Social Cognition, Vol. 27, No. 2, 2009, pp. 222–248

TWO FACES ARE BETTER THAN ONE: ELIMINATING FALSE TRAIT ASSOCIATIONS WITH FACES

Amir Goren and Alexander Todorov
Princeton University

Traits spontaneously inferred from behaviors bind to actors' faces (spontaneous trait inferences, or STI), and can inappropriately bind to faces of persons deemed irrelevant to the behaviors (spontaneous trait transference, or STT). Experiments 1 and 2 showed that STT effects, while weaker than STI, nevertheless occurred in the face of clear irrelevance instructions. Experiment 3 showed that STT—but not STI—can be eliminated when faces and behaviors are separated perceptually. Experiments 4, 5, and 6 found no STT when a relevant and an irrelevant actor were presented simultaneously, suggesting that inferred traits bind exclusively to faces of relevant targets when these targets are present and bind to irrelevant targets only when relevant targets are absent and irrelevant targets are present. The experiments rule out an attentional account of this phenomenon, demonstrating STT elimination when attention to concomitant face-behavior pairs was ensured. The findings suggest a spontaneous attributional process.

Imagine that you run into Jill, and during the course of conversation, she mentions that a friend of hers, Jackie, never gives tips at restaurants. Perhaps surprisingly, you now get a sense that *Jill* is stingy, having inferred this from her description of Jackie and having misattributed it to Jill instead. This is precisely what studies on spontaneous trait transference (STT) show (Carlston, Skowronski, & Sparks, 1995; Skowronski, Carlston, Mae, & Crawford, 1998). Traits spontaneously inferred from behaviors can become associated with the wrong face. Although this association is inappropriate, if people have the goal of forming impressions, they would use Jill's description of Jackie to form an evaluative impression of Jill (Wyer, Budesheim, Lambert, & Swan, 1994). What happens when the face is clearly irrelevant to the person description? That is, what happens if Jill does not know Jackie and just happened to be around when Jackie was described as stingy?

We wish to thank Manish Pakrashi and Valerie Loehr for helping with data collection. This research was supported by National Science Foundation Grant BCS-0446846 to Alexander Todorov, Assistant Professor of Psychology and Public Affairs, Department of Psychology, Princeton University, Princeton, NJ 08540; E-mail: atodorov@princeton.edu.

Correspondence concerning this article should be addressed to Amir Goren, now at the Department of Psychology, Yale University, 309 Edwards Street, New Haven, CT 06520-8369. E-mail: amir.goren@yale.edu.

The first objective of this article is to provide convergent evidence for the generality of the STT phenomenon. In the first three experiments, we tested whether STT occurs even when the behavioral information is clearly irrelevant to the person shown, and we tested whether it is possible to eliminate the false association between trait inferences and faces.

Now, imagine the same "Jackie/Jill" scenario, except that this time, Jackie is standing right next to Jill as she describes Jackie's stingy behavior. Under the circumstances, you would infer that Jackie is stingy, as would be expected by research on spontaneous trait inferences (STI; Carlston & Skowronski, 1994; Carlston, Skowronski, & Sparks, 1995; Todorov & Uleman, 2002). But would you also misattribute this personality trait to Jill, as predicted by STT research?

The second objective of this article is to delineate the boundary conditions under which STT are formed. In the third experiment, we investigated the influence of perceptual separation on STT, and in the last three experiments, we tested whether false associations between trait inferences and faces are formed when the face of the relevant actor is present.

Researchers have long been interested in the attributions that people make for others' behaviors (e.g., Jones & Davis, 1965; Kelley, 1967; for a review, see Gilbert, 1998), and the focus has gradually shifted toward implicit and automatic processes by which people form impressions (Uleman, Blader, & Todorov, 2005). Research on STI has demonstrated that people engage in STI from behaviors even when they do not intend to do so (Uleman, Newman, & Moskowitz, 1996). For example, from a behavioral description, "Henry kicked the dog out of his way," one is likely to infer the trait "cruel," even though this trait was only implied by the behavior. Moreover, subsequent research has shown that such inferences are bound to the actor's representation (e.g., "Henry is a cruel person.") even when participants do not recall the original behavior (Carlston & Skowronski, 1994; Carlston et al., 1995; Todorov & Uleman, 2002) and are depleted in terms of cognitive resources (Todorov & Uleman, 2003).

Whereas these studies describe highly efficient cognitive processes for person perception, they also suggest potential attribution errors. Correspondent trait inferences, in which people attribute personality traits to others on the basis of their behaviors despite the presence of behavior-constraining situational factors (Jones & Harris, 1967; Gilbert, 1989, 1998; Gilbert, Pelham, & Krull, 1988; Ross, Amabile, & Steinmetz, 1977; Trope, 1986), are an example of one such error. Another error is the misattribution of STI to the wrong person. Carlston et al. (1995) provided such evidence by showing that STI were bound to the faces of communicators describing the trait-implying behaviors of an unseen, opposite-sex target (the actor). Skowronski et al. (1998; see also Carlston & Skowronski, 2005; Mae, Carlston, & Skowronski, 1999), building on the previous work, developed the notion of STT, whereby trait inferences are spontaneously formed about and linked to irrelevant actors. For example, from Claudia's claim, "Henry kicked the dog out of his way," one will get the sense that *Claudia* is a cruel person.

In the above studies, STT effects have been obtained after participants were presented with faces paired with paragraphs describing trait-implying behaviors. In the current studies, we used a false recognition paradigm with short presentations of single trait-implying behaviors, and examined the conditions and processes by which STT effects are eliminated. Both Skowronski et al. (1998) and Todorov and Uleman (2004) have argued that STT can be described as shallow associative processes different from person attribution processes. Interestingly, before the savings and false recognition paradigms were introduced, this was one of the interpretations of STI (Bassili, 1989; Wyer & Srull, 1989, pp. 57-58). However, both of these paradigms, using faces as critical stimuli, have provided convincing evidence that STI are more properly characterized in terms of attributional processes (Carlston & Skowronski, 2005; Todorov & Uleman, 2003).

To the extent that STT are driven by shallow associative processes, it should be difficult to eliminate STT effects. In fact, STT effects have been obtained even when participants are warned about the STT effect (Carlston & Skowronski, 2005) and when the communicators are well-known (Mae et al., 1999). However, if the formation of inappropriate person-trait associations is conditional and depends on the absence of an appropriate target (an intriguing new hypothesis put forth by Crawford, Skowronski, & Stiff, 2007), then STT effects may easily be prevented. Todorov and Uleman (2004) have argued that "a spontaneous rule-based inference mechanism that operates automatically" (p. 492) can explain why the effects for STI are stronger than the effects for STT. This hypothesis that trait inferences are automatically associated with the relevant face (e.g., Todorov & Uleman, 2003) implies that these inferences are *not* associated with irrelevant (concomitantly present) faces. However, Todorov and Uleman's studies did not include the proper control conditions, as explained below, to establish whether STT were truly eliminated. The only effect that these studies showed was that associations with a relevant face were stronger than associations with an irrelevant face.

The process by which traits are bound to faces is still not entirely well specified. The hypothesis we are putting forth in this article is that in trait-person binding, exclusive, nontransferring trait-target associations are created except when relevant, appropriate targets are absent. In other words, traits may become associated with any concomitant object (if any is available) when a relevant target is missing (in accord with a simple associational account), but traits will bind exclusively to the relevant target when that target is present (suggesting an important limitation on free association with respect to faces and traits). A type of spontaneous selection may therefore be the mechanism involved in trait-person binding that, depending on the situation, is activated in either an exclusive or an indiscriminate manner.

