influences the relative use of these traits in subsequent impression formation tasks (e.g., Bargh, Lombardi, & Higgins, 1988; Bargh & Pietromonaco, 1982; Higgins, Rholes, & Jones, 1977; Srull & Wyer, 1979). For instance, Higgins et al. (1977) demonstrated that prior activation of the word *brave* caused participants to think of an actor's behavior as adventurous instead of reckless. Some trait constructs may become temporarily more accessible than other relevant ones if they have been used or primed recently. The specific traits that are used to interpret behavior partly seem to depend on this relative accessibility of available trait domains (e.g., Bargh & Pietromonaco, 1982; Higgins et al., 1977).

In line with Dijksterhuis and van Knippenberg (1996), we argue that the activation of a stereotype prior to the encoding of behavioral information not only makes stereotype-consistent traits temporarily more accessible but also makes stereotype-inconsistent traits temporarily less accessible. When a behavior is consistent with an activated stereotype, the heightened accessibility of consistent traits does not interfere with the STI process; it may even foster this process. However, inferring a trait that is inconsistent with an already activated stereotype may be harder, as the trait that has to be inferred is temporarily inaccessible. In the latter case, the STI process becomes more difficult or may even be effectively obstructed, because the activated stereotype renders the implied trait concept temporarily less accessible.[1] Thus, although, in general, a person who correctly answers a Trivial Pursuit question is likely to be perceived as relatively smart, the prior knowledge that this person is a garbage man may inhibit the concept of *smart*. Therefore, one is less likely to spontaneously infer the trait *smart* in the case of a garbage man giving a correct answer. As we have noted, in this way, one's stereotype of garbage men may conveniently stay intact. In the following, we present five studies that were designed to test the encoding process outlined above.

Study 1

The research paradigm used in our studies is based on the recognition probe paradigm (see McKoon & Ratcliff, 1986; Uleman, Hon, et al., 1996). On the computer screen, participants were presented with sentences describing concrete behaviors. The category labels with which the actors in these behavioral sentences were described were either consistent or inconsistent with respect to the behavior described in the sentences (e.g., "The professor [garbage man] wins the science quiz"). In all studies, each sentence was followed by a probe word, for which participants were asked to indicate as quickly and accurately as possible whether the word was included in the preceding sentence. In the STI trials, probe words were traits that were implied by the behavioral sentences (e.g., *smart*) but were not actually included in these sentences.

In our first study, we presented participants with stereotype-consistent and stereotype-inconsistent sentences. In line with the encoding process outlined above, we expected participants to make stronger STIs on the basis of stereotype-consistent behavioral information than on the basis of stereotype-inconsistent information. Thus, we hypothesized that, in general, participants would have more difficulty performing their task and would thus respond more slowly to relevant trait probes when the behavior in the preceding sentence was stereotype-consistent with the category label presented than when this behavior was stereotype-inconsistent with the category label presented.

Method

Participants

A total of 42 Dutch undergraduate students (14 men and 28 women) from the University of Nijmegen participated in this study ($M = 22.24$ years old). Participants were recruited on campus and received 5 Fl (approximately $2.50) for their participation. All participants were native Dutch speakers.

Overview and Design

Participants were presented with sentences that were followed by probe words and were asked to indicate as quickly as possible whether these words were part of the preceding sentence. Critical STI trials consisted of behavioral sentences (verb sentences) that were either stereotype-consistent or stereotype-inconsistent with the category label presented (e.g., "The skinhead [girl] hits the saleswoman") and were followed by a relevant trait probe (e.g., *aggressive*). Several types of additional trials were presented to prevent the correct response always being "no." First, verb-sentence trials were added in which the verb sentences were the same as in the STI trials but the probe word following these verb sentences was a repetition of the verb in the sentence (e.g., *hit*) and, thus, a "yes" answer was correct. Second, stereotype-consistent and stereotype-inconsistent trait-sentence trials were added (e.g., "The skinhead [girl] is aggressive") that were followed by either a relevant verb probe (e.g., *hit*) or trait probe (e.g., *aggressive*). Note that for these trait sentences, the verb probe elicited the "no" answer, and the trait probe elicited the "yes" answer.

The trials that participants reacted to thus consisted of stereotype-consistent and stereotype-inconsistent sentences that were formulated either in concrete terms, with action verbs, or in abstract terms, with traits as adjectives (see Semin & Fiedler, 1991), and were followed by either a verb or a trait probe. The experiment consisted of a 2 (stereotype consistency: consistent vs. inconsistent) $\times$ 2 (sentence type: verb vs. trait) $\times$ 2 (probe type: verb vs. trait) within-subject design.

Procedure

The experiment was run on Apple computers (iMac 266 MHz, with a monitor refresh rate set at 117 Hz) and was described to participants as a study on reading speed. All participants were instructed individually by the experimenter. Before the actual experiment started, two practice rounds were run. For the first practice round, participants were asked to read short sentences that appeared (in black letters) in the center of the screen and stayed on the screen for 1,000 ms. After a blank screen of 500 ms, a probe word appeared (in blue letters) in the center of the screen. Participants' task was simply to indicate as quickly and accurately as possible whether this probe word was present in the preceding sentence. Participants could indicate a "yes" answer by pressing the green-labeled *6* key on the numeric

[1] It should be noted that we do not mean that people do not make any kind of STIs when confronted with behaviors that imply a trait that is stereotype inconsistent. Rather, we argue that the spontaneous inference of this particular trait may become obstructed by the prior activation of the stereotype. If there exists an alternative trait implication of the behavior that is not stereotype inconsistent, the spontaneous inference of this alternative trait may not be affected by the activated stereotype.

pad of the keyboard and a "no" answer by pressing the red-labeled *A* key on the keyboard. Participants were instructed to position their hands over those keys throughout the experiment. After each response, there was a blank screen of 1,000 ms before the next sentence appeared on the screen. The first practice round consisted of 10 such trials.

For the second practice round, participants were asked to perform the same task. However, this time only half of the sentences were followed by a probe word. That is, each sentence was followed by either a blank screen for 500 ms and then a probe word or just a blank screen for 1,000 ms and then the next sentence. Therefore, participants could not predict when a probe would appear. The second practice round consisted of 10 sentences with and 10 without a probe word in random order.

After the two practice rounds, the experimenter left the participants alone to complete the experimental trials. In total, participants went through 96 trials with a probe word and 120 trials without a probe word. We randomly divided these trials into six trial groups, each consisting of 16 trials with and 20 trials without a probe word. The presentation order of the trials within each trial group was randomized. To reduce the effects of fatigue, we paused the experiment halfway through the trials, and participants were allowed to take a 2-min break. Half of the participants went through Trial Groups 1–3 before the break and proceeded with Trial Groups 4–6. For the other half of the participants, this order was reversed. Within each half of the experiment, the order of the three trial groups was also randomized. Moreover, to prevent participants from becoming too bored with the tedious task they had to perform, we promised to reward the fastest, most accurate participant with a book token of 25 Fl (approximately $12.50). To avoid start-up problems, the first 5 trials of the experiment before and after the break were filler trials consisting of two sentences with and three sentences without a probe word.

Stimuli

Experimental trials. On the basis of pretesting, six trait-implying sentences were developed, each describing one specific behavior of an actor. For each sentence, two category labels were selected, one for which the behavior in the sentence was stereotype consistent, and one for which the same behavior was stereotype inconsistent. For instance, the sentence "The actor hits the saleswoman," implying the trait *aggressive*, may be considered stereotype consistent for a skinhead but stereotype inconsistent for a girl. In a pilot study ($N = 19$), participants were asked to indicate on a 7-point scale to what extent each of the resulting 12 sentences was *stereotype consistent* (3) or *stereotype inconsistent* (−3). Results indicated that the stereotype-consistent sentences ($M = 2.30$, $SD = 0.39$) were significantly more stereotype consistent than the stereotype-inconsistent sentences ($M = -2.32$, $SD = 0.59$), $t(18) = 26.06$, $p < .001$ (see Table 1 for a list of the sentences and probes). In the same pilot study, participants were asked to indicate on a 7-point scale (1 = *not applicable*, 7 = *very applicable*) to what extent each of the six traits applied to each of the six behaviors. For all of the selected behavior–trait combinations, applicability scores were greater than 6. Moreover, in a small pilot study ($N = 8$) in which participants were asked to name three traits that come to mind after reading each sentence, selected traits were named by all participants.

