CAPACITY AND COMPREHENSION: SPONTANEOUS STEREOTYPING UNDER COGNITIVE LOAD

Daniël H. J. Wigboldus
University of Amsterdam

Jeffrey W. Sherman & Heather L. Franzese
Northwestern University

Ad van Knippenberg
University of Nijmegen

People tend to make spontaneous trait inferences (STIs) when confronted with the behavior of others. Recent research has demonstrated that these STIs may be moderated by contextual cues such as stereotypic category labels. The central aim of the current research was to investigate the role of cognitive resources in this process. Two experiments were performed in which STIs were measured using a probe recognition paradigm under low or high cognitive load. Under high cognitive load, STIs were less likely for stereotype–inconsistent than stereotype–consistent behaviors. Compared to baseline, STIs for inconsistent behaviors were less strong under high cognitive load. Under low-load, no differences in STIs as a function of stereotypes were found. These findings support the idea that stereotypes are especially likely to affect STIs when resources are low.

Spontaneous trait inferences (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). For instance, when you observe someone carrying an elderly woman's groceries across the

street, the trait "helpful" may be spontaneously activated (Winter, Uleman, & Cunniff, 1985). Convincing evidence for the occurrence of STIs has been obtained using several research paradigms (e.g., Carlston & Skowronski, 1994; Carlston, Skowronski, & Sparks, 1995; Newman, 1991, 1993; Uleman, Hon, Roman, & Moskowitz, 1996; Winter & Uleman, 1984; Winter et al., 1985).

As yet, most research into STIs has focused solely on behaviors, whereas the features of the actors have remained nondescript (e.g., John, Mary) or have been described by an irrelevant category label (for an exception, see Mae, Carlston, & Skowronski, 1999). In real life, however, we know considerably more about the actors whose behaviors we observe. For example, gender, age, and skin color are salient features that perceivers will use to rapidly categorize a person.[1] These social categories and the stereotypes associated with such categories may greatly influence interpretations of an actor's behavior. For example, the same ambiguous behavior is perceived as more aggressive when performed by a Black than a White target (Duncan, 1976; Sagar & Schofield, 1980; see also Bodenhausen & Wyer, 1985; Darley & Gross, 1983; Krueger & Rothbart, 1988; Kunda & Sherman–Williams, 1992). As such, it might be expected that stereotypes would influence the likelihood that trait inferences would be drawn spontaneously from observed behaviors.

In fact, recent research conducted by Wigboldus, Dijksterhuis, and van Knippenberg (2003; see also Howard & Uleman, 1997) demonstrated just such effects. Participants were presented with trait-implying sentences that described behaviors that were either stereotype–consistent or stereotype–inconsistent with a category label that was activated subliminally immediately preceding each sentence. For example, the sentence "X wins the science quiz" was preceded by the subliminally presented category label "professor" or "garbage man." Across five experiments, results showed that STIs were more likely to occur for consistent than for inconsistent behaviors. Moreover, STIs based on stereotype–inconsistent behaviors were significantly weaker than STIs made from those same behaviors when the stereotype was not primed. No differences were found between STIs made from stereotype–consistent behaviors and the neutral (no prime) baseline. These results suggest that a "default" STI process was obstructed in the case of stereotype–inconsistent category–behavior combinations. Stereotype–consistent category labels, however, did not seem to enhance the "default" STI process.

Wigboldus et al. (2003) propose that the process underlying these findings is based on accessibility effects. It is assumed that, by default, traits

---

1. Although some researchers included information about these target attributes in their stimuli (e.g., Carlston & Skowronski, 1994; Carlston et al., 1995), the attributes were not systematically varied or analyzed.

get spontaneously activated on the basis of trait-implying behaviors. Similarly, the activation of a stereotype will make stereotype–consistent traits temporarily more accessible and simultaneously will make stereotype–inconsistent traits temporarily less accessible (e.g., Dijksterhuis & van Knippenberg, 1996). In the case of stereotype–consistent category–behavior combinations, there is no competition between both trait activations. However, in the case of stereotype–inconsistent category–behavior combinations, there will be competition between the different activations or trait accessibilities. In two studies in which the order of presentation of the category information (stereotype activation) and the behavioral information (trait-implying sentences) were varied, Wigboldus et al. (2003, Exp. 4 & 5) found that STI generation was interfered with only when a stereotype–inconsistent category label preceded the behavioral information. When an inconsistent category label was presented after the trait-implying behavior, no such effect was found. These findings strongly support the idea that stereotype labels activate traits that are inconsistent with the trait implications of a behavior, and this competing activation interferes with the generation of STIs at encoding.

The central aim of the current research was to examine the extent to which these stereotype–based inference processes are moderated by the availability of processing capacity. Research into the efficiency of STIs has shown that, in general, the STI process is neither disrupted nor enhanced by concurrent load (Lupfer, Clark, & Hutcherson, 1990; Winter et al., 1985; but also see Skowronski, Carlston, Mae & Crawford, 1998). This suggests that the STI process is quite efficient (although some interference with the STI process has been found at very high-load levels; Uleman, Newman, & Winter, 1992).