In all six experiments, we used the false recognition paradigm (Todorov & Uleman, 2002, 2003, 2004). As in Todorov and Uleman's studies, photos of faces are presented along with individual, trait-implying behavioral sentences. Participants view many of these person-behavior pairs during the presentation phase. Subsequently, they are presented with a series of face-trait pairs for which they have to indicate whether the trait has previously appeared in the behavioral sentence paired with the particular person shown. Todorov and Uleman's studies have shown that participants are more likely to falsely recognize implied traits in the context of the correspondent actors' faces than in the context of other actors' faces (control traits). High relative rates of false recognition indicate spontaneous inference of traits from behaviors, because participants misremember having seen traits that were merely implied, not actually included, in behavioral sentences. False recognition rates also indicate association of inferred traits with their corresponding faces, because participants recognize more of those traits that were originally implied in behaviors presented with corresponding faces than traits presented with faces from noncorresponding trials.

Building on the false recognition paradigm, participants in the current experiments—with the exception of Experiment 5—were told that some of the sentences described the person whose face was presented with the sentence and that others were randomly assigned to the faces. Moreover, in three experiments, pairs of faces and behaviors were presented simultaneously, such that each behavior was clearly relevant to one face and irrelevant to the other. Experiment 5 was modeled after the classic "communicator-target" STT studies (Carlston et al., 1995; Crawford et al., 2007) in which one of the faces described the other face and vice versa. In addition to the false recognition trials, in two experiments we included explicit trait judgments to provide convergent evidence that trait associations are formed with a relevant face but not with an irrelevant face.

The complete process of STI involves not only the spontaneous inference of a trait from a behavior, but also the binding of that trait to an actor's face. Normatively, binding of STI to faces should be eliminated by the knowledge that the trait-implying behaviors do not pertain to a particular target person. However, if this binding reflects shallow associative processes as opposed to attribution processes (e.g., Carlston & Skowronski, 2005), this knowledge should not eliminate the binding of trait inferences to the "wrong" nonactor faces, hence revealing an STT effect. On the other hand, to the extent that binding of STI to actor faces is an automatic attribution process (Todorov & Uleman, 2003), the false recognition effects should be stronger for STI than for STT (where attribution is inappropriate).

It could be argued that attribution processes (binding of STI to actor faces) are built over automatic shallow association processes (binding of STI to nonactor faces). According to this reasoning, the default process is association (resulting in STT when actors are absent, for example), which is subsequently buttressed by attribution processes (resulting in the stronger STI when actors are present).

Alternatively, it may be that there is no "default" or underlying process at all, but rather, that simple associations (underlying STT) and exclusive attributions (underlying STI) represent two distinct processes (a view consistent with Carlston & Skowronski, 2005). With the current experiments, we build on this view, suggesting that the two processes, rather than simply being independent of each other, are actually exclusive, such that only one process can occur at a time, like a switch between attributions and associations. Figure 1 depicts the switch-like process of the formation of STI versus STT. On the left, we see that when a relevant target is present—regardless of the presence or absence of an irrelevant target—STI are formed, to the exclusion of STT. On the right, we see that when a relevant target is absent, if an irrelevant target is available, STT occurs, to the exclusion of STI. If no irrelevant target is available, neither STI nor STT occurs, in the sense that there is no target to which the "free floating" trait can bind.

The current experiments allow us to test whether the conceptual diagram validly describes STI and STT as a function of the presence or absence of relevant and irrelevant targets. Experiments 1 and 2 test whether STT occurs when only irrelevant targets are present. Experiment 3 tests whether STT is eliminated in the perceptual absence of an irrelevant target. Experiments 4, 5, and 6 test whether STI are formed to the exclusion of STT when both relevant and irrelevant targets are present. These experiments are not intended to provide an exhaustive or conclusive test of the conceptual model. Rather, they are meant to add to the existing literature on trait inference processes and to build up a base of evidence for the model proposed.

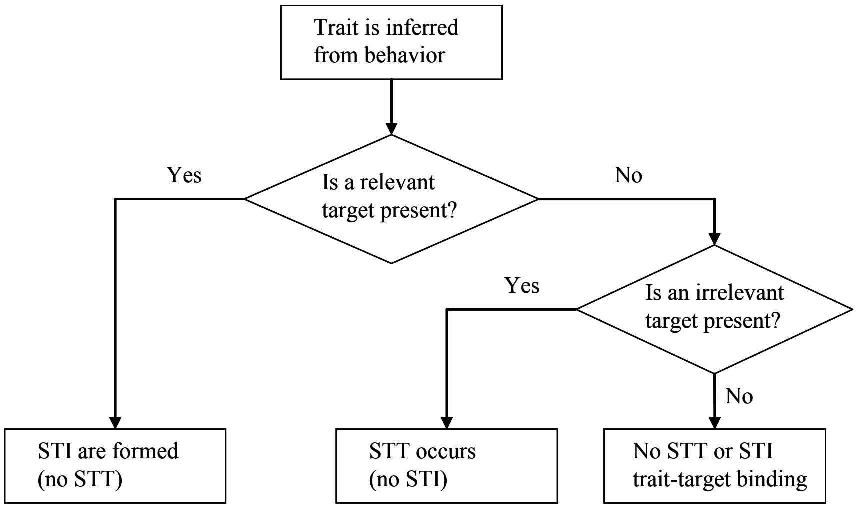

FIGURE 1. Conceptual diagram of trait inferences leading to STI vs. STT (with respect to binding of traits to targets) as a function of the presence or absence of relevant or irrelevant targets.

EXPERIMENT 1

The first experiment was designed to replicate STI and STT effects using a false recognition paradigm and instructions indicating the relevance or irrelevance of target faces for trait attributions. The irrelevance instructions used in the current studies are less ambiguous than many mundane situations, in that participants are explicitly told that behaviors have nothing to do with the faces shown. In mundane situations, the relevance of connections is more ambiguous. For example, one person may be communicating something about an unseen other (e.g., Carlston et al., 1995). In such cases, it may be reasonable to infer that communicators share the traits that they imply about others (e.g., because friends share similar characteristics, or because communicators' messages may be taken to imply that the implied traits are salient to them and therefore legitimately attributable to their own personalities). With these considerations in mind, the current manipulation of irrelevance is a conservative test of the occurrence of STT, because there is truly no reason for participants to associate implied traits with their irrelevant targets. In fact, Skowronski et al. (1998, Exp. 3) used a similar manipulation and obtained evidence for STT. However, their experiment did not include an STI control condition.

In both the irrelevance and "communicator-target" manipulations–used elsewhere (e.g., Carlston et al., 1995) and in our Experiment 5, in which one person communicates a piece of information about another person–the information is normatively irrelevant to one's impression of the person with whom the information is originally presented. In other words, both manipulations offer no logical basis on which to attribute a trait to the person. Thus, any consequent associations between person and (irrelevant) trait must reflect simple associative processing and evidence of STT, if one considers STT broadly as the transference of traits to unintended or inappropriate targets.

In the first stage of the experiment, participants were presented with 60 face-behavior pairs. The relevance of the behavior was manipulated by informing participants that the behavior either corresponds to the person pictured or is randomly paired up with the person and has nothing to do with him or her. We used a fast 3-second presentation of face-behavior pairs. If binding of trait inferences to faces is an automatic process, this presentation should be sufficient. Prior research has shown that although participants study a face-behavior pair on average for 6 seconds, a 2-second presentation is sufficient for binding of STI to actor faces to occur (Todorov & Uleman, 2003).