We hypothesized that for stereotype-consistent behavioral information, people make STIs that make it harder to quickly indicate that these traits actually are not included in the sentences. To test this hypothesis, we planned to compare the RTs on stereotype-consistent verb sentences followed by a trait probe with the RTs on stereotype-inconsistent verb sentences followed by a trait probe. However, when only these types of trials are presented to participants, the correct answer to every trial is "no," which makes the task too simple for us to discriminate between the different answers. Therefore, we added several types of additional trials.

In total, 48 experimental trials were created in which the stereotype consistency of the sentences, the type of sentence, and the type of probe were varied systematically. As was described above, we manipulated stereotype consistency by varying the type of actor. For each of the 12 verb sentences that resulted from this manipulation, we developed a trait-sentence counterpart to manipulate the type of sentence. We created trait sentences using the traits implied by the behaviors described in the verb sentences. For instance, the sentence "The girl is aggressive" served as the trait-sentence counterpart for the verb sentence "The girl hits the saleswoman." In the pilot study mentioned earlier ($N = 19$), participants were also asked to indicate on a 7-point scale to what extent each of the resulting 12 trait sentences was *stereotype consistent* (3) or *stereotype inconsistent* (−3). Results indicated that the stereotype-consistent trait sentences ($M = 2.43$, $SD = 0.55$) were also judged as significantly more stereotype consistent than were the stereotype-inconsistent sentences ($M = -2.39$, $SD = 0.47$), $t(18) = 29.34$, $p < .001$.

To manipulate the type of probe, we followed each of these sentences with a literal repetition of the verb from the verb sentence (e.g. *hits*) or the trait from the trait-sentence counterpart (e.g., *aggressive*). In total, these

Table 1
Stimuli Used for the Experimental Trials in Studies 1–5 as a Function of Stereotype Consistency

Study and sentence	Probe
Consistent (Inconsistent)	
Studies 1 and 2	
The skinhead (girl) hits the saleswoman.[a]	Aggressive
The professor (garbageman) wins the science quiz.[a]	Smart
The boyscout (punk) helps the handicapped person.[a]	Helpful
The soccer fan (nurse) shouts at the waiter.	Rude
The priest (junkie) brings back the found purse.	Honest
The Moroccan (manager) comes home from work early.	Lazy
Consistent (Inconsistent)	
Study 2	
The girl (skinhead) comforts the shopowner.[a]	Sweet
The garbageman (professor) does not know the answer.[a]	Stupid
The punk (boyscout) pushes the tottler aside.[a]	Anti-social
The nurse (soccer fan) puts the elderly woman into bed.	Careful
The junkie (priest) puts the found money in his own pocket.	Dishonest
The manager (Moroccan) works late till midnight.	Industrious
Neutral	
Study 2	
The skinhead (girl) walks through town.	Aggressive (sweet)
The professor (garbageman) cycles through the street.	Smart (stupid)
The boyscout (punk) eats a sandwich.	Helpful (anti-social)
The soccer fan (nurse) buys a new coat.	Rude (careful)
The priest (junkie) collects his groceries.	Honest (dishonest)
The Moroccan (manager) orders a cup of coffee.	Lazy (industrious)

Note. In all studies, each sentence was used twice: once in combination with a consistent category label, and once with an inconsistent category label.
[a] Sentence was used in Studies 3, 4, and 5.

three manipulations resulted in 48 experimental trials. Because type of sentence and type of probe were manipulated orthogonally within these trials, participants were not able to predict from the type of sentence what kind of probe word would follow. Thus, participants were not able to predict from the sentence whether a "yes" or a "no" answer was required. Moreover, both types of probe words could elicit a "yes" as well as a "no" answer.

Filler trials. Because of the importance of stereotype consistency in the current experiment, it was crucial that participants pay attention to the actors in the sentences. However, for the 48 experimental trials described above, paying attention to only the verb or trait in the sentence is sufficient to perform the task. To make sure participants paid attention to the actors in the sentences as well, we created 48 filler trials using the 24 verb and trait sentences described above. In these filler trials, we used the category labels as probes (e.g., *skinhead* or *girl*) to make sure participants attended to the actor of each sentence as well.

Finally, 120 filler sentences were added that were not followed by a probe word. This was done to prevent participants from predicting when a probe word would appear (see, for a similar procedure, McKoon & Ratcliff, 1986; Uleman, Hon, et al., 1996). To serve this function, filler sentences had to describe the same kind of behaviors and actors as the experimental sentences, to prevent participants from learning to discriminate between sentences that are followed by a probe and sentences that are not. Filler sentences were thus constructed using all possible combinations among the 6 verb sentences, the 6 trait sentences, and the 10 category labels that were not used for each sentence in the experimental trials (e.g., "The Boy Scout comes home from work early").

Table 2
Mean Response Latencies in Milliseconds as a Function of Stereotype Consistency, Sentence Type, and Probe Type: Study 1

Probe type	Stereotype consistency: Consistent	Inconsistent
Verb sentences		
Verb probe		
M	831_{c}	$841_{b,c}$
SD	219	211
Trait probe		
M	913_{a}	$873_{b,c}$
SD	259	246
Trait sentences		
Verb probe		
M	871_{b}	$858_{b,c}$
SD	203	214
Trait probe		
M	741_{d}	747_{d}
SD	171	171

Note. $N = 42$. Cell means in rows and columns that do not share the same subscripts differ significantly from each other ($p < .05$).

Results

RTs

The principal dependent variable was the time (RTs in ms) it took participants to indicate that a probe word was or was not a part of the preceding sentence. Incorrect answers on this task were infrequent ($M = 4.12\%$, range = 0.00–16.67%) and were excluded from the RT analyses. RTs faster than 200 ms and slower than 2,000 ms were regarded as outliers and thus disregarded.[2] As a result of these cut-offs, only six responses (0.30%) had to be dropped from the statistical analysis.

The remaining RTs were analyzed in a 2 (stereotype consistency: consistent vs. inconsistent) × 2 (sentence type: verb vs. trait) × 2 (probe type: verb vs. trait) within-subject analysis of variance (ANOVA). A significant three-way interaction among stereotype consistency, sentence type, and probe type was obtained, $F(1, 41) = 5.65$, $p < .05$ (see Table 2). Analyses of the simple main effects for stereotype consistency revealed that, as predicted, only for verb sentences followed by a trait probe was there a significant difference, $F(1, 41) = 4.19$, $p < .05$. On average, it took participants 40 ms longer to indicate that the trait probe was not a part of the verb sentence when the sentence was stereotype consistent than when the sentence was stereotype inconsistent. For all the other cells of the design, this difference was not significant.

Besides the three-way interaction, a two-way interaction between sentence type and probe type was also found, $F(1, 41) = 37.87$, $p < .001$. It took participants less time to react when sentence type and probe type were the same ($M = 836$ ms, $SD = 209$ ms, for verb sentence and verb probe; $M = 744$ ms, $SD = 165$ ms, for trait sentence and trait probe) than when sentence type and probe type were different ($M = 893$ ms, $SD = 244$ ms, for verb sentence and trait probe; $M = 865$ ms, $SD = 204$ ms, for trait sentence and verb probe). Apparently, "no" answers resulted in longer RTs than did "yes" answers.

Finally, two main effects were found. Participants took longer to react to verb sentences ($M = 864$ ms, $SD = 219$ ms) than to trait sentences ($M = 804$ ms, $SD = 180$ ms), $F(1, 41) = 27.38$, $p < .001$. Also, participants took longer to react to verb probes ($M = 850$ ms, $SD = 201$ ms) than to trait probes ($M = 818$ ms, $SD = 196$ ms), $F(1, 41) = 18.52$, $p < .001$. No other effects were found.