However, there are reasons to expect that the effects of stereotypes on STIs may be influenced by the availability of cognitive resources. Stereotypes are commonly characterized as efficient social perception tools that provide a relatively large amount of social information and require very little effort to activate and apply. Consistent with this characterization, stereotypes have been found to be especially likely to guide social perception when processing capacity is low (e.g., for reviews, see Fiske, 1998; Hamilton & Sherman, 1994; Hilton & von Hippel, 1996; Sherman, Macrae, & Bodenhausen, 2000). With respect to intentional inferences, Sherman, Lee, Bessenoff, and Frost (1998; see also Sherman & Frost, 2000) demonstrated that when resources were low, perceivers were particularly likely to extract the conceptual gist (trait meanings) from consistent but not inconsistent behaviors in an intentional impression formation task. In contrast, when cognitive resources were high, both types of behaviors were understood equally well. We would like to argue that spontaneous and unintentional inference processes also would be more likely to be influenced by stereotypes when resources are de-

pleted. That is, it is especially likely that stereotypes will be activated when cognitive resources are depleted. Therefore, in these resource–depleted conditions, the competition between traits activated on the basis of the stereotype and traits activated on the basis of the behavior will be greatest and the resources available to resolve the competition will be minimal. Although Wigboldus et al. (2003) found no evidence that spontaneous inferences of stereotype–consistent behaviors were facilitated by stereotype labels, we hypothesized that such effects might be more likely when cognitive resources are low and comprehension is more difficult. Two experiments were conducted to test these hypotheses.

EXPERIMENT 1

OVERVIEW

Both experiments relied on a probe recognition task developed by McKoon and Ratcliff (1986) to examine spontaneous inference processes, and modified by Newman (1991, 1993) to measure STIs. This is also the task that was used in the previously described studies by Wigboldus et al. (2003) that found initial evidence that STIs are moderated by stereotypes. In this task, participants are presented with trait-implying sentences, such as "He stepped on his girlfriend's feet during the foxtrot." After reading the sentence, participants are presented with the target trait ("clumsy") and are asked to indicate as quickly and as accurately as possible whether or not this probe word was included in the preceding sentence. To the extent that the trait is spontaneously inferred upon reading the sentence, it is more difficult to correctly report that the trait was not actually included in the sentence. Activation of a trait at encoding is thus assumed to interfere with the subsequent trait probe task. In line with this assumption, trait–implying sentences result in longer response times to trait probes than to control sentences (Uleman, Hon, et al., 1996).

In Experiment 1, behavioral descriptions that implied traits that were either consistent or inconsistent with the stereotype of Asian Americans were presented. The name of the actor in each behavior was either a typically Caucasian or a typically Asian name. Responses to trials with Caucasian names provide a baseline for responses to consistent (typically Asian) and inconsistent (typically non–Asian) behaviors. Half of the participants were put under a cognitive load as they read the behaviors and performed the probe word recognition task. On the basis of past results (Wigboldus et al., 2003), we expected that, even in the low-load condition, stereotypes might influence the likelihood of inferring traits from consistent and inconsistent behaviors. If so, then participants would require more time to correctly reject traits implied by consistent than by inconsistent behaviors when the behaviors were paired with

Asian names. No such differences were expected for behaviors paired with Caucasian names.

On the basis of Wigboldus et al. (2003), specific predictions regarding the absolute facilitation (in the case of stereotype–consistent behaviors) and inhibition (in the case of stereotype–inconsistent behaviors) can also be made. In line with Wigboldus et al., we expected that in the case of stereotype–inconsistent category–behavior combinations, stereotypes obstruct the spontaneous inferences that would be made solely from the behaviors. Thus, participants would require less time to correctly reject traits implied by Asian–inconsistent behaviors when those behaviors were paired with Asian names than in the baseline condition.

Thus, in the stereotype–inconsistent behavior condition, the STI process will be obstructed due to the competition between traits activated on the basis of the stereotypical category label and the trait-implying behavior. No such competition will take place in the case of stereotype–consistent category–behavior combinations. Therefore, no difference was expected in the time that it took to correctly reject traits implied by Asian–consistent behaviors when those behaviors were paired with Asian names compared to the baseline condition, when they were paired with Caucasian names, particularly in the low-load condition.

Most importantly, however, in all cases, we expected the influence of stereotypes on STIs to be greater in the high-load condition than in the low-load condition. When attentional capacity is depleted, the effects of stereotypes on the comprehension of stereotype–relevant material should be maximized, perhaps even leading to facilitatory effects on the inferences made from stereotype–consistent behaviors paired with stereotypic names.