METHOD

Participants. Participants were 30 Princeton undergraduates who received partial course credit.

Stimulus Material. We borrowed 60 behavioral, third-person sentences from Uleman (1988), replacing pronouns with proper names. Additional, minor modifications were made in order to modernize the statements. The sentences represented a wide range of behaviors, implying various, distinct traits. In the recognition test, the implied traits themselves were presented, and each trait word appeared only once. For visual "person" representation, we used black-and-white headshots of college students from somewhere other than Princeton University. In order to avoid potential confounding of faces (e.g., positive expressions) with particular types of behaviors (e.g., positive behaviors), the sentences were randomly paired with the faces, and names were randomly paired with faces with the sole constraint that they match gender.

Procedures. Participants were told that the study examined how people remember information about others. Participants worked individually in isolated rooms. Detailed instructions were given via computer. The experiment was said to consist of two parts. In the first part, participants were told that they would be shown faces accompanied by sentences. Participants were urged to memorize both faces and sentences in preparation for an unspecified memory test later on—there were no explicit inducements to form impressions. The memorization instruction was also meant to encourage participants to pay attention to both faces and sentences, especially in the irrelevant trials.

During the study phase of the experiment, participants were each presented with 60 trials consisting of single faces paired with single, trait-implying behavioral sentences displayed below the faces. Trial order was randomized by computer for each participant. The 60 face-behavior pairs were subdivided into two sets of 30 faces that were each paired up with sentences said to be relevant or irrelevant to the person with whom they were presented. The sentences were presented in either green or red font to indicate their relevance (green) or irrelevance (red) to the corresponding faces. The cover story was that "green" sentences were about the person, whereas "red" sentences were "randomly assigned" to the faces and were therefore completely "irrelevant" to the face with which they were presented. Each trial consisted of a 3-second face-behavior presentation, followed by a 2-second inter-trial interval.

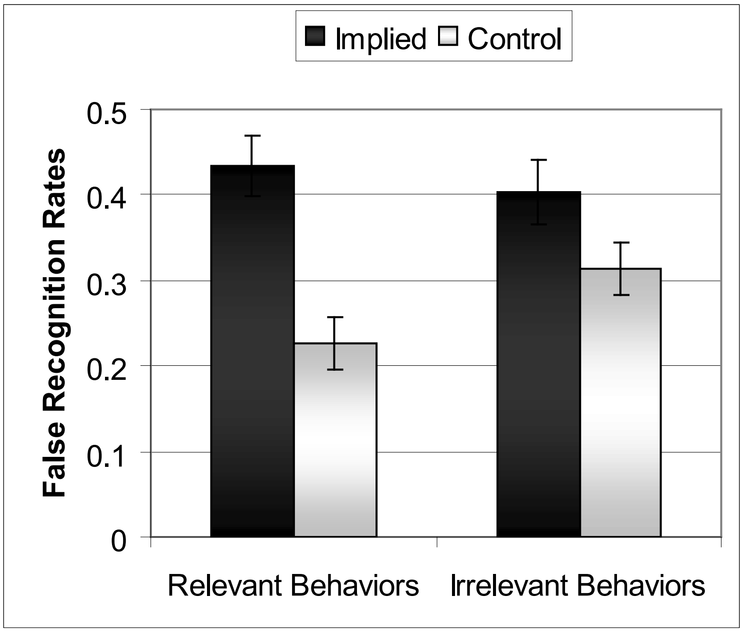

FIGURE 2. Mean proportion of false recognition ("yes") responses as a function of face-trait pairing and stated relevance of the behavior to the person (Experiment 1). Standard error bars are shown.

On 20 trials, the behavioral sentences contained the trait implied by the behavior (filler sentences; e.g., "Douglas was curious and checked out three books on the village's history"—this sentence both implies and includes the trait "curious"). These trials were included to ensure that participants were vigilant for correct "yes" responses during the subsequent recognition task—otherwise, all "yes" responses would indicate *false* recognition, and participants might have been tempted to respond "no" across the board if they suspected that no implied traits were ever actually included among the sentences. In the remaining 40 trials, the sentences described behaviors that only implied, but did not explicitly include, the traits (trait-implying sentences; e.g., "Henry solved the mystery halfway through the book"—this implies the trait "clever."). Within each of these trial subsets, half of the sentences were presented as relevant and half as irrelevant. Relevance was counterbalanced such that the sentences presented as relevant in one condition were irrelevant in the other condition, and vice versa.

The study phase began with two practice trials (one "green" and one "red" behavior). Upon completion of the 60 study trials, participants were prepared for the recognition task, during which they were told they would be presented with pairs consisting of faces from the study phase and single words (the implied traits). The participants' task was to indicate whether, during the study phase, they saw the word in the sentence that corresponded to the particular person shown. They did this by pressing either the "Yes"-labeled key (the "/" key) or the "No"-labeled key (the "Z" key) on the keyboard. Participants were given two practice trials preceded by 3-second re-presentations of the two faces and sentences from the previous practice trials. For the practice trials themselves, each face was presented with a trait word that was either a part of the sentence presented with the face (in one trial) or not (in the other). Participants indicated, with the "Yes" and "No" keys, whether they thought the word had appeared in the sentence corresponding to the face, and they were given feedback on whether their response was correct. None of the participants expressed difficulty in understanding the instructions. Before continuing to the actual recognition task, participants were reminded to work as quickly and accurately as possible.

The recognition task consisted of 60 face-trait pairs presented sequentially in randomized order for each participant. Each face had been presented in the study phase, and each trait had been implied by a sentence in the study phase, such that none of the stimuli were truly novel (this conveniently rules out any alternative explanations of false recognition effects being due to the relative novelty of control versus implied trials). Trait words were presented below the faces, as were the prompts "Yes" and "No," which remained on screen until participants made their choice. Subsequent trials followed immediately upon participants' responses. On 20 trials, the faces were the ones corresponding to the filler sentences—those that explicitly contained the otherwise implied traits—paired with the actual trait words included in those sentences during the study phase. In all of these cases, and only in these cases, a "Yes" response indicated *correct* (rather than false) recognition of a trait word as having been included in the original sentence presented with the face. Because these trials were included to make the false recognition task plausible, they were not analyzed.

In preparing the stimulus materials, the 40 faces that were not paired with filler sentences during the study phase had been randomly subdivided into two fixed sets of 20. These sets were used for the 40 trials of the recognition task that showed implied-only (i.e., not explicitly mentioned) traits.

For the 20 random-pairing (*control*) trials, we took one of these sets of 20 faces and paired them with 20 traits implied by sentences corresponding to *other*, different faces from within the same set. In other words, for the 20 control trials, each trait word was implied by a different sentence than the one originally presented with its corresponding face—in this subset, a "Yes" response indicated false recognition of a noncorresponding trait. Although the two sets of faces were fixed, the pairing of specific traits and faces was randomized for each participant (basically, face-trait pairs during control trials differed randomly between participants). In order to prevent the accidental pairing of faces with their own corresponding sentences from the presentation, faces and sentences were further subdivided into two sets of 10, each of which was paired with faces from the opposite, noncorresponding set before being randomized. Therefore, the randomization was somewhat limited, in that each item in one set could only be paired with one of 10, rather than 20, items in the other set. This limitation is trivial, in that it only affects the variety of control pairings.

For the 20 remaining, systematic-pairing (*implied*) trials of the recognition task, we paired the other set of faces with the 20 traits implied by sentences corresponding to the faces shown. For these trials, each trait word was implied by the sentence originally presented with the concomitant face—in this subset, a "Yes" response indicated false recognition of a corresponding implied trait. The two groups of 20 face-trait pairs were counterbalanced such that control trials in one condition served as implied trials in the other condition.