Error Rates

As is the case with most counts of errors in cognitive or perceptual tasks (Cohen & Cohen, 1975), error rates were seriously skewed. A square root transformation on the error proportions was performed to reduce skew (see Cohen & Cohen, 1975). The resulting scores were analyzed in a 2 (stereotype consistency: consistent vs. inconsistent) × 2 (sentence type: verb vs. trait) × 2 (probe type: verb vs. trait) within-subject ANOVA. This analysis revealed a significant interaction effect between sentence type and probe type, $F(1, 41) = 5.47$, $p < .05$. Fewer errors were made when sentence type and probe type were the same ($M = 2.98\%$, $SD = 5.10\%$, for verb sentence and verb probe; $M = 2.98\%$, $SD = 4.40\%$, for trait sentence and trait probe) than when sentence type and probe type were different ($M = 4.56\%$, $SD = 8.10\%$, for verb sentence and trait probe; $M = 5.95\%$, $SD = 7.20\%$, for trait sentence and verb probe). Apparently, "no" answers resulted not only in longer RTs but also in more errors than did "yes" answers. No other effects were found.

[2] In line with Ratcliff's (1993) recommendations, we used two different methods to deal with RT outliers. That is, we confirmed the current analyses by using the inverse transformation (1/x) of all RTs. In all studies presented, these two different methods yielded a similar pattern of results.

Discussion

In line with our hypothesis, participants took more time to indicate that the relevant trait was not part of the preceding sentence when the category label and the behavior presented in the sentence were stereotype consistent than when they were stereotype inconsistent. For the number of errors participants made, no effects due to stereotype consistency were found. This does not seem surprising, considering the simplicity of the task and the low overall error rate.

It is interesting that effects due to stereotype consistency on participants' RTs only were found in the STI condition—that is, when concrete behavioral sentences were followed by trait probes. No differences between stereotype-consistent and stereotype-inconsistent sentences were found when behavioral sentences were followed by verb probes or when trait sentences were followed by verb or trait probes. In other words, stereotype-consistency effects only emerged when trait inferences were made on the basis of behavioral sentences. Results thus seem to confirm that participants made stronger STIs on the basis of stereotype-consistent behavioral information than on the basis of stereotype-inconsistent information.

An alternative explanation for the current results is that the difference in RTs between stereotype-consistent and stereotype-inconsistent sentences may be due solely to stereotype activation and not to the biased spontaneous trait encoding process outlined in the introduction. After all, the presentation of the category label may have automatically activated a relevant trait concept (e.g., Devine, 1989; Dijksterhuis & van Knippenberg, 1996), subsequently making it harder for participants to indicate that the trait probe was not in the sentence. In other words, did participants spontaneously infer the trait *smart* on the basis of the sentence "The professor wins the science quiz," or did activation of the stereotype of professors lead to automatic activation of this trait? Although a stereotype activation account of the Study 1 results fails to explain the finding that stereotype-consistency effects were found only on STI trials (i.e., only on trials in which concrete behavioral sentences were followed by trait probes), we conducted a second study in which a mere stereotype activation explanation was addressed explicitly.

Study 2

In Study 2 we aimed to check for stereotype activation effects more directly by having participants react to stereotype-neutral sentences (e.g., "The professor [garbage man] cycles through the street") that were followed by trait probes that might be activated on the basis of the category labels presented in these sentences (e.g., *smart* [*stupid*]). When the effects of Study 1 are based solely on stereotype activation, stereotype-neutral sentences should lead to trait activation effects that are just as strong as the ones based on stereotype-consistent sentences. However, we expect that stereotypical trait activation based on category labels will not lead to the same probe–response interference process as do the STIs based on stereotype-consistent behaviors.

Although a category label may activate stereotype-consistent traits, such traits are less likely to be associated with the content of the neutral behavior presented in the stimulus sentence, because the traits involved do not play a role in the interpretation of the behavior (i.e., in sentence comprehension). Although in stereotype-consistent behaviors, the primed stereotype may facilitate abstract sentence comprehension, the activated stereotype bears no relation to the behavior described in neutral stimulus sentences. Hence, in the latter case the potential confusion concerning whether the trait was in the sentence is less likely to occur. Therefore, we predict that participants will respond more slowly to relevant trait probes when the behavior in the preceding sentence is stereotype consistent with the category label presented than when the behavior is stereotype neutral with respect to the category label presented. Moreover, as was the case in Study 1, we predict that participants will respond more slowly to relevant trait probes when the behavior in the preceding sentence is stereotype consistent with the category label presented than when this behavior is stereotype inconsistent.

Method

Participants

A total of 45 Dutch undergraduate students (13 men and 32 women) from the University of Nijmegen participated in this study ($M = 21.84$ years old). Participants were recruited on campus and received 5 Fl (approximately $2.50) for their participation. All participants were native Dutch speakers.

Design

The critical trials consisted of stereotype-consistent, stereotype-inconsistent, and stereotype-neutral verb sentences that were followed by a trait probe. The experimental design thus consisted of one within-subject factor with three levels (stereotype consistency: consistent vs. neutral vs. inconsistent).

Procedure

The experiment was presented and set up in exactly the same way as described in Study 1. This time, however, participants went through 144 trials with a probe word and 120 trials without a probe word. Trials were divided once more into six trial groups, each consisting of 24 trials with and 20 trials without a probe word. The presentation of these six trial groups was the same as in Study 1.

Stimuli

Experimental trials. In addition to the stimuli used for the experimental trials in Study 1, new stimuli were used in Study 2. In Study 1, the actors in the stereotype-consistent and stereotype-inconsistent sentences were always the same. For instance, the skinhead was always part of a stereotype-consistent sentence, whereas the girl was always part of the stereotype-inconsistent equivalent. For the current study, additional verb sentences were created for which the same actors were used; however, this time the stereotype consistency was reversed. Thus, in these additional sentences, the skinhead engaged in stereotype-inconsistent behavior, whereas the girl performed stereotype-consistent behavior. Moreover, to test to what extent category labels are responsible for the activation of relevant traits, we added stereotype-neutral verb sentences (see Table 1).

In total, this resulted in 36 behavioral descriptions of 12 actors. In a pilot study ($N = 19$), participants were asked to indicate on a 7-point scale to what extent each of these 36 sentences was *stereotype consistent* (3) or *stereotype inconsistent* (−3). Results indicated that stereotype-consistent sentences ($M = 2.14$, $SD = 0.45$) were significantly more stereotype consistent than stereotype-neutral sentences ($M = 0.30$, $SD = 0.68$),

$t(18) = 9.59$, $p < .001$; and stereotype-inconsistent sentences ($M = -2.25$, $SD = 0.44$), $t(18) = 27.57$, $p < .001$. Also, stereotype-inconsistent sentences were significantly more stereotype inconsistent than were stereotype-neutral sentences, $t(18) = 13.30$, $p < .001$. In the same pilot study, participants were asked to indicate on a 7-point scale (1 = *not applicable*, 7 = *very applicable*) to what extent each of the 12 traits applied to each of the 12 behaviors. For all of the selected behavior–trait combinations, applicability scores were greater than 6. Moreover, in a small pilot study ($N = 8$) in which participants were asked to name 3 traits that come to mind after reading each sentence, selected traits were named by all participants.

In Study 2, all experimental trials were followed by trait probes that consisted of traits implied by the behaviors described in the stereotype-consistent and stereotype-inconsistent sentences (e.g., *aggressive* or *sweet*). Note, however, that trait probes following stereotype-neutral sentences consisted of the same traits as were used for the stereotype-consistent sentences for each target. Thus, the neutral sentence "The skinhead walks through town" was followed by the trait probe *aggressive*, whereas the neutral sentence "The girl walks through town" was followed by the trait probe *sweet*. We did this to test to what extent category labels may elicit stereotype-consistent trait activation.

Filler trials. As expected, in Study 1, differences due to stereotype consistency were found only for verb sentences followed by trait probes. Moreover, as noted before, to test our hypothesis, a simple comparison between the RTs on trait probes following stereotype-consistent, -inconsistent, and -neutral verb sentences suffices. Therefore, probe type and sentence type were no longer systematically manipulated in Study 2 (nor in Studies 3, 4, and 5). However, as was the case in Study 1, the same 36 sentences as the experimental ones were also presented with verb probes to prevent participants from indicating only "no" answers; 48 trait versions of the sentences were added to make sure each type of probe could elicit a "yes" as well as a "no" response (24 stereotype-consistent and 24 stereotype-inconsistent sentences, half of them with verb probes and half of them with trait probes); the 12 neutral sentences were presented twice, with the two different actors serving as probes, to make sure participants attended to the actor of each sentence. Finally, in accordance with Study 1, 120 filler sentences were added that were not followed by a probe word, to prevent participants from predicting when a probe word would appear.