METHOD

Participants. One hundred and sixteen Northwestern University students enrolled in an Introduction to Psychology class participated for partial course credit. Participants were run in sessions of 1–5 people.

Materials. Based on pretesting, 12 sentences were selected that implied traits relevant to the stereotype of Asian Americans. Of these sentences, six implied traits stereotypic of Asian Americans (two each implied shy, smart, and hardworking), and six implied traits counterstereotypic of Asian Americans (two each implied aggressive, conceited, and dumb; Rothbart & John, 1993). In addition to these 12 sentences, six sentences that implied but did not include traits irrelevant to the stereotype of Asian Americans were also presented to prevent participants from guessing the purpose of the experiment. Eighteen filler sentences were also developed in which the implied and probed trait word was actually

present in the stimulus sentence. These sentences were all irrelevant to the Asian American stereotype.

Based on pilot testing, names that were highly indicative of Caucasian (e.g., Richie Fairfox) and Asian (e.g., Henry Wong) ethnicities were selected to be paired with the sentences. The Caucasian and Asian names were each paired with six key sentences (three consistent, three inconsistent). Filler sentences were paired with a mixture of Caucasian, Asian, Latino, and African American names. The six key behaviors that were paired with Caucasian and Asian names were reversed in a between-subjects factor that provided an internal replication.

Procedure. Participants were told that they were taking part in a memory experiment. Because interest centered on spontaneously occurring trait inferences, nothing was mentioned about inferring personality traits or forming impressions of others. Participants were told that they would be seeing a series of sentences on their computers, and that after each one they would be asked to respond to a word that would appear on the screen. They were instructed that, upon the presentation of this word, they were to press the button on their keyboard marked "yes" if that exact word had been in the sentence they had just read, and to press the button marked "no" if the word had not been in the sentence. They were instructed to respond as quickly and as accurately as possible. To familiarize participants with the procedure, prior to responding to the behavioral items, participants were given practice items having nothing to do with personality traits or stereotypes. During the experiment, participants paced themselves through the behavior–trait pairs, which were presented randomly. The reaction time for each item was recorded and served as the dependent measure.

As they performed this task, half of the participants were placed in a high cognitive load condition. These participants were further informed that the experiment was concerned with people's ability to perform multiple tasks at the same time. A cognitive load was manipulated by asking these participants to hold an eight–digit number in memory as they performed the experimental task. This method has been used frequently and successfully in past research to deprive participants of processing resources (e.g., Gilbert & Hixon, 1991; Sherman & Frost, 2000; Sherman et al., 1998).[2]

RESULTS

Incorrect "yes" responses to the key items were removed from the data set. There were only four such responses. Reaction times greater than 2

---

2. Analyses that both included and excluded four participants who incorrectly reported four or more of the digits produced the same results. The reported analyses include these participants.

standard deviations above the mean were replaced with the mean for the particular item in question. This resulted in the replacement of approximately 2.5% of the responses. Reaction times for relevant items were averaged to form indices of stereotype–consistent and –inconsistent responses for each participant. All means are reported in milliseconds.

A 2 (cognitive load: high vs. low) × 2 (internal replication) × 2 (target ethnicity: Caucasian vs. Asian) × 2 (stimulus type: Asian stereotype–consistent vs. –inconsistent) ANOVA with repeated measures on the last two factors was conducted on the reaction times. This analysis produced a marginally reliable three–way interaction between target ethnicity, stimulus type, and load, $F(1, 112) = 2.77$, $p < .10$. As planned, separate analyses were conducted on responses in the high- and low-load conditions. In the low-load condition, a 2 (replication) × 2 (target ethnicity: Caucasian vs. Asian) × 2 (stimulus type: Asian stereotype–consistent vs. –inconsistent) ANOVA, with repeated measures on the last two factors, produced no reliable results, all Fs < 1 (see Figure 1, top panel). In the high-load condition, the same analysis produced a reliable interaction between target ethnicity and stimulus type, $F (1, 58) = 4.21$, $p < .05$ (see Figure 1, bottom panel). Further analyses demonstrated that when the behaviors were paired with Asian names, more time was required to correctly reject Asian–consistent ($M = 1,051$) than –inconsistent ($M = 1,003$) traits, $F(1, 59) = 3.68$, $p = .06$. The time required to correctly reject Asian–consistent ($M = 1,016$) and –inconsistent ($M = 1,045$) traits did not differ in the baseline condition in which the same behaviors were paired with Caucasian names, $F(1, 59) = 1.61$, $p = .21$.