The overall design was a 2 (trait pairing: implied about person shown versus control) x 2 (stated relevance of behavior: relevant versus irrelevant) within-subjects factorial.

RESULTS AND DISCUSSION

As shown in Figure 2, there was a significant interaction between relevance and pairing, $F(1, 29) = 5.80$, $p = .023$, $\eta^2 = .167$. Planned comparisons showed that while the difference between implied and control pairings was significant in all cases, main effect $F(1, 29) = 32.87$, $p < .001$, $\eta^2 = .531$, it was larger for relevant traits, $t(29) = 7.00$, $p < .001$, $\eta^2 = .628$, than for irrelevant traits, $t(29) = 2.22$, $p = .034$, $\eta^2 = .146$. There was no main effect for stated relevance, $F(1, 29) = 1.35$, $p = .255$, $\eta^2 = .044$, indicating that false recognition rates were similar for traits implied by relevant and irrelevant behaviors.

According to Skowronski et al.'s (1998) model of STI, the first stage in STI is the activation of trait concepts. The second stage is the forming of associations between traits and targets. In terms of this model, given that all traits were implied, the overall false recognition rate provides a measure of trait activation. On the other hand, the difference between the false recognition rates for implied traits associated with corresponding faces (implied trials) and implied traits associated with noncorresponding faces (control trials) provides a measure of trait associations.[1]

Although for half of the trials the trait-implying behaviors did not pertain to the person depicted with the behavior, participants associated the inferred trait with the nonactor face, revealing an STT effect in the false recognition paradigm. This finding replicates Experiment 3 of Skowronski et al. (1998). However, as expected, the binding of STI to actors' faces was stronger than the binding to nonactor faces. This is the first demonstration of STT in the false recognition paradigm, extending previous findings in the savings paradigm (Carlston & Skowronski, 2005; Crawford et al., 2007; Mae et al., 1999; Skowronski et al., 1998) with an explicit manipulation of irrelevance and providing convergent evidence for the robustness of the STT phenomenon.

EXPERIMENT 2

One limitation of Experiment 1 is that it is unclear whether participants, when presented with red irrelevant versus green relevant behaviors, were encoding both sets of behaviors in a similar manner. For example, it is possible that in the irrelevant condition participants did not pay attention to the behavior, although this

1. Note that the relevant unit of analysis is the *difference* between implied and control trials, not the absolute level of either one. It is difficult to interpret control and implied false recognition rates in isolation, because other factors that differ between irrelevant and relevant conditions (e.g., difficulty discriminating between implied and control traits) may drive overall false recognition rates either higher or lower in either condition. In any case, although the difference in Exp. 1 seems to have been driven primarily by the control false recognition rates, this was not the case in any of the remaining five experiments (see Fig. 2 to 9).

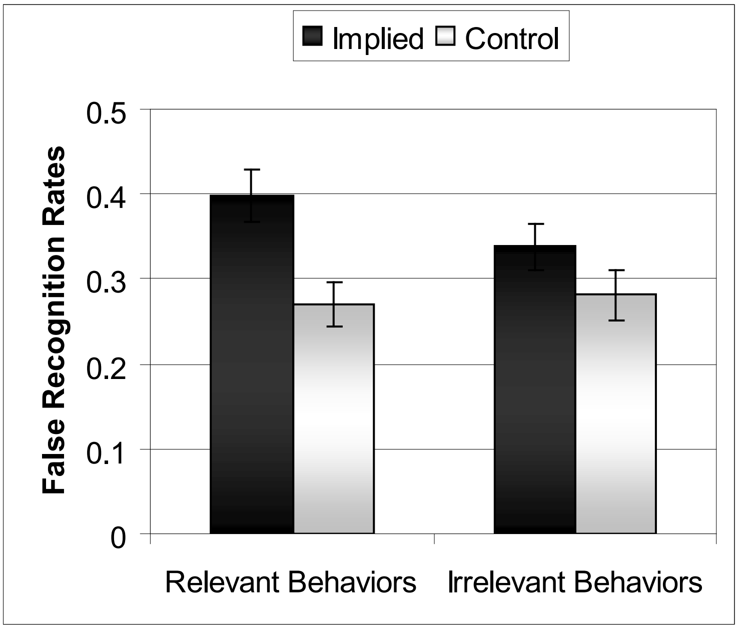

FIGURE 3. Mean proportion of false recognition ("yes") responses as a function of face-trait pairing and stated relevance of the behavior to the person (Experiment 2). Standard error bars are shown.

possibility is inconsistent with the STT effect observed in Experiment 1. Nevertheless, Experiment 2 ensured that encoding was similar for both irrelevant and relevant behavioral sentences, by presenting relevance information *after* the presentation of face-behavior pairs. Given the limited time between encoding trials, this experiment provides a more stringent test of the hypothesis that STI effects are stronger than STT effects. A decrease in STT could occur only to the extent that participants were able to overwrite their initial encoding of trait inferences and faces as bound to one another.

METHOD

Participants. Participants were 47 Princeton undergraduates who received partial course credit. One additional participant's data was lost due to a program malfunction.

Stimulus Material. The same 60 face-behavior pairs were used as in Experiment 1.

Procedures. The procedure was identical to that of Experiment 1, with two minor modifications. First, all behaviors were presented in black font, not in green or red. Second, relevance was indicated by either green (relevant) or red (irrelevant) dots flashed for 1 second immediately following each face-behavior pair. Therefore, each trial consisted of a 3-second face-behavior presentation, followed immediately by a 1-second relevance indicator (a large colored dot in the middle of the screen), and then a 2-second inter-trial interval.

The design was again a 2 (trait pairing: implied about person shown versus control) x 2 (stated relevance of behavior: relevant versus irrelevant) within-subjects factorial.

RESULTS AND DISCUSSION

As shown in Figure 3, although the interaction between relevance and pairing did not reach significance, $F(1, 46) = 2.60$, $p = .114$, $\eta^2 = .054$, planned comparisons showed that whereas the difference between implied and control pairings was significant for relevant traits, $t(46) = 3.94$, $p < .001$, $\eta^2 = .252$, it was only marginally significant for irrelevant traits, $t(46) = 1.76$, $p = .085$, $\eta^2 = .063$. The overall main effect for implied versus control traits was significant, $F(1, 46) = 14.70$, $p < .001$, $\eta^2 = .242$. There was no main effect for relevance, $F(1, 46) = 1.52$, $p = .225$, $\eta^2 = .032$, indicating that false recognition rates were similar for traits implied by relevant and irrelevant behaviors.

The findings from Experiment 2 were very similar to those of Experiment 1. However, there were slight differences between the statistical tests. Whereas the interaction between relevance and pairing was significant in Experiment 1, it was not significant in Experiment 2. At the same time, whereas the simple STT effect was significant in Experiment 1, it was marginally significant in Experiment 2. Given that the participants were drawn from the same population and that the experimental procedures were almost identical, we decided to test these effects in a three-way ANOVA including Experiment as a between-subjects factor.

The only significant effects to emerge from this analysis were a main effect for pairing, $F(1, 75) = 43.26$, $p < .001$, $\eta^2 = .366$, indicating that participants were more likely to recognize implied traits than control traits, and a significant interaction between relevance and pairing, $F(1, 75) = 7.81$, $p = .007$, $\eta^2 = .094$ ($F < 1$ for the triple interaction of relevance, pairing, and experiment), such that the difference between implied and control traits was smaller for irrelevant than for relevant traits. Both simple effects for relevant traits, $t(75) = 7.11$, $p < .001$, $\eta^2 = .403$, and irrelevant traits, $t(75) = 2.83$, $p = .006$, $\eta^2 = .097$, were significant.