Results

Incorrect answers on the task were infrequent ($M = 2.41\%$, range = 0.00–13.88%) and were excluded from the RT analyses. Analyses were again performed using RT cut-offs of 200 ms and 2,000 ms. Twenty-eight responses (0.49%) had to be dropped from the statistical analysis because of the cut-off criteria.

The RTs remaining after the cut-off were analyzed in a 3 (stereotype consistency: consistent vs. neutral vs. inconsistent) within-subject ANOVA. As predicted, a significant main effect was found, $F(2, 43) = 17.50$, $p < .001$ (see Table 3). Specific comparisons indicated that, as was the case in Study 1, participants were slower after stereotype-consistent than stereotype-inconsistent sentences, $F(1, 44) = 7.77$, $p < .01$. On average, it took participants 28 ms longer to indicate that the trait probe was not a part of the verb sentence when the sentence was stereotype consistent than when the sentence was stereotype inconsistent. Moreover, not only did it take participants less time to react to stereotype-neutral sentences than to stereotype-consistent sentences, $F(1, 44) = 35.18$, $p < .001$; it also took them significantly less time to react to stereotype-neutral sentences than to stereotype-inconsistent sentences, $F(1, 44) = 7.63$, $p < .01$. Thus, category labels in combination with neutral behaviors elicited less interference than did category labels in combination with either stereotype-consistent or stereotype-inconsistent behaviors.

For the error rates, no effects were found. This does not seem surprising, considering the simplicity of the task and the low overall error rate.

Table 3
Mean Response Latencies in Milliseconds as a Function of Stereotype Consistency: Study 2

	Stereotype consistency		
Statistic	Consistent	Neutral	Inconsistent
M	862_a	803_c	834_b
SD	183	162	174

Note. $N = 45$. Cell means that do not share the same subscripts differ significantly from each other ($p < .01$).

Discussion

As expected and in accordance with Study 1, participants made stronger STIs when presented with stereotype-consistent category–behavior combinations than when presented with stereotype-inconsistent category–behavior combinations. Moreover, the results of Study 2 indicate that this effect was not solely due to stereotype activation. Responses on trait probes following stereotype-consistent category–behavior sentences (e.g., "The professor wins the science quiz" followed by the trait *smart*) were significantly slower than responses on trait probes following stereotype-neutral category–behavior combinations (e.g., "The professor cycles through the street" followed by the trait *smart*). If stereotype activation was responsible for the differences in RTs found in Study 1 and Study 2, no difference between stereotype-consistent and stereotype-neutral sentences would have been obtained. Apparently, the current results cannot be explained by an account solely based on stereotype activation.

In both Study 1 and Study 2, participants' STIs were moderated by stereotype consistency. However, it is impossible to conclude on the basis of these studies whether participants made stronger STIs for stereotype-consistent information or weaker STIs for stereotype-inconsistent information, because no baseline measure was included. Although in Study 2 neutral sentences were added to check for the alternative stereotype activation account, in these sentences the behavior presented was neutral with respect to the category labels used. To investigate whether stereotypes have a facilitating influence on STIs (i.e., stronger STIs for stereotype-consistent information), an impeding influence on STIs (i.e., weaker STIs for stereotype-inconsistent information), or both, we need to include a stereotype-neutral STI baseline condition in which the category presented is neutral with respect to the behaviors used. This was the goal of our third study.

Study 3

In Study 3 we presented participants with behavioral sentences in which the actor was simply described with the letter *X* (e.g., "X wins the science quiz"). However, just prior to the sentence, a subliminal prime was presented containing a category label that was either stereotype consistent, stereotype inconsistent, or stereotype neutral with respect to the behavior presented in the sentence. The subliminal presentation of category labels has the advantage that it is easy to use a neutral category label such as *human* without the sentences becoming a bit odd (e.g., "The human wins the science quiz"). Moreover, subliminal category presentation makes the purpose of the research more obscure to participants because no extreme stereotypical or counterstereotypical information is presented explicitly.

In the case of stereotype-consistent category–behavior combinations, we reasoned that the increased accessibility of stereotype-consistent traits (e.g., Devine, 1989) and decreased accessibility of stereotype-inconsistent traits (e.g., Dijksterhuis & van Knippenberg, 1996) would not interfere with the STI process evoked by the behavior. It might even foster it. However, when the behavior is inconsistent with the activated stereotype, the increased accessibility of behavior-inconsistent traits and decreased accessibility of behavior-consistent traits are likely to interfere with the STI process evoked by the behavior. The STI process, therefore, is likely to become obstructed in the case of the prior activation of a stereotype that is inconsistent with the behavior or, more specifically, with the trait implied by the behavior. On the basis of this reasoning, we expected stronger STIs and, thus, slower responses to relevant trait probes when the behavior in the preceding sentence was stereotype consistent or stereotype neutral with the category label presented than when this behavior was stereotype inconsistent. With respect to the difference between stereotype-consistent and stereotype-neutral category–behavior combinations, we had no clear expectations. In both cases, the relevant trait concepts will get activated and thus will lead to slower RTs on the trait recognition probes. On the one hand, the prior activation of the relevant trait may foster the STI process and thus slow down the responses to trait probes even further in the case of stereotype-consistent combinations. On the other hand, it could be argued that once a trait is activated, it may be difficult to activate it even more, and thus responses to trait probes will be as slow for stereotype-consistent combinations as for stereotype-neutral combinations.

Method

Participants

A total of 56 Dutch undergraduate students (14 men and 42 women) from the University of Nijmegen participated in this study ($M = 21.30$ years old). Participants were recruited on campus and received 5 Fl (approximately $2.50) for their participation. All participants were native Dutch speakers.

Design

The critical trials that participants reacted to consisted of stereotype-consistent, stereotype-inconsistent, and stereotype-neutral category–sentence combinations. The experimental design thus consisted of one within-subject factor with three levels (category label: consistent vs. neutral vs. inconsistent).

Procedure

The experiment was presented in the same way as described in Study 1. Similar to Studies 1 and 2, participants were asked to read short sentences that appeared in the center of the screen. This time, however, each sentence described the behavior of an unknown actor who was indicated with an *X* (e.g., "X hits the saleswoman"). Each sentence stayed on the screen for 1,000 ms and was preceded by an *XXXXX* mask of 1,000 ms. Between the mask and the sentence, a category label was presented subliminally for two computer ticks (approximately 33 ms). After a blank screen of 500 ms, a probe word appeared in the center of the screen, preceded by an *XXXXX* mask of 500 ms. The participants' task was simply to indicate as quickly and accurately as possible whether the probe word was present in the preceding sentence.

Participants indicated their answers by pressing the *6* key on the numeric pad of the keyboard with their right index finger or the *A* key with their left index finger. During the instructions, participants were asked whether they preferred to indicate a "yes" answer with their right hand or with their left hand. In line with their preference, the *6* key and the *A* key functioned either as the "yes" key or the "no" key. Participants were instructed to position their hands over these keys throughout the experiment. After each response there was a blank screen of 1,000 ms before the next mask and sentence appeared on the screen.

In total, participants went through 72 trials, which were presented in random order. Before the actual experiment began, participants completed a practice round consisting of 15 sentences and probes. No subliminal primes were presented during this practice round.

Stimuli

Experimental trials. The experimental stimuli used in Study 3 consisted of a selection from the stimuli used in Studies 1 and 2 (see Table 1). In total, six different behaviors were used. Note that in this study, the actor in each sentence was represented by an *X*. Before each sentence, the category label was presented subliminally. Moreover, next to the behavior-consistent and behavior-inconsistent category labels that are presented in Table 1, a neutral category label was added. For all sentences, this category label was the same—namely, *human*. In total, this resulted in 18 experimental trials in which a verb sentence was followed by a trait probe.

Filler trials. As was the case in Study 1 and Study 2, we added verb probes and trait sentences to prevent participants from indicating only "no" answers. That is, the same 18 sentences as the experimental ones were also presented with verb probes. Moreover, we added 36 trait versions of the sentences to make sure each type of probe could elicit a "yes" as well as a "no" response (18 stereotype-consistent and 18 stereotype-inconsistent sentences, half of them with verb probes and half of them with trait probes). In the current experiment, no nonprobe filler sentences were added. Thus, each sentence was followed by a probe word.