The comparison between reaction times for Asian–consistent behaviors paired with Asian and Caucasian baseline names was not reliable, $F(1, 59) = 1.46$, $p = .23$. Thus, there was no evidence for absolute facilitation effects. However, the comparison between reaction times for Asian–inconsistent behaviors paired with Asian and Caucasian names was marginally reliable, $F(1, 59) = 3.05$, $p < .09$, and the means indicated that STIs for stereotype–inconsistent behaviors were obstructed in the high-load condition.[3]

---

3. In this experiment, comparisons of responses to consistent and inconsistent items may be of limited value, given that the items include both different behaviors and different probes. Thus, for example, correct responses to inconsistent items may require more time than correct responses to consistent items because the inconsistent behavior–trait pairs imply their trait more strongly than do the consistent behavior–trait pairs. Such differences may be magnified by a cognitive load, which interferes with comprehension. In fact, in the Caucasian baseline in the high-load condition, responses to inconsistent items did require somewhat more time than did responses to consistent items. This is important to keep in mind in interpreting the absolute facilitation and inhibition results.

Low Cognitive Load
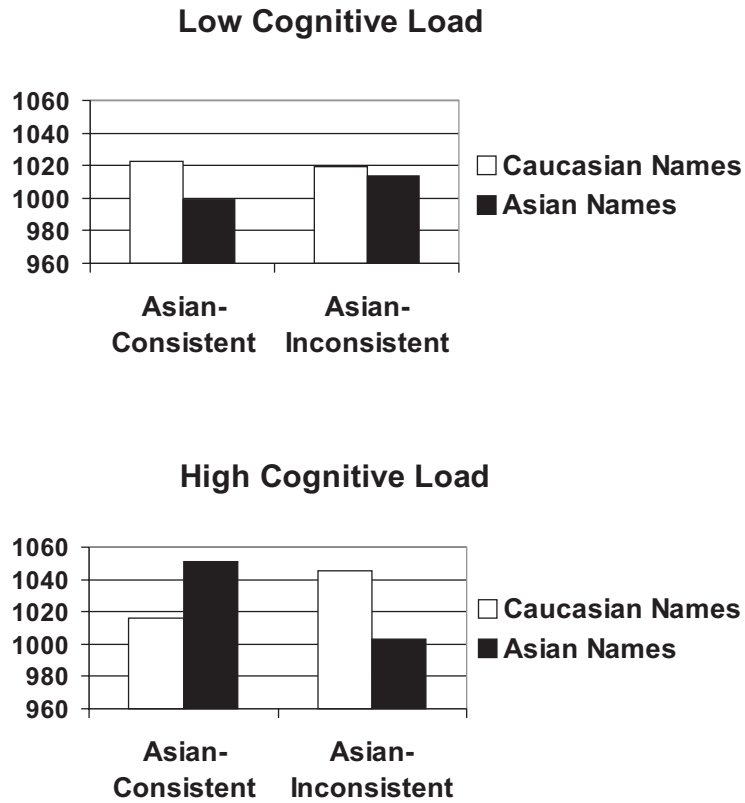

High Cognitive Load

FIGURE 1. Mean reaction times in milliseconds to trait probes as a function of target ethnicity and stereotype consistency under low cognitive load (top panel) and high cognitive load (bottom panel) in Study 1.

DISCUSSION

The results from this experiment provide initial evidence that stereotypes are particularly likely to influence STIs when processing capacity is depleted. In the high-load condition, participants required more time to correctly reject Asian–consistent than Asian–inconsistent traits when stereotypically Asian behaviors were paired with Asian names. No such effect was observed when participants enjoyed full processing capacity

or when behaviors were paired with Caucasian names (regardless of processing capacity). Indeed, just the opposite tendency was observed for Caucasian names in the high-load condition. These data suggest that when capacity is low, stereotype–inconsistent traits are less likely to be spontaneously inferred than stereotype–consistent traits. In addition, in the high-load condition, Asian–inconsistent traits were rejected more quickly when behaviors were paired with Asian names than when they were paired with Caucasian names, providing evidence of absolute inhibition of inconsistent STIs. However, although Asian–consistent traits were rejected somewhat more slowly when behaviors in the high-load condition were paired with Asian rather than Caucasian names, this difference did not approach significance. These results are consistent with the findings of Wigboldus et al. (2003), who also found evidence for absolute inhibition of inconsistent traits but not absolute facilitation of consistent traits.

The data from Experiment 1 provide initial support for our hypotheses, but they are not conclusive. Although the critical two–way interaction was significant in the low-capacity condition and not in the high-capacity condition, the overarching three–way interaction was only marginally significant. Moreover, the use of Caucasian names as controls might not have been ideal, because these names might themselves have invoked a stereotype. Therefore, replication and extension of the Experiment 1 findings seemed to be in order, and a second experiment was performed.

EXPERIMENT 2

The aims of Experiment 2 were to replicate the findings from Experiment 1 with a different methodology and a wider variety of stereotypes. In Experiment 2, we used a modification of the probe recognition task developed by McKoon and Ratcliff (1986). In order to be able to use multiple stereotypes, but at the same time keep the total number of trials as small as possible, we presented multiple probes after each behavioral sentence. That is, each sentence was followed by five different probes. In this way, many fewer filler sentences were necessary. We wanted to keep the total number of trials as low as possible to prevent the task itself from becoming too demanding (which would more or less eliminate our capacity manipulation).