These findings show that even when encoding was guaranteed to be similar for relevant and irrelevant face-behavior pairs, the STT effect was weaker than the STI effect. Regardless of the type of manipulation used—green versus red font that indicates relevance during encoding, as in Experiment 1, or green versus red dots that indicate relevance after face-behavior presentation, as in Experiment 2—the impact on false recognition is similar. Across both experiments, there was evidence of STI, as well as (weaker) STT.

EXPERIMENT 3

Both Carlston and Skowronski (2005) and Todorov and Uleman (2004) argued that whereas shallow associative processes underlie STT effects, attributional processes underlie STI effects. One implication of this hypothesis is that binding of STI to nonactor faces should be more perceptually driven than binding of STI to actor faces. In the previous experiment, faces and behaviors were presented simultaneously. This presentation can facilitate the formation of associations between faces and STI. In fact, Brown and Bassili (2002) have shown that people can form associations between STI and inanimate objects (e.g., a "superstitious banana") when these objects and behaviors are presented simultaneously.

If binding of STI to nonactor faces reflects simple perceptual binding, as argued by Todorov and Uleman (2004), then separating faces and behaviors perceptually should eliminate the STT effect. To test this hypothesis, we manipulated relevance during the presentation of behaviors, but subsequent to face-only presentation. Behavioral descriptions came after the face-only presentations, ensuring that participants had no simultaneous, perceptual presence of stimuli (i.e., the faces) with which to associate the irrelevant behaviors. Participants had to encode the faces up front and subsequently link the behaviors to the previously appearing faces (in the "relevant" condition).

METHOD

Participants. Participants were 41 Princeton undergraduates who received partial course credit.

Procedures. Procedures were identical to those in Experiment 1, with the exception that behaviors were presented separately for 3 seconds immediately following each 3-second presentation of faces by themselves (a total face-plus-sentence presentation time of 6 seconds). There was a 2-second inter-trial interval. As in Experiment 1, the sentences themselves were presented in either green or red font, indicating relevance or irrelevance, respectively.

The overall design was a 2 (trait pairing: implied versus control) x 2 (stated relevance of behavior: relevant versus irrelevant) within-subjects factorial.

RESULTS AND DISCUSSION

As shown in Figure 4, there was a significant interaction between relevance and pairing, $F(1, 40) = 7.31$, $p = .010$, $\eta^2 = .155$. Whereas the difference between implied and control pairings was significant for relevant traits, $t(40) = 4.81$, $p < .001$, $\eta^2 = .366$, it was not significant for irrelevant traits, $t(40) = 1.07$, $p = .289$, $\eta^2 = .028$. The overall main effect for implied versus control traits was significant, $F(1, 40) = 18.94$, $p < .001$, $\eta^2 = .321$. There was a marginally significant main effect for stated relevance, $F(1, 40) = 3.90$, $p = .055$, $\eta^2 = .089$, indicating that false recognition rates tended to be higher for traits implied by relevant ($M = .36$, $SE = .026$) than by irrelevant ($M = .32$, $SE = .027$) behaviors.

Although the STT effect seemed to disappear in this experiment, it is possible that the lack of significant difference was due to insufficient statistical power. However, two things should be noted. First, the STT effect was substantially reduced in this experiment relative to the first two experiments. In terms of the more intuitive correlation measure of effect size (Rosenthal, Rosnow, & Rubin, 2000), the STT effect was reduced from .31 to .17. The contrast test of the effect sizes for the first two experiments and the effect size for the third experiment was significant, $z = 1.67$, $p < .05$ (one-tailed). Second, the STI effect in the third experiment did not seem to be reduced relative to the first two experiments. The corresponding correlation mea-

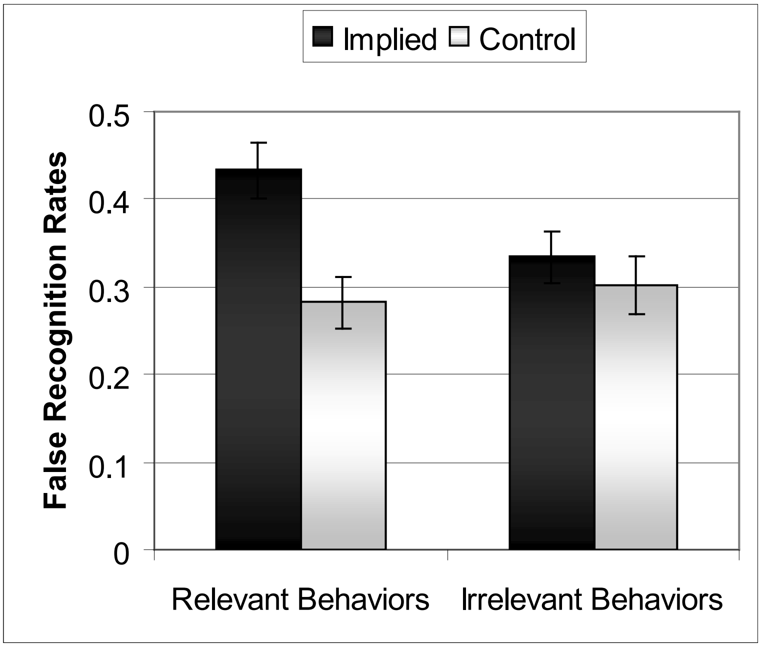

FIGURE 4. Mean proportion of false recognition (“yes”) responses as a function of face-trait pairing and stated relevance of the behavior to the person (Experiment 3). Standard error bars are shown.

sures of effect size were .61 and .63. For comparison, the average effect size of the six false recognition experiments reported in Todorov and Uleman (2002) was .66.

This pattern suggests that the binding of STI to nonactors—that is, STT effects—may be eliminated when the normal perceptual processes leading to spontaneous binding are interrupted. STT effects disappeared when participants were told that behaviors were irrelevant to faces presented prior to, and separately from, those behaviors. In contrast, the binding of STI to actor faces was intact despite the perceptual separation of faces and behaviors. These findings are consistent with the Carlston and Skowronski (2005) hypothesis that whereas STT are driven by shallow associative processes, STI are driven by attributional processes.

STT was prevented only when (a) there were no appropriate targets to which to attribute the inferred traits (due to the irrelevance instruction), and (b) there were no irrelevant faces with which to associate the traits (due to the perceptual separation of stimuli). This strongly suggests that irrelevant stimuli must be present at the time of trait inference in order for traits to bind to them. The current experiment shows that STT elimination can occur via a novel approach, perceptual separation, which may prevent simple associations between irrelevant traits and faces from being formed during encoding. This is consistent with the attentional account proffered by Crawford et al. (2007) and described below. The current study also provides additional support for the attributional nature of STI, which should reasonably bridge stimuli across small temporal and perceptual gaps. What is unclear from the current experiment is whether STT can be prevented when both relevant and irrelevant stimuli are present at the same time.

EXPERIMENT 4

The fourth experiment sought evidence for STT elimination in a false recognition paradigm similar to that used by Todorov and Uleman (2004), with pairs of faces and behaviors shown simultaneously, rather than one face-behavior pair presented at a time. Todorov and Uleman found that false recognition of an inferred trait paired with its assigned face on screen was more likely to occur than false recognition of a trait paired with the other face shown concomitantly. Although this demonstrated that STT effects were significantly reduced *relative* to STI, the authors did not provide a control condition against which to test whether STT effects, while smaller, might nevertheless have still been present—that is, there was no way of testing STT *elimination* in Todorov and Uleman's studies. Therefore, the current experiment included an appropriate control condition.