Results

In accordance with the earlier studies, incorrect answers on the task were infrequent ($M = 0.21\%$, range $= 0.00–16.67\%$) and were excluded from the RT analyses. Analyses were again performed with RT cut-offs of 200 ms and 2,000 ms. Only one response (0.10%) had to be dropped from the statistical analysis because of the cut-off criteria.

The remaining RTs were analyzed in a 3 (category label: consistent vs. inconsistent vs. neutral) within-subject ANOVA. As predicted, the ANOVA revealed a significant main effect for stereotype category, $F(2, 54) = 5.23$, $p < .01$ (see Table 4). Specific comparisons indicated that participants took more time to indicate that the trait probe was not a part of the verb sentence when the category label was stereotype consistent with the behavior in the sentence, rather than when an inconsistent category was activated, $F(1, 55) = 6.75$, $p < .05$. When the neutral category label was activated, participants also took more time than they did when an inconsistent category was activated, $F(1, 55) = 7.07$, $p <$

Table 4
Mean Response Latencies in Milliseconds as a Function of Category Label: Study 3

Statistic	Category label: Consistent	Inconsistent	Neutral
M	618_a	595_b	622_a
SD	94	91	109

Note. $N = 56$. Cell means that do not share the same subscripts differ significantly from each other ($p < .05$).

.01. Moreover, no significant difference was found in participants' RTs when consistent or neutral category labels were presented, $F(1, 55) = 0.12$, *ns*. Apparently, trait activation in these two conditions was similar.

For the error rates, no effects were found.

Discussion

In line with our general hypothesis, participants again made stronger STIs on the basis of stereotype-consistent category–behavior combinations than on the basis of stereotype-inconsistent category–behavior combinations. Moreover, the results of Study 3 indicate that this effect is not due so much to the facilitation of STIs in the case of stereotype-consistent information as to the blocking of STIs in the case of stereotype-inconsistent information. That is, although we found slower RTs for stereotype-neutral category–behavior combinations than for stereotype-inconsistent combinations, we found no difference in RTs for stereotype-neutral and stereotype-consistent category–behavior combinations. In other words, it seems that it is time to rephrase our general finding: Weaker STIs are made for stereotype-inconsistent information than for stereotype-consistent information and for neutral information.

With this, the current findings support our assumption that the increased accessibility of behavior-inconsistent traits and decreased accessibility of behavior-consistent traits that results from the prior activation of category primes that are inconsistent with the behavior interferes with the STI process that is evoked by the behavior in a sentence. As a result, the STI process based on this behavior becomes obstructed. If such an encoding-interference process is indeed underlying the current findings, activation of stereotype-inconsistent information after the presentation of a behavior should not interfere with the STI process that is evoked spontaneously on the basis of the behavior. After all, in that case, the relevant trait inference on the basis of the behavior has already been made at the moment the (inconsistent) category label is presented.

However, there is an alternative account for our finding. It could be argued that responses to consistent and neutral category–sentence combinations are relatively slow not because the individual spontaneously inferred the trait while reading the sentence but because of a relation between the trait and the sentence that became apparent at the time the trait probe was presented. The recognition probe paradigm we used in the current studies is not conclusive in this respect because it does not discriminate between inferences that occur while the individual is reading the behavioral sentence and inferences that occur during the testing of the trait probe (for a similar argument, see Keenan et al., 1990). If backward integrative processes rather than differences in encoding are underlying the current results, activation of stereotype-inconsistent information after presentation of the behavior may still interfere with the reconstruction process that starts when participants are presented with the trait probe in the recognition probe paradigm. Thus, by presenting category labels after the behaviors, we may obtain conclusive evidence regarding the encoding versus backward integration issue in STI research (see Keenan et al., 1990; Uleman, Newman, & Moskowitz, 1996; Wyer & Srull, 1989). To test further our biased encoding assumptions, we performed two experiments in which category labels were presented subliminally (Study 4) or supraliminally (Study 5) before or after the presentation of the behaviors.

Study 4

This study to a large extent is based on Study 3. Again, participants were presented with behavioral sentences in which the actor is simply described with the letter *X* (e.g., "X wins the science quiz"). Like in Study 3, some sentences were preceded by a subliminal prime containing a category label that was either stereotype consistent or stereotype inconsistent with the behavior in the sentences. Additionally, however, some sentences were not preceded but succeeded by the subliminal category primes. These primes did, however, precede the presentation of the trait probes. Finally, some sentences were neither preceded nor succeeded by a category prime. These sentences make up the STI baseline condition. In accordance with Study 3, we expected weaker STIs and thus faster responses to relevant trait probes for stereotype-inconsistent category–behavior combinations than for stereotype-consistent combinations and the no-category baseline when the category prime preceded the behavior presentation. Furthermore, on the basis of our biased encoding account, we expected no differences between stereotype-consistent and stereotype-inconsistent behavior–category combinations when the category presentation succeeded the behavior presentation. After all, in these cases an STI has already occurred on the basis of the behavior in the sentence.

Method

Participants

A total of 63 Dutch undergraduate students (23 men and 40 women) from the University of Nijmegen participated in this study ($M = 21.92$ years old). Participants were recruited on campus and received 5 Fl (approximately $2.50) for their participation. All participants were native Dutch speakers.

Design

The critical trials consisted of stereotype-consistent and stereotype-inconsistent category–sentence combinations that were either preceded or succeeded by a subliminally presented category prime. The experiment thus consisted of a 2 (category label: consistent vs. inconsistent) × 2 (prime condition: prime before vs. prime after) within-subject design. The same set of stimuli also was presented without a category prime. This no-prime condition functioned as an STI baseline.

Procedure

The experiment was presented and set up almost the same way as described in Study 3. This time, however, participants went through 102 trials, which were presented in random order, with a short pause when participants were halfway through to prevent errors due to fatigue. Similar to Study 3, each sentence described the behavior of an unknown actor who was indicated with an *X* (e.g., "X hits the saleswoman"). All behavioral sentences were presented five times: once with a consistent and once with an inconsistent category label that was presented subliminally for two computer ticks (approximately 33 ms) immediately before the stimulus sentence was presented; once with a consistent and once with an inconsistent category label that was presented subliminally immediately before the probe was presented and thus after the stimulus sentence; and once with no category label presented. Before the actual experiment began, participants completed a practice round consisting of 15 sentences and probes. No subliminal primes were presented during this practice round.

Stimuli

Experimental trials. The behavioral sentences that were used for the experimental trials were the same as the ones used for Study 3 (see Table 1). Note that in this study, too, the actor in each sentence was represented by an "X." In total, 30 experimental trials were created. That is, each of the six behavioral sentences was presented in each of the five ways described above.

Filler trials. Similar to the other studies, verb probes and trait sentences were added to prevent participants from indicating only "no" answers. That is, the six stereotype-consistent and six stereotype-inconsistent sentences mentioned above were also presented with verb probes to prevent participants from indicating only "no" answers. Moreover, we added 24 trait versions of the sentences to make sure each type of probe could elicit a "yes" as well as a "no" response (12 stereotype-consistent and 12 stereotype-inconsistent sentences, half of them with verb probes and half of them with trait probes). In total, this resulted in 36 filler trials that were presented two times (with the category label presented subliminally once before the sentence and once before the probe). In the current experiment, no nonprobe filler sentences were added. Thus, each sentence was followed by a probe word.

Results

Three participants were removed from the analyses described in the following because of error rates of more than 30% of the experimental trials. Closer inspection of the data showed that these participants kept one response key pushed down for a long period of time during the experiment.

In accordance with the earlier studies, incorrect answers on the task were infrequent ($M = 0.28\%$, range = 0.00–10.00%) and were excluded from the RT analyses. Analyses were again performed with RT cut-offs of 200 ms and 2,000 ms. Only 11 responses (0.61%) had to be dropped from the statistical analyses because of the cut-off criteria.

The remaining RTs were analyzed in a 2 (category label: consistent vs. inconsistent) × 2 (prime condition: prime before vs. prime after) within-subject ANOVA. On the basis of our hypothesis, we expected longer RTs for stereotype-consistent than for stereotype-inconsistent category–sentence combinations, but only when a category label was presented before a stimulus sentence. A significant two-way interaction between category label and prime condition supported these predictions, $F(1, 59) = 9.08$, $p < .005$ (see Table 5).