Another modification in Experiment 2 was that stereotypes were activated by means of a category label (e.g., garbage man) presented subliminally prior to each sentence. This subliminal presentation obscured the purpose of the experiment and minimized social desirability concerns. In addition, it allowed us to create control trials in which, instead of a stereotypical category label, the word "human" was presented sublimi-

nally before a sentence (for a similar method, see Wigboldus et al., 2003). This avoids the potential pitfalls of presenting Caucasian names, which may have activated Caucasian stereotypes, on the control trials of Experiment 1.

METHOD

Participants. Sixty–eight University of Nijmegen students participated in this study. Participants were recruited on campus and received Fl. 5 (approximately $ 2.5) for their participation. All participants were native Dutch speakers.

Materials. Based on pretesting, four pairs of trait-implying sentences were selected, each sentence describing one specific behavior of an actor. Sentences were selected such that for each pair of sentences, one sentence implied a trait stereotypic of a specific social category, whereas the other sentence implied a trait counterstereotypic of this category. The reverse was the case for another social category. For instance, the sentence "X hits the saleswoman," implying the trait aggressive, may be considered stereotype–consistent for a skinhead but stereotype–inconsistent for a girl, whereas the sentence "X comforts the little boy," implying the trait sweet, may be considered stereotype–consistent for a girl but stereotype–inconsistent for a skinhead. The other three pairs of trait-implying sentences (and associated social categories) described smart versus stupid behavior (professor vs. garbage man), helpful versus antisocial behavior (boy scout vs. punk), and industrious versus lazy behavior (manager vs. Surinamese).

Procedure. Participants were told that they were participating in a dual tasking experiment; nothing was mentioned about inferring personality traits or forming impressions of others. It was explained that they would be presented with a series of sentences on their computers, and that after each sentence they would be asked to respond to five probe words that would appear on the screen. They were instructed to respond as quickly and as accurately as possible upon the presentation of these probe words. For each probe word, they were asked to press the green button on their keyboard marked "yes" if that exact word had appeared in the sentence they had just read and to press the red button marked "no" if that word had not been in the sentence. A brief five–sentence practice session using sentences unrelated to the purpose of the experiment was used to familiarize participants with this task.

Next, the second task of this dual-tasking experiment was described to participants. Immediately preceding each sentence, a randomly chosen one–digit (low-load condition) or five–digit number (high-load condition) was presented for 5 seconds. Participants were instructed to memorize this number. At the end of each trial, participants were asked to type the number provided at the start of the trial. Feedback about the ac-

curacy of this response was provided immediately after the response.[4] Participants were familiarized with this procedure via another short practice round presenting the same five sentences used in the first practice round.

After the practice trials, participants were presented with 24 experimental trials. These trials were presented in random order and consisted of the eight stereotype–relevant sentences described above. Each was presented three times: once preceded by a masked, subliminally presented (33 ms) stereotype–consistent category label, once preceded by a subliminally presented stereotype–inconsistent category label, and once preceded by a subliminally presented stereotype–neutral label (human). Responses to the sentence preceded by the neutral label provide a baseline that can be used to evaluate the effects of the stereotype–consistent and stereotype–inconsistent labels on recognition performance.[5]

Each sentence stayed on the screen for 1,000 ms and was followed by five probe words that were presented in random order. The critical probe word was the trait implied by the behavior in each sentence. The other four probe words served as fillers to prevent participants from having to respond "no" to each probe word. One filler probe consisted of the verb presented in the sentence ("yes" response), one consisted of the object presented in the sentence ("yes" response), one consisted of a trait that was not implied by the sentence ("no" response), and, finally, one consisted of an article that was presented in some cases but not in others (sometimes "yes," sometimes "no" response). For instance, the sentence "X hits the saleswoman" was followed by the probe words "aggressive," "hits," "saleswoman," "sweet," and "the." Reaction times to the probe words were recorded by the computer.

---

4. In the low-load condition (one–number digit), on average, participants reported the wrong number on 3.1% of the trials, whereas in the high-load condition (five–number digit) the wrong number was reported on 23.7% of the trials, $F(1, 66) = 57.30$, $p < .001$. This difference confirms our load manipulation. Moreover, the rather high percentage of wrong numbers justifies the use of a five–digit number instead of an eight–digit number in the high-load condition (cf. Gilbert & Hixon, 1991). Analyses both including and excluding these responses produced the same results. The reported results included response times from these trials.