The current experiment was also modeled after Crawford et al.'s (2007) experiments. These authors argued that the elimination of STT in a dual-face-presentation paradigm was due to participants' lack of attention to the second, irrelevant person shown. Crawford et al. presented participants with pairs of photographs of people and a trait-implicative paragraph that was either used to describe one person (the actor) in the presence of the other (the observer) or used by one person (the informant) to describe the other (the target). Although there was no evidence of STT (as indicated by the lack of facilitation in relearning traits in the presence of nonactors) to either the observers or the informants, there also was no need for participants to encode the observers or informants in the first place. The authors concluded quite reasonably that the typical associations formed between traits and nonactors (STT) need not occur under certain circumstances, such as those of their paradigm. However, they left unanswered the question of whether they prevented STT via mere distraction or diverted attention, or whether STT can be prevented even when attention to irrelevant faces is ensured. At the very least, Crawford et al.'s participants could have spent more time studying the relevant face than the irrelevant face on each trial (see also Todorov & Uleman, 2004), and so attention was not entirely controlled.

The current experiment ensured attention to appropriate and inappropriate target faces and behaviors following Todorov and Uleman's (2004) procedures. Specifically, the two faces were presented with two behaviors, each ostensibly communicated by one of the faces and describing the other face. This allows for a more conservative test of STT robustness and a more impressive demonstration of STT elimination that goes beyond a simple attentional account.

With the encoding of nonactors ensured, if STT fails to occur, we can rule out a constantly active associational process as accounting for the STT effects, because there is no a priori reason to assume that the inappropriate associations will be prevented by the mere presence of appropriate attributional targets—in other words, why would the traits not become associated with both the relevant target *and* the nonactor, as shown in prior research (Brown & Bassili, 2002)? The elimination of STT in the context of mandated attention suggests instead that underlying STT effects is an associational process that is deactivated when relevant targets are present. Crawford, Skowronski, Stiff, and Leonards (2008) have also ruled out differences in visual attention, and our current studies therefore complement and strengthen their findings using different methods.

The actual process may involve sequential processing of two face-behavior pairs rather than simultaneous processing of all stimuli, but it still has real-world relevance in that such processing is likely to occur even in mundane situations (an observer may choose to encode Jane's face with Jill's statement about Jane and then encode Jill's face with Jane's statement about Jill, but the fact that this feat is even possible without transferring inferred traits from the appropriate target to the other one is informative and contrary to the notion that inappropriate, multiple associations are inevitable). As Todorov and Uleman (2004) have argued, spontaneous allocation of attention to relevant faces is part of the process of binding STI to these faces.

METHOD

Participants. Participants were 38 Princeton undergraduates who received partial course credit. One additional participant was not included because she misunderstood the sentence-face connection.

Procedures. The basic procedures were similar to those of Experiment 1, with the notable exception that during the initial presentation phase, instead of single face-behavior pairs, two face-behavior pairs were presented one above the other on screen, such that the top behavior referred to the face just above it, whereas the bottom behavior referred to the face just above it and below the top face-behavior pair. Further, all sentences were placed within quotation marks, and participants were instructed that each sentence was a description of the face with which it was presented, as communicated by the other person shown on screen. This was clarified in detail and with examples—the bottom sentence was a description of the bottom face by the top person, whereas the top sentence was a description of the top face by the bottom person.

Pairings of face-behavior pairs within each study trial were prepared randomly, with the constraint that gender of top and bottom faces always matched. In Experiment 4, the 60 faces and behaviors were presented over 30 trials, rather than 60, and participants were therefore given 12 seconds, not 3, to memorize each slide, with a 2-second inter-trial interval (most participants volunteered that this was a difficult memorization task, even given the additional time). In Experiment 4, all behaviors were described as relevant to the faces with which they were presented, and the sentences were presented in the same black font, not green or red as in Experiments 1 and 3.

The recognition task was almost identical to the one in the first experiments, with 60 face-trait pairs presented sequentially, preceded by two practice trials. As before, the 40 faces that were not paired with filler sentences during the study phase were subdivided and paired up with different traits from the original study trials. Ten faces were paired with traits that were implied by the behaviors corresponding to those same faces during the study phase (implied-same face trials, used to measure STI)—in this subset, a "Yes" response indicated false recognition of a corresponding implied trait. Ten faces were paired with traits implied by the behaviors corresponding to the *other* faces shown simultaneously on screen during the study phase (implied-other face trials, used to assess STT effects)—in this subset, a “Yes” response indicated false recognition of a noncorresponding trait. Ten faces were paired with traits corresponding to faces from other study trials in the same position on screen (control-same trials, used as a control condition for STI analysis), and 10 were paired with traits corresponding to other study trials in the opposite position on screen (control-other trials, used as a control condition for STT analysis). The four groups of 10 face-trait pairs were counterbalanced such that control trials in two conditions served as implied trials in the other two conditions, and faces shown on the bottom in two conditions were shown on top in the other two.

The overall design was a 2 (trait pairing: implied about person on screen versus control) x 2 (face relevance: relevant versus irrelevant) within-subjects factorial.[2]

RESULTS AND DISCUSSION

As shown in Figure 5, there was a significant Relevance x Pairing interaction, $F(1, 37) = 6.64$, $p = .014$, $\eta^2 = .152$. Planned comparisons showed that whereas false recognition was significantly higher for implied versus control pairings with face relevant traits, $t(37) = 3.85$, $p < .001$, $\eta^2 = .286$, there was no difference for face irrelevant traits, $t(37) = 0.28$, $p = .782$, $\eta^2 = .002$. The main effect for implied versus control traits was significant, $F(1, 37) = 11.41$, $p = .002$, $\eta^2 = .236$, as was a main effect for relevance, such that face relevant traits were falsely recognized more often than face irrelevant traits, $F(1, 37) = 6.85$, $p = .013$, $\eta^2 = .156$.

Replicating Crawford et al. (2007), spontaneous trait transference to an irrelevant concomitant face did not occur in the current experiment, in spite of the finding that STI did occur, as expected, with both faces shown simultaneously. The dual face-behavior nature of the design, as well as the significant STI results across stimuli, rules out the possibility that participants simply failed to attend to both pairs of faces and behaviors on screen. On the contrary, associations were obviously formed between corresponding faces and inferred traits (as indicated by STI), and furthermore, those associations were limited to those corresponding pairs in an exclusive manner that precluded the inappropriate association of traits with noncorresponding faces.

This experiment provides a conceptually and methodologically significant improvement over previous studies that did not ensure attention to all stimuli or establish whether STT was indeed eliminated or simply reduced relative to STI (e.g., Todorov & Uleman, 2004). The findings show that it is possible to eliminate STT even when attention is allocated to stimuli (irrelevant face-behavior pairs) that in our prior experiments (and in other similar studies), participants were unable to avoid associating with each other.

2. The reader should note that relevance was operationalized in two different ways throughout the experiments, and Experiment 6 manipulated relevance in both ways. Therefore, to avoid confusion in the results and discussion, “face relevance” refers to whether traits are paired with the same, corresponding face (relevant) or other, noncorresponding face (irrelevant) of the pair that was shown on screen with the trait-implicative behavior during the initial study phase, whereas “stated relevance” refers to whether trait-implying behaviors were actually said to be relevant or irrelevant to the faces shown during the study phase.