Table 5
Mean Response Latencies in Milliseconds as a Function of Category Label and Prime Condition: Study 4

Prime condition	Category label: Consistent	Category label: Inconsistent
Prime before		
M	700_{a}	649_{c}
SD	155	117
Change with baseline	17	−33*
Prime after		
M	$659_{b,c}$	$675_{a,b}$
SD	140	138
Change with baseline	−23	−7

Note. $N = 60$. Values in rows and columns that do not share the same subscripts differ significantly from each other ($p < .05$). Values that differ significantly from the no-prime baseline ($M = 683$, $SD = 146$) are indicated with an asterisk.
* $p < .05$.

Analyses of the simple main effects for category label revealed that when category labels were activated immediately before the sentences, as expected, participants were slower in responding to stereotype-consistent category–sentence combinations than to stereotype-inconsistent combinations, $F(1, 59) = 11.30$, $p < .001$. On average, it took participants 51 ms longer to indicate that the trait probe was not included in the sentence when a behavior-consistent category was activated before the sentence than when a behavior-inconsistent category was activated. It is important to note that the difference between the stereotype-consistent category–sentence combinations with the no-prime baseline (M = 683 ms, SD = 146 ms) was not significant, $t(59) = 1.06$, *ns*, whereas the difference between the baseline and the stereotype-inconsistent combinations was significant, $t(59) = 2.63$, $p < .05$. Again, participants showed blocking of a trait inference when the category label was inconsistent.

When category labels were activated after the sentences, no significant difference between consistent and inconsistent category labels was found, $F(1, 59) = 1.27$, *ns*. Also, differences between the no category baseline condition and the consistent and inconsistent category conditions did not reach significance, $t(59) = 1.65$, *ns*; and $t(59) = 0.50$, *ns*, respectively. In both conditions in which the category was presented after the sentence, participants made as strong a trait inference as in the baseline condition. No other effects were found.

For the error rates, no effects were found.

Discussion

With respect to the presentation of a category label prior to a trait-implying behavioral sentence, the results of Study 4 nicely replicate the results of our third study. Again, participants made weaker STIs on the basis of stereotype-inconsistent category–behavior combinations than on the basis of stereotype-consistent category–behavior combinations and when no category was presented. Similar to Study 3, the current results thus indicate that the difference in STIs between stereotype-consistent and stereotype-inconsistent category–behavior combinations is not due to the facilitation of STIs in the case of stereotype-consistent information but rather to the blocking of STIs in the case of stereotype-inconsistent information.

In line with our biased encoding account, we found no difference between stereotype-consistent and stereotype-inconsistent behavior–category combinations when the category presentation succeeded the behavior presentation. It is interesting that when the category was presented after the behavioral sentence, neither stereotyping condition differed significantly from the no-category baseline. Apparently, participants spontaneously engaged in default STI processing on the basis of the behavioral information when the behavior was presented first. The category label that subsequently was presented came too late to interfere with this process. After all, at that point, the spontaneously inferred trait already had been activated.

However, it should be noted that the subliminal and therefore implicit nature of the category primes in Study 4 in part may have been responsible for the obtained pattern of results. Subliminal presentation of the category labels has the advantage that it obscures the stereotypic nature of the task and tells us something about the uncontrollability of the effects of stereotypes on STIs. However, the disadvantage is that we could not instruct participants explicitly that these labels described the actors in the sentences. Therefore, there is the possibility that presenting these labels before the behavioral sentences led to a stronger link between the activated category and the actor in the sentence than presenting them afterward did. To overcome this alternative account, we performed a fifth experiment.

Study 5

Study 5 is similar to Study 4, only this time category labels were presented supraliminally before or after the behavioral sentences. Moreover, participants were instructed explicitly that these labels referred to the actor in each sentence. On the basis of our encoding explanation, we expected the same pattern of results as was predicted and found for Study 4.

Method

Participants

A total of 53 Dutch undergraduate students (34 men and 19 women) from the University of Amsterdam participated in this study ($M = 21.56$ years old). Participants were recruited on campus and received 5 Fl (approximately $2.50) for their participation. All participants were native Dutch speakers.

Design and Procedure

The experiment was designed, presented, and set up in the same way as described in Study 4. This time however, the category labels were presented supraliminally immediately before the sentences or the probes for 300 ms. These category labels were presented in red. We did this so that participants would make a distinction between these labels and the probe word they had to react to. Participants were explicitly told that these red words that were flashed very briefly on the screen right before or after each sentence indicated something about the actor (the *X*) in the sentence.

Results

Incorrect answers on the task were infrequent ($M = 2.70\%$, range $= 0.00–13.33\%$) and were excluded from the RT analyses. Analyses were again performed with RT cut-offs of 200 ms and 2,000 ms. Only 18 responses (1.13%) had to be dropped from the statistical analysis because of the cut-off criteria.

The remaining RTs were analyzed in a 2 (category label: consistent vs. inconsistent) × 2 (prime condition: prime before vs. prime after) within-subject ANOVA. On the basis of our hypothesis, we expected longer RTs for stereotype-consistent than for stereotype-inconsistent category–sentence combinations, but only when a category label was presented before a stimulus sentence. A significant two-way interaction between category label and prime condition supported these predictions, $F(1, 52) = 5.38$, $p < .05$ (see Table 6).

Analyses of the simple main effects for category label revealed that when category labels were activated immediately before the sentences, as expected, participants were slower in responding to stereotype-consistent category–sentence combinations than to stereotype-inconsistent combinations, $F(1, 52) = 8.51$, $p < .005$. On average, it took participants 46 ms longer to indicate that the trait probe was not included in the sentence when a behavior-consistent category was activated before the sentence than when a behavior-inconsistent category was activated. It is important to note that the difference among the stereotype-consistent category–sentence combinations with the no-prime baseline ($M = 698$ ms, $SD = 117$ ms) was not significant, $t(52) = 1.00$, *ns*, whereas the difference between the baseline and the stereotype-inconsistent combinations was significant, $t(52) = 2.33$, $p < .05$. Again, participants showed blocking of a trait inference when the category label was inconsistent.

When category labels were activated after the sentences, no significant difference between consistent and inconsistent category labels was found, $F(1, 52) = 0.63$, *ns*. Also, differences between the no-category-baseline condition and the consistent and inconsistent category conditions did not reach significance, $t(52) = 0.49$, *ns*; and $t(52) = 1.19$, *ns*, respectively. In both conditions in which the category was presented after the sentence, participants made as strong a trait inference as in the baseline condition. No other effects were found.

For the error rates, no effects were found.

Table 6
Mean Response Latencies in Milliseconds as a Function of Category Label and Prime Condition: Study 5

	Category label	
Prime condition	Consistent	Inconsistent
Prime before		
M	713_a	667_b
SD	154	125
Change with baseline	15	−31*
Prime after		
M	704_a	718_a
SD	151	155
Change with baseline	6	20

Note. $N = 53$. Values in rows and columns that do not share the same subscripts differ significantly from each other ($p < .05$). Values that differ significantly from the no-prime baseline ($M = 698$, $SD = 117$) are indicated with an asterisk.
* $p < .05$.

Discussion

The results of Study 5 perfectly replicate the results of Study 4. When the category label preceded the behavior, again, participants made weaker STIs on the basis of stereotype-inconsistent category–behavior combinations than on the basis of stereotype-consistent category–behavior combinations and when no category was presented. It is important to note, however, that no differences between consistent and inconsistent category–behavior combinations and the no-category baseline were found when the category label succeeded the behavior. Thus, whereas the activation of a stereotype-inconsistent category label before a behavior interfered with on-line trait encoding, the same stereotype-inconsistent category label could no longer prevent on-line trait encoding when it was presented after the fact (i.e., after the encoding of the original behavioral information). If, in this situation, there had been backward integrative processes starting off as soon as the trait probe appeared on the screen, the just-presented inconsistent category label would have prevented correspondent trait encoding just the way it did when the category label was presented preceding the behavior description. In other words, the current findings constitute a serious problem for a backward integrative account of the STI effects found and provide strong evidence that STIs indeed are made at the encoding of behavioral information.

Taken together, the results of Studies 4 and 5 support our assumption that the increased accessibility of behavior-inconsistent traits and decreased accessibility of behavior-consistent traits that result from the prior activation of a behavior-inconsistent category interferes with the STI process that is evoked by the behavior in a sentence. As a result, the STI process based on this behavior becomes blocked. However, when the category prime is presented after the behavioral information or when no category information is available, the STI encoding process is not influenced, and relevant traits are spontaneously activated while the individual is encoding the behavioral information.