5. After the experiment, participants were asked whether they had noticed any flashes during the presentation of the sentences. Sixteen participants indicated that they had noticed some flashes. Subsequently, these participants were asked whether they had recognized these flashes as words. Two participants indicated that they had recognized words. When asked what words were presented, however, neither of them mentioned any of the category labels that were used.
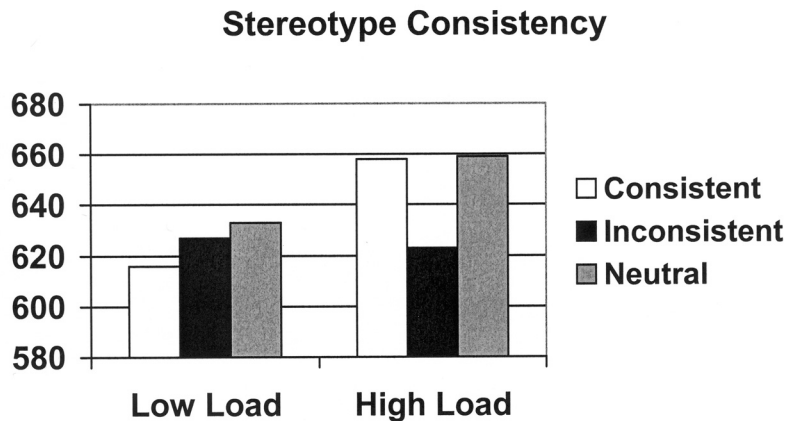
Stereotype Consistency

FIGURE 2. Mean reaction times in milliseconds to trait probes as a function of stereotype consistency and cognitive load in Study 2.

RESULTS

Incorrect "yes" responses to key trait probe items were removed from the data set. There were only nine such responses. Reaction times greater than 2 standard deviations above the mean were replaced with the mean for the particular item in question. This resulted in the replacement of approximately 4.5% of the responses. Reaction times for relevant items were averaged to form indices of stereotype–consistent, stereotype–inconsistent, and stereotype–neutral baseline responses for each participant. All means are reported in milliseconds.

A 2 (cognitive load: high vs. low) × 3 (stereotype consistency: consistent vs. inconsistent vs. neutral) mixed model ANOVA, with repeated measures on the last factor was conducted on the reaction times. This analysis yielded a main effect for stereotype consistency, $F(2, 65) = 3.77$, $p < .05$, which was qualified by the expected interaction between cognitive load and stereotype consistency, $F(2, 65) = 5.75$, $p < .005$ (see Figure 2). Analyses of simple main effects demonstrated that, in the low-load condition, participants required the same amount of time to correctly reject traits implied by stereotype–consistent, stereotype–inconsistent, and stereotype–neutral sentences, $F < 1$. However, in the high-load condition, reaction times differed as a function of stereotype consistency, F

(2, 65) = 7.90, $p < .001$. Specific comparisons indicated that participants required more time to correctly reject traits when the sentences were preceded by stereotype–consistent ($M = 658$) and stereotype–neutral primes ($M = 659$) than when they were preceded by stereotype–inconsistent primes ($M = 623$), t (33) = 3.34, $p < .005$ and t (33) = 2.98, $p < .005$, respectively.

DISCUSSION

In line with our expectations and the results of Experiment 1, when capacity was depleted, participants made weaker trait inferences for stereotype–inconsistent category–behavior combinations than for stereotype–consistent combinations. Further replicating Experiment 1, trait inferences from inconsistent category–behavior combinations were inhibited compared to baseline in the high-load condition. Moreover, as in Experiment 1, no absolute facilitation was found for consistent items relative to baseline in any condition. In addition, no differences in spontaneous trait inferences were found as a function of stereotype consistency in the low-load condition. Thus, as in Experiment 1, these results provide evidence for the inhibition of counterstereotypic STIs under cognitive load, but not for the facilitation of stereotypic STIs.

GENERAL DISCUSSION

In two experiments and across multiple social categories, it was shown that the influence of stereotypes on STIs is moderated by the availability of processing capacity. When capacity was depleted, weaker STIs were found for stereotype–inconsistent than stereotype–consistent behaviors. In addition, STIs for inconsistent behaviors were inhibited compared to a neutral baseline in these conditions. None of these effects emerged when participants enjoyed full processing capacity (Experiment 1) or were under a very low-load (Experiment 2).

Two findings here deserve further comment. First, the results from both experiments showed evidence for the obstruction of STIs based on stereotype–inconsistent behaviors, but no faciliatory effects for STIs based on stereotype–consistent behaviors. Notably, Wigboldus et al. (2003) similarly found evidence for inhibition but not facilitation. It is tempting to conclude from these results that stereotypes influence social comprehension primarily by interfering with the encoding of inconsistent behaviors, and not by facilitating the encoding of consistent behaviors. However, we believe that at this moment a conclusion based only on competition of different trait activations could be premature. The

particular stimuli chosen for these studies make it very difficult to observe facilitation effects. This is, because the key behavior/trait pairs that were selected for use were selected specifically because the behaviors strongly imply the traits. As such, there may be ceiling effects in terms of the spontaneity with which the traits are inferred from the behaviors, leaving room only for inhibition. Studies that employ more ambiguous stimuli may be more likely to demonstrate facilitation effects. Given the facilitatory effects of stereotypes on all sorts of other dependent measures (as described in the introduction), it seems likely that the probability of spontaneously inferring traits from behaviors would also be facilitated for ambiguous behaviors.