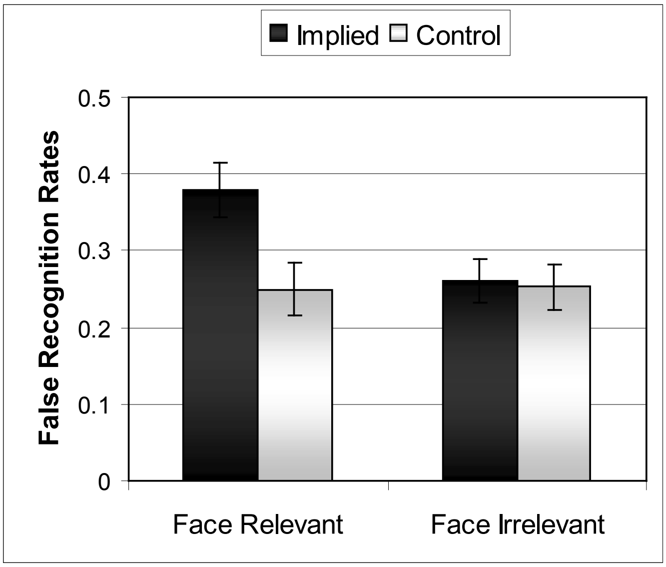

FIGURE 5. Mean proportion of false recognition ("yes") responses as a function of face-trait pairing and face relevance (relevant—same face—vs. irrelevant—other face on the same screen) (Experiment 4). Standard error bars are shown.

EXPERIMENT 5

Although the "communication context" of Experiment 4 is consistent with the possibility of legitimate trait attributions to the communicator, there was no evidence for STT. For example, even though Jackie says that *Jill* is stingy, an observer might infer, consciously and deliberately, that Jackie is also stingy, perhaps based on the fact that Jackie chose to talk about stingy behaviors and seemed to focus on that particular trait. Such a conscious inference would thereby enhance any automatic tendency to attribute stinginess to Jackie on the sole basis of concomitant face-behavior presentation. However, as in Crawford et al. (2007), we failed to find evidence for STT.

Experiment 5 replicated Experiment 4 without using "communication" instructions. If STT occurs in the presence of completely irrelevant faces, this would be strong evidence of the tenacity of the effect. This experiment was conceptually similar to the first three experiments and improved upon Experiment 4 by presenting participants with less complex encoding trials. In addition to the false recognition trials, we included an explicit trait rating task to provide convergent evidence that trait associations are formed only with relevant targets. If this is the case, participants should apply implied traits only to relevant faces.

METHOD

Participants. Participants were 26 Princeton undergraduates who received partial course credit.

Procedures. The procedures were identical to those of Experiment 4, with two minor modifications. The sentences were not placed within quotation marks, and

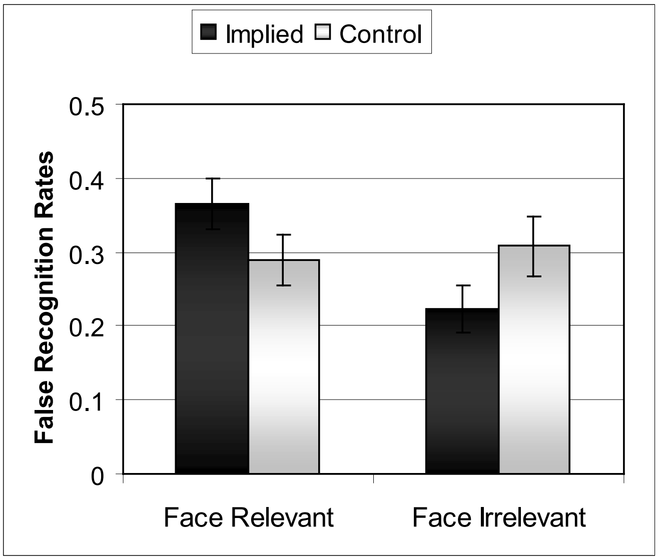

FIGURE 6. Mean proportion of false recognition ("yes") responses as a function of face-trait pairing and face relevance (Experiment 5). Standard error bars are shown.

participants were simply told that the bottom sentence was a description of the bottom face, whereas the top sentence was a description of the top face. The recognition task was the same as in Experiment 4.

As a more explicit test of trait attribution, a trait rating task was added to this experiment, in which participants were presented with the same 60 face-trait pairs as in the recognition task, but in a new random order for each participant. Participants were asked to indicate, based on their memory for the information from the behavioral sentences originally paired with the faces, "How certain are you that this person is [implied trait]?" Responses were on a scale of 1 ("Certain he/she ISN'T") to 3 ("Not Sure") to 5 ("Certain he/she IS").

The design was a 2 (trait pairing) x 2 (face relevance) within-subjects factorial.

RESULTS AND DISCUSSION

As shown in Figure 6, there was a significant interaction between relevance and pairing, $F(1, 25) = 12.28$, $p = .002$, $\eta^2 = .329$. Planned comparisons showed that whereas false recognition was significantly higher for implied versus control pairings with face relevant traits, $t(25) = 2.21$, $p = .036$, $\eta^2 = .164$, there was a marginally significant reversal with face irrelevant traits, $t(25) = 1.94$, $p = .063$, $\eta^2 = .131$. There was no main effect for pairing, $F < 1$, but participants were more likely to falsely recognize face relevant versus irrelevant traits, $F(1, 25) = 3.99$, $p = .057$, $\eta^2 = .138$.

Ratings of trait applicability mirrored the false recognition findings of interest. As shown in Figure 7, there was a significant interaction between relevance and pairing, $F(1, 25) = 15.80$, $p < .001$, $\eta^2 = .387$. Planned comparisons showed that whereas trait applicability was significantly higher for implied versus control pairings with face relevant traits, $t(25) = 5.69$, $p < .001$, $\eta^2 = .565$, there was no difference

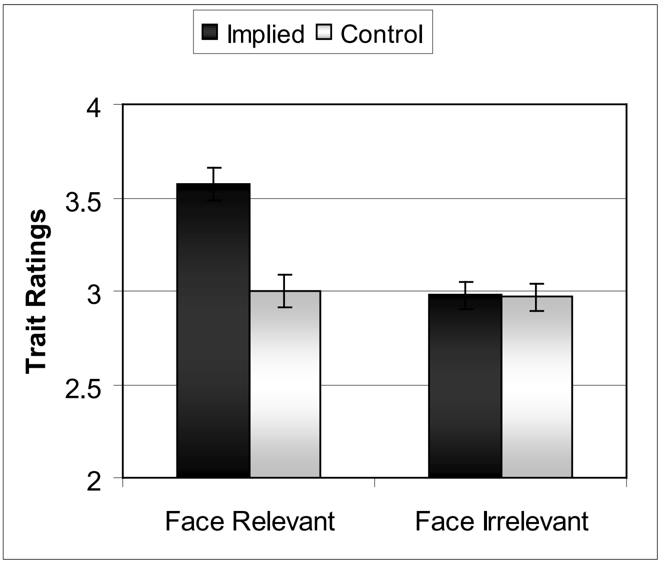

FIGURE 7. Mean ratings of trait applicability, from 1 (“Certain he/she ISN’T”) to 3 (“Not Sure”) to 5 (“Certain he/she IS”), as a function of face-trait pairing and face relevance (Experiment 5). Standard error bars are shown.

for face irrelevant traits, $t < 1$. There was a main effect for pairing, $F(1, 25) = 18.75$, $p < .001$, $\eta^2 = .429$, and a main effect for face relevance, $F(1, 25) = 16.43$, $p < .001$, $\eta^2 = .397$.