The important question thus becomes which piece of information gets noticed first, the category a person belongs to or the behavior this person demonstrates. As noted in the introduction, on encountering an individual, perceivers automatically categorize this actor on the basis of salient features, such as gender, age, and skin color (e.g., Bodenhausen, Macrae, & Sherman, 1999; Devine, 1989; Fiske, 1998). The current findings demonstrate that such categorizations subsequently influence the STIs people make while encoding behavioral information. With this, the present results demonstrate the dominant role that categorical information plays in social information processing (see Bodenhausen et al., 1999; van Knippenberg & Dijksterhuis, 2000). Once activated, stereotypes simply seem to overrule the impact of individuating information. As noted by Kunda (1999), however, behavioral information may, in principle, also be observed and processed before an actor has been categorized (e.g., when one learns afterward that the person who saved one's life by calling an ambulance was a passing skinhead). The present results also seem to indicate that, in that case, individuating behavioral information may exert its own influence on subsequent spontaneous trait activation without immediate interference of the category information presented later on.

The important role of order in the presentation of stereotype-relevant information does not seem to be restricted to cognitive measures of encoding. In line with the present findings, Bodenhausen (1988) found that a Hispanic defendant of a crime was judged more guilty than was a Caucasian defendant only when the ethnicity of the defendant was known before the evidence was evaluated. When participants learned the ethnicity of the defendant after the evidence was evaluated, no difference was found. In summary, the present findings suggest that the order in which categorical and individuating information are presented (or perceived) can be of crucial importance with respect to the inferences a perceiver makes when encoding the information.

General Discussion

The research presented in this article focuses on the influence of stereotypes on the spontaneous inference of traits (e.g., Uleman, Newman, & Moskowitz, 1996). In five studies, it was found that participants made weaker relevant STIs on the basis of stereotype-inconsistent information than on the basis of stereotype-consistent information. Also, in Studies 3, 4, and 5, evidence was found that participants made weaker relevant STIs on the basis of stereotype-inconsistent information than on the basis of stereotype-neutral information. Apparently, the STI process was obstructed when an inconsistent category label was presented prior to a behavior. In the present studies, no significant difference was found between STIs based on stereotype-consistent information and those based on stereotype-neutral information. Thus, the STI process was not enhanced when a category label that was consistent with the behavior was presented prior to the behavior.

The finding that STI processes are impeded as a function of stereotype inconsistency is in line with other recent research and theory that point to the centrality of inhibitory mechanisms in cognition and impression formation (Anderson & Spellman, 1995; Kunda & Thagard, 1996). Especially in impression formation, inhibitory mechanisms may be very beneficial. In general, people prefer to maintain their existing beliefs (e.g., stereotypes) rather than change them. Although collecting evidence in favor of one's existing beliefs may be helpful in this respect, it may even be more beneficial to avoid obtaining counterevidence (see also Dijksterhuis & van Knippenberg, 1996; van Knippenberg & Dijksterhuis, 2000). The inhibition of trait activation on the basis of counterevidence such as stereotype-inconsistent behavioral information, therefore, may be a more useful mechanism than the facilitation of trait activation on the basis of stereotype-consistent behavioral information.

In addition to the functional difference between inhibitory and facilitatory processes in impression formation outlined above, the absence of facilitatory effects due to stereotyping in the current research may, in part, also be due to the stimulus material used. Similar to other research on STIs (e.g., Winter & Uleman, 1984), the behavioral sentences that were used in the present experiments were selected specifically on the basis of their trait-implying nature. This may have resulted in particularly strong STIs on the basis of the behaviors themselves. As a result, the consistent category labels may not have had much room to add to the trait activation. Future research in which more ambiguous stimulus material is used may also demonstrate facilitatory effects in STIs due to stereotyping. In line with this assumption, research based on ambiguous stimuli demonstrated that stereotypes influence the tacit inferences people make on the basis of ambiguous information in a stereotype-confirming way (Dunning & Sherman, 1997).

There is still some debate concerning whether the effects found in STI studies can indeed be attributed to traits that are spontaneously inferred during the encoding of behavioral information (e.g., Uleman, Newman, & Moskowitz, 1996) or whether these effects are due to a backward integrative process that takes place on the basis of the trait cue or trait probe that participants are presented with to test the STI (e.g., Keenan et al., 1990; Wyer & Srull, 1989). The recognition probe paradigm seems to provide convincing evidence in favor of the assumption that traits are spontaneously inferred at the encoding of behavioral information for two reasons. First, making trait inferences in the recognition probe paradigm actually deteriorates participants' performance on the task they have to perform. Therefore, it seems highly unlikely that participants engage in some sort of retrieval or reconstruction process when confronted with the trait probe, because that would be dysfunctional for the task at hand. Second, in the recognition probe paradigm, trait inferences are measured directly after encoding. Thus, there is not much time left to engage in reconstructive processes (e.g., McKoon & Ratcliff, 1986).

Still, the recognition probe paradigm is not conclusive with regard to the encoding issue because it does not discriminate between inferences that occur while the individual is reading and inferences that occur during testing (Keenan et al., 1990). The results of Study 4 and 5, in our view, convincingly demonstrate that STIs are made during encoding. Category labels presented immediately before the stimulus sentences did affect participants' responses to the trait probes; however, category labels presented after the stimulus sentences but before the trait probes did not affect participants' responses to these trait probes. If a postencoding process had been underlying the current results, activation of inconsistent category labels after presentation of the behaviors would have interfered with reconstruction processes occurring during the recognition probe paradigm. The fact that this was not the case, in our view, constitutes strong evidence in favor of the assumption that STIs are made during encoding.

Taken together, the current experiments clearly demonstrate that the STI process may be obstructed in the case of prior activation of an inconsistent stereotype. Although STIs may occur spontaneously, they are apparently not inevitable. This finding is in line with recent theorizing and research in the social judgment literature (for an overview, see Gilbert, 1998). In this literature, social judgment is generally thought of as a multiple-stage process (e.g., Gilbert, 1998; Trope, 1986) that initially involves two stages. First, perceivers automatically identify a behavior. Second, they engage in spontaneous attributional inferences on the basis of the identified behavior. These initial spontaneous attributions, in general, are thought to be of a dispositional nature (see Gilbert, Krull, & Pelham, 1988; Gilbert, Pelham, & Krull, 1988; Uleman, Newman, & Moskowitz, 1996). However, recent research shows that they may be situational as well (see Krull & Dill, 1996; Krull & Erickson, 1995). Whether a spontaneous dispositional or situational inference occurs depends on the context (e.g., Krull & Dill, 1996; Krull & Erickson, 1995).

With respect to the spontaneous inference stage, Trope and Gaunt (1999) showed that situational activation of factors such as salience, accessibility, and applicability affects the use of this information in the dispositional inferences that participants make. They demonstrated that when contextual information is salient, accessible, and applicable, it may produce strong discounting effects on dispositional inferences even when perceivers are under cognitive load (Trope & Gaunt, 1999). In line with these findings, Krull and Erickson (1995) showed that once behavior is identified, people may spontaneously make situational inferences as well as dispositional inferences, depending on their epistemic goals. On the basis of these findings, Gilbert (1998) concluded that "dispositional inferences may be easily launched and easily completed, but they are not inevitable" (p. 114). The current findings support this view. Although people seem to be inclined to make a spontaneous dispositional inference on the basis of behavioral information, this process may be inhibited as a function of the social context—for example, because of the category label of an actor.

As noted in the introduction, the majority of theorizing and research into impression formation has focused on intentional inferences or causal reasoning (for an overview, see Gilbert, 1998; Gilbert & Malone, 1995)—that is, trying to explain how people make sense of the behavior of others when explicitly asked. The present results clearly demonstrate that people also make inferences unintentionally on the basis of the behavior of others. It is interesting that the predictions one would make for stereotype-inconsistent behaviors on the basis of intentional inferences are the opposite of what was predicted in the current article for unintentional inferences. For example, suppose a nurse helps an elderly woman across the street. This behavior is not very distinctive for nurses—on the contrary, it is part of their job to help the elderly. Thus, at an explicit level, no correspondent inference should be made (e.g., Jones & Davis, 1965; Kelley & Michela, 1980). Now imagine for a moment that a skinhead helps the elderly woman cross the street. Now the behavior is very distinctive and, consequently, a correspondent inference should be made at an explicit level. The data on unintentional, spontaneous inferences presented in the current article show exactly the opposite pattern. Stronger STIs were made in the case of stereotype-consistent information than for stereotype-inconsistent information.