The second finding that deserves comment is our failure to find any evidence of stereotypic influences on STIs in the low-load conditions. Wigboldus et al. (2003) demonstrated such effects. However, although those experiments relied on the same trait probe technique as the current ones, there are some important differences between the studies. One important difference is that participants in the Wigboldus et al. experiments were required to perform a much larger number of experimental and filler trials than in both current experiments. As such, it is possible that fatigue was a factor in the studies, and that participants were, in fact, capacity depleted. This could mean that the results presented by Wigboldus et al. (2003) are restricted to low-capacity conditions. However, other differences in the studies having to do with the particular behaviors, traits, names, and social categories employed (Experiment 1) and the multiple probe method used in Experiment 2 may also have contributed to the different findings. For instance, in Experiment 1, names were used as a basis for categorization. STI effects with names maybe harder to get because they require a processing step that goes beyond that necessary when stereotype labels are directly presented. That is, the names must first lead to category inferences before those inferences can affect STIs. This could be one reason why in Experiment 1 no effect was found in the high-capacity condition. With respect to Experiment 2, it could be argued that in a recognition probe paradigm, the strongest effect will be found when each sentence is followed by one probe. With multiple probes, the effect of the inferences made on the basis of the sentence may slowly fade out. Therefore, STI effects may be harder to obtain in a multiple probe paradigm. This could explain why in Experiment 2 no effects were found in the high-capacity condition. In any case, we do not wish to claim that stereotypes will necessarily only influence STIs when capacity is low. To the contrary, it seems likely that such effects may occur in many contexts. We claim only that stereotypes are more likely to influence STIs when capacity is low than when it is high.

ON THE INEVITABILITY AND APPLICATION OF SPONTANEOUS TRAIT INFERENCES

One of the central issues in research on STIs concerns their inevitability. Recent research suggests that they are very difficult to prevent. STIs have been observed with both comprehension and memory instruction sets (e.g., Carlston et al., 1995). Likewise, STIs were demonstrated in the present research, despite the absence of any explicit instruction to form target impressions. STIs also have been shown to be relatively impervious to fluctuations in processing capacity (e.g., Lupfer et al., 1990; Winter et al., 1985). However, robust as they may be, the present results demonstrate that STIs are not inevitable. Wigboldus et al. (2003) demonstrated that stereotypes may inhibit STIs for counterstereotypic behaviors. Our results show that this may be particularly likely under cognitive load.

Another central issue in the STI literature concerns the nature of the inference that is made. In particular, participants may infer only the trait meaning of the behavior, may implicitly associate the trait with the actor, or may infer that the actor actually possesses the trait. Although considerable energy and resources have been devoted to the resolution of this issue, consensus remains elusive (e.g., Brown & Bassili, 2002; Carlston et al., 1995; Todorov & Uleman, 2002). Certainly, we are in no position to settle the issue based on our data. However, we would note that the effects demonstrated in our studies have important implications, regardless of the extent to which participants' inferences are directly applied to the actors. Stereotypes may be changed by stereotype–inconsistent behavior only when the behavior is identified as such. Although identification of inconsistent behavior certainly does not guarantee stereotype change, the absence of such identification precludes the possibility. Even if the trait implied by the behavior is not directly applied to the actor, the enhanced accessibility of identified but not applied counterstereotypic traits may yet have important influences on social cognition and behavior. For example, the threshold for perceiving future counterstereotypic behavior by the individual or other group members may be lowered. In addition, the accessibility of the counterstereotypic trait may diminish the likelihood that the perceiver will act toward the target in ways that produce self–fulfilling hypotheses (e.g., Chen & Bargh, 1997).

CONCLUSION

These results add to a large body of research on the manner in which stereotypes influence social comprehension processes. Past research has shown that stereotypes can affect the meanings ascribed to consistent

and inconsistent behaviors (Bodenhausen & Wyer, 1985; Darley & Gross, 1983; Duncan, 1976; Krueger & Rothbart, 1988; Kunda & Sherman–Williams, 1992; Sagar & Schofield, 1980). Prior research has also shown that traits are more likely to be inferred spontaneously from stereotype–consistent than stereotype–inconsistent behaviors (e.g., Wigboldus et al., 2003). Finally, when explicitly attempting to form an impression of another person, perceivers under a cognitive load have been shown to be more likely to extract the gist meaning of consistent rather than inconsistent behaviors (Sherman et al., 1998). The present results extend these findings by demonstrating that stereotypes are especially likely to affect the extent to which traits are spontaneously inferred from behaviors when processing capacity is depleted.