In accordance with the difference between intentional and unintentional inferences outlined above, Gilbert and colleagues (Gilbert, Krull, & Pelham, 1988; Gilbert, Pelham, & Krull, 1988) demonstrated that participants made other inferences when they were motivated or had the ability to engage in intentional processing than when they were not motivated or did not have the ability to engage in intentional processing. Future research in which intentional and unintentional inferences on the basis of stereotypical information are compared may shed more light on the differences between these two types of inferences and their role in stereotype maintenance (see Uleman, 1999).

The current studies clearly demonstrate that stereotypes influence the amount and nature of information an individual encodes (e.g., von Hippel et al., 1995). Although most researchers will agree with the assumption that stereotypes may lead to biased encoding of information, relatively little empirical attention had been paid to this notion (for exceptions, see von Hippel et al., 1995). The present research shows that stereotypical views have a significant impact on the way people encode information by impeding the STIs that people make on the basis of trait-implying behaviors. A person who correctly answers a Trivial Pursuit question by default is likely to be perceived as relatively smart; however, the prior knowledge that this person is a garbage man will inhibit the *smart* concept. At the cost of misjudging this particular garbage man's intelligence, in this way one may conveniently maintain one's stereotype of garbage men. Put more generally, people encode social information in a stereotype-maintaining way. Expected behaviors are spontaneously encoded in more abstract conceptions that fit one's existing ideas about the personality of a person. Unexpected behaviors, however, are spontaneously perceived and encoded as dissonant behaviors and do not get incorporated in the overall picture one has of others. That is why, unfortunately, stereotypical views are so resistant to change. They keep reinforcing themselves.

References

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review, 102,* 68–100.

Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology, 41,* 258–290.

Bargh, J. A., Lombardi, W. J., & Higgins, E. T. (1988). Automaticity of chronically accessible constructs in Person × Situation effects on person perception: It's just a matter of time. *Journal of Personality and Social Psychology, 55,* 599–605.

Bargh, J. A., & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside awareness on impression formation. *Journal of Personality and Social Psychology, 43,* 437–449.

Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology, 55,* 726–737.

Bodenhausen, G. V., Macrae, C. N., & Sherman, J. W. (1999). On the dialectics of discrimination: Dual processes in social stereotyping. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 271–290). New York: Guilford Press.

Carlston, D. E., & Skowronski, J. J. (1994). Savings in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology, 66,* 840–856.

Carlston, D. E., Skowronski, J. J., & Sparks, C. (1995). Savings in relearning II: On the formation of behavior-based trait associations and inferences. *Journal of Personality and Social Psychology, 69,* 420–436.

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

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

Dijksterhuis, A., & van Knippenberg, A. (1995). Timing of schema-activation and memory: Inhibited access to inconsistent information. *European Journal of Social Psychology, 25,* 383–390.

Dijksterhuis, A., & van Knippenberg, A. (1996). The knife that cuts both ways: Facilitated and inhibited access to traits as a result of stereotype activation. *Journal of Experimental Social Psychology, 32,* 271–288.

Dunning, D., & Sherman, D. A. (1997). Stereotypes and tacit inference. *Journal of Personality and Social Psychology, 73,* 459–471.

Fiske, S. T. (1998). Stereotypes, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 357–411). New York: McGraw-Hill.

Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 89–150). New York: McGraw-Hill.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509–517.

Gilbert, D. T., Krull, D. S., & Pelham, B. W. (1988). Of thoughts unspoken: Social inference and the self-regulation of behavior. *Journal of Personality and Social Psychology, 55,* 685–694.

Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin, 117,* 21–38.

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology, 54,* 733–740.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Higgins, E. T., Rholes, W. S., & Jones, C. R. (1977). Category accessibility and impression formation. *Journal of Experimental Social Psychology, 13,* 141–154.

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,* 69–84.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). New York: Academic Press.

Keenan, J. M., Potts, G. R., Golding, J. M., & Jennings, T. M. (1990). Which elaborative inferences are drawn during reading? A question of methodologies. In D. A. Balota, G. B. F. d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 377–402). Hillsdale, NJ: Erlbaum.

Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology, 31,* 457–501.

Krull, D. S., & Dill, J. C. (1996). On thinking first and responding fast: Flexibility in social inference processes. *Personality and Social Psychology Bulletin, 22,* 949–959.

Krull, D. S., & Erickson, D. J. (1995). On judging situations: The effortful process of taking dispositional information into account. *Social Cognition, 13,* 417–438.

Kunda, Z. (1999). Parallel processing of stereotypes and behaviors. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 314–322). New York: Guilford Press.

Kunda, Z., & Thagard, P. (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103,* 284–308.

Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72,* 275–287.

Maass, A., Milesi, A., Zabbini, S., & Stahlberg, D. (1995). Linguistic intergroup bias: Differential expectancies or in-group protection? *Journal of Personality and Social Psychology, 68,* 116–126.

Macrae, C. N., Bodenhausen, G. V., & Milne, A. B. (1995). The dissection of selection in person perception: Inhibitory processes in social stereotyping. *Journal of Personality and Social Psychology, 69,* 397–407.

Macrae, C. N., Stangor, C., & Milne, A. B. (1994). Activating social stereotypes: A functional analysis. *Journal of Experimental Social Psychology, 30,* 370–389.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 82–91.

Perdue, C. W., & Gurtman, M. B. (1990). Evidence for the automaticity of ageism. *Journal of Experimental Social Psychology, 26,* 199–216.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510–532.

Semin, G. R., & Fiedler, K. (1991). The linguistic category model, its bases, applications and range. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 2, pp. 1–30). Chichester, England: Wiley.

Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998).

Stereotype efficiency reconsidered: Encoding flexibility under cognitive load. *Journal of Personality and Social Psychology, 75,* 589–606.

Srull, T. K., & Wyer, R. S., Jr. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology, 37,* 1660–1672.

Stangor, C., & Lange, J. E. (1994). Mental representation of social groups: Advances in understanding stereotypes and stereotyping. *Advances in Experimental Social Psychology, 26,* 357–416.

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

Trope, Y., & Gaunt, R. (1999). A dual-process model of overconfident attributional inferences. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 161–178). New York: Guilford Press.

Uleman, J. S. (1999). Spontaneous versus intentional inferences in impression formation. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 141–160). New York: Guilford Press.

Uleman, J. S., Hon, A., Roman, R., & Moskowitz, G. B. (1996). On-line evidence for spontaneous trait inferences at encoding. *Personality and Social Psychology Bulletin, 22,* 377–394.

Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreters: Evidence and issues from spontaneous trait inference. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 28, pp. 211–279). San Diego, CA: Academic Press.

van Knippenberg, A., & Dijksterhuis, A. (2000). Social categorization and stereotyping: A functional perspective. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 11, pp. 105–144). Chichester, England: Wiley.

van Overwalle, F., Drenth, T., & Marsman, G. (1999). Spontaneous trait inferences: Are they linked to the actor or to the action? *Personality and Social Psychology Bulletin, 25,* 450–462.

von Hippel, W., Sekaquaptewa, D., & Vargas, P. (1995). On the role of encoding processes in stereotype maintenance. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 27, pp. 177–253). San Diego, CA: Academic Press.

Wigboldus, D. H. J., Semin, G. R., & Spears, R. (2000). How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology, 78,* 5–18.

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. Also see correction in *Journal of Personality and Social Psychology* (1986), *50,* 355.

Winter, L., Uleman, J. S., & Cunniff, C. (1985). How automatic are social judgments? *Journal of Personality and Social Psychology, 49,* 904–917. Also see correction in *Journal of Personality and Social Psychology* (1986), *50,* 381.

Wyer, R. S., & Srull, T. K. (1989). *Memory and cognition in its social context.* Hillsdale, NJ: Erlbaum.

Received July 10, 2000
Revision received September 25, 2002
Accepted October 1, 2002 ■