Stereotypes have long been heralded as efficient social perception tools that allow perceivers to infer a great deal about other persons with relatively little effort (e.g., Fiske, 1998; Hamilton & Sherman, 1994; Hilton & von Hippel, 1998; Sherman et al., 2000). The present results demonstrate just how pervasive is the use of these tools. Even if perceivers have no intention to extract impressions from others' behavior, stereotypes may be recruited and applied spontaneously, particularly if cognitive resources are low. As is the case with other stereotypic influences on social inference processes, spontaneously applied stereotypes bias comprehension in a self–perpetuating manner. Given that perceivers may be unlikely to even be aware that they are using stereotypes in these contexts, efforts to decontaminate social perception from the influence of stereotypes may be very difficult indeed.

REFERENCES

Bodenhausen, G. V., & Wyer, R. S., Jr. (1985). Effects of stereotypes on decision making and information processing strategies. *Journal of Personality and Social Psychology*, *48*, 267–282.

Brown, R. D., & Bassili, J. N. (2002). Spontaneous trait associations and the case of the superstitious banana. *Journal of Experimental Social Psychology, 38,* 87–92.

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.

Chen, M., & Bargh, J. A. (1997). Nonconscious behavioral confirmation processes: The self–fulfilling consequences of automatic stereotype activation. *Journal of Experimental Social Psychology, 33,* 541–560.

Darley, J. M., & Gross, P. H. (1983). A hypothesis–confirming bias in labeling effects. *Journal of Personality and Social Psychology, 44*, 20–33.

Dijksterhuis, A., & van Knippenberg, A (1996). The knife that cuts both ways: Facili-

tated and inhibited access to traits as a result of stereotype activation. *Journal of Experimental Social Psychology, 32,* 271–288.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology, 34,* 590–598.

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert & S. T. Fiske, (Eds.), *The handbook of social psychology*, (4th ed., Vol. 2, pp. 357–411). 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.

Hamilton, D. L., & Sherman, J. W. (1994). Stereotypes. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd Ed., Vol. 2, pp. 1–68). Hillsdale, NJ: Erlbaum.

Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology, 47,* 237–271.

Howard, P. B., & Uleman, J. S. (1997, May). *Prejudiced trait inferences: The automatic effects of an actor's race.* Paper presented at the 9th annual convention of the American Psychological Society, Washington, DC.

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

Kunda, Z., & Sherman–Williams, B. (1992). Stereotypes and the construal of individuating information. *Personality and Social Psychology Bulletin, 19,* 90–99.

Lupfer, M. B., Clark, L. F., & Hutcherson, H. W. (1990). Impact of context on spontaneous trait and situational attributions. *Journal of Personality and Social Psychology, 58,* 239–249.

Mae, L., Carlston, D. E., & Skowronski, J. J. (1999). Spontaneous trait transference to familiar communications: Is a little knowledge a dangerous thing? *Journal of Personality and Social Psychology*, *77*, 233–246.

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

Newman, L. S. (1991). Why are traits inferred spontaneously? A developmental approach. *Social Cognition, 9,* 221–253.

Newman, L. S. (1993). How individualists interpret behavior: Idiocentrism and spontaneous trait inference. *Social Cognition, 11,* 243–269.

Rothbart, M. & John, O. P. (1993). Intergroup relations and stereotype change: A social–cognitive analysis and some longitudinal findings. In P. M. Sniderman, P. E. Tetlock, & E. G. Carmines (Eds.), *Prejudice, politics, and the American dilemma* (pp. 32–59). Stanford, CA: Stanford University Press.

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

Sherman, J. W., & Frost, L. A. (2000). On the encoding of stereotype–relevant information under cognitive load. *Personality and Social Psychology Bulletin, 26,* 26–34.

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.

Sherman, J. W., Macrae, C. N., & Bodenhausen, G. V. (2000). Attention and stereotyping: Cognitive constraints on the construction of meaningful social impressions. *European Review of Social Psychology, 11,* 145–175.

Skowronski, J. J., Carlston, D. E., Mae, L., & Crawford, M. T. (1998). Spontaneous trait transference: Communicators take on the qualities they describe in others. *Journal of Personality and Social Psychology, 74,* 837-848.

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology, 83,* 1051–1065.

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

Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreters: Evidence and issues from spontaneous trait inference. *Advances in Experimental Social Psychology, 28,* 211–279.

Uleman, J. S., Newman, L. S., & Winter, L. (1992). Can personality traits be inferred automatically? Spontaneous inferences require cognitive capacity at encoding. *Consciousness and Cognition, 1,* 77–90.

Wigboldus, D. H. J., Dijksterhuis, A., & van Knippenberg, A. (2003). When stereotypes get in the way: Stereotypes obstruct stereotype–inconsistent trait inferences. *Journal of Personality and Social Psychology, 84,* 470–484

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.