Replicating the first four experiments, as well as Todorov and Uleman’s (2004) studies, this experiment showed that the binding of STI to actor faces was stronger than binding to nonactor faces. In fact, as in Experiments 3 and 4, STT effects were eliminated, but under conditions in which both relevant and irrelevant targets were simultaneously present. This experiment also goes beyond Todorov and Uleman’s studies because those studies did not include proper conditions to estimate whether STT effects occur at all. The experiment also provided convergent evidence from two different measures of trait inferences—implicit (false recognition) and explicit (trait ratings). This parallels the findings of Crawford et al. (2007). The findings support the notion that an exclusive, automatic attribution process—not a simple associational mechanism—is responsible for binding traits to appropriate, relevant targets.

EXPERIMENT 6

The sixth experiment sought further evidence for STT elimination by replicating Experiment 5 and adding a stated irrelevance condition, in which the automatic and exclusive attribution process should not occur, and only simple face-trait associations should be formed. In other words, when participants are told that behaviors are irrelevant to the faces with which they are presented, we should see transference of traits to noncorresponding faces even in a dual-face paradigm. However, when behaviors are said to be relevant, STT should be eliminated.

That is, the objective of this experiment was not only to demonstrate how STT can be eliminated but also to delineate the conditions under which STT can be *created* in a dual-face paradigm, allowing for greater understanding of the underlying mechanisms and providing more support for our conceptual model. By including a stated irrelevance condition in this experiment, any resulting STT effect serves as strong evidence of the tenacity of the effect, because, as mentioned previously, STT in the presence of completely irrelevant faces—that is, ones for which the behavior is merely a statement, not a communication—occurs in the absence of any plausible attributional reason for transference. As in Experiment 5, this experiment included an explicit trait rating task. We expected that these ratings would mirror the false recognition findings.

METHOD

Participants. Participants were 62 Princeton undergraduates who received partial course credit.[3]

Procedures. The procedures were identical to those of Experiment 5, with the addition of an irrelevance condition, in which participants were initially instructed that all behaviors are irrelevant to the faces with which they are presented: "The sentences have NOTHING to do with the faces! We are interested in people's visual versus verbal memory, and therefore whether people can keep memory for faces and sentences separate."

The overall design was a 2 (trait pairing: implied about person on screen versus control) x 2 (face relevance: relevant versus irrelevant) x 2 (stated relevance of behavior: relevant versus irrelevant) factorial with stated relevance as a between-subjects factor.

RESULTS AND DISCUSSION

Most pertinent to our hypothesis (see Figure 8), there was a marginally significant three-way interaction between stated relevance, face relevance, and pairing, $F(1, 60) = 3.73$, $p = .059$, $\eta^2 = .059$. Planned comparisons showed that in the stated relevant condition, there was a significant relevance by pairing interaction, $F(1, 30) = 12.48$, $p = .001$, $\eta^2 = .172$. Further planned comparisons revealed that whereas false recognition was significantly higher for implied versus control pairings with face relevant traits, $t(30) = 6.45$, $p < .001$, $\eta^2 = .410$, there was no difference with face irrelevant traits, $t < 1$. As in Experiments 4 and 5, therefore, STT effects were eliminated when behaviors were said to be about their corresponding faces.

In the stated irrelevant condition, there was only a main effect for pairing, $F(1, 30) = 9.26$, $p = .005$, $\eta^2 = .134$, and no interaction between relevance and pairing, $F < 1$. This demonstrates the typical STT effect, wherein traits bind to any target present, whether the same face or another face shown simultaneously. The

3. Data from six participants were not available for the trait applicability rating task, due to initial use of an older version of the experimental software. Although the false recognition data (which were collected prior to the rating task) were recorded, the program failed to record the trait rating data.

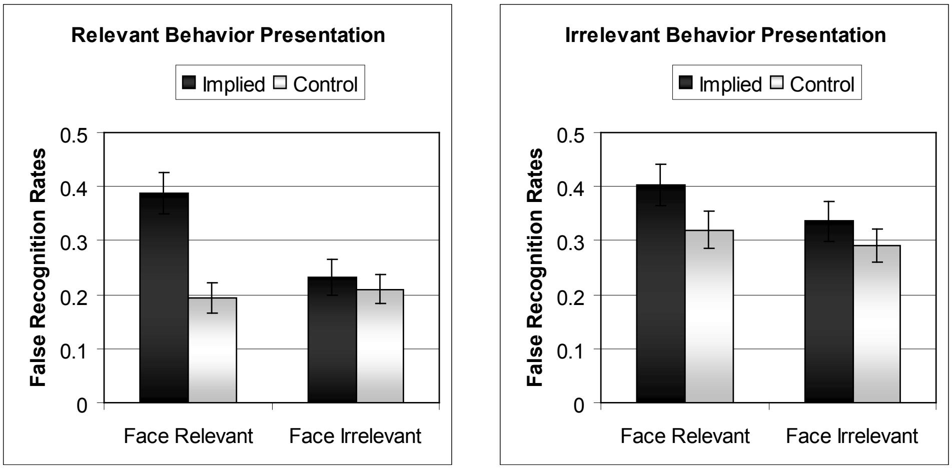

FIGURE 8. Mean proportion of false recognition ("yes") responses as a function of face-trait pairing, face relevance, and stated relevance of behavior (left panel = "relevant", right panel = "irrelevant"; Experiment 6). Standard error bars are shown.

reader may note that even when traits are paired with their corresponding faces in the stated irrelevant condition, false recognition differences between implied and control pairings are still indicative of STT, not STI; therefore, both same-face and other-face relevance in the irrelevant condition reveal STT effects, which renders the failure to reveal an interaction unsurprising.

There were the usual main effects for pairing, $F(1, 60) = 33.12$, $p < .001$, $\eta^2 = .336$, and for face relevance $F(1, 60) = 10.97$, $p = .002$, $\eta^2 = .155$. Additionally, there was a main effect for stated relevance, $F(1, 60) = 4.76$, $p = .033$, $\eta^2 = .074$, indicating that false recognition was higher for stimuli pairings said to be irrelevant ($M = .34$, $SE = .027$) than for those said to be relevant ($M = .26$, $SE = .026$). This finding may be due to the relatively greater difficulty of the task when stimuli were said to be irrelevant to each other, leading participants to be poorer at discriminating stimuli overall, since they were not supposed to form any explicit associations between the four stimulus objects (two faces, two behaviors) presented at a time. The face relevance by pairing interaction was significant, $F(1, 60) = 9.38$, $p = .003$, $\eta^2 = .135$, reflecting the smaller difference in false recognition rates between implied and control pairings for face irrelevant versus relevant traits. In other words, face irrelevant STT effects were on average smaller than face relevant effects, replicating the findings of the first five experiments.

As before, ratings of trait applicability mirrored the false recognition findings (see Figure 9). The three-way interaction between stated relevance, face relevance, and pairing was significant, $F(1, 54) = 8.25$, $p = .006$, $\eta^2 = .133$. Planned comparisons showed that in the stated relevant condition, there was a significant relevance by pairing interaction, $F(1, 27) = 23.07$, $p < .001$, $\eta^2 = .299$. Further planned comparisons revealed that whereas trait ratings were significantly higher for implied versus control pairings with face relevant traits, $t(27) = 7.03$, $p < .001$, $\eta^2 = .478$, there was no difference with face irrelevant traits, $t < 1$, again showing STT elimination when behaviors were said to be about their corresponding faces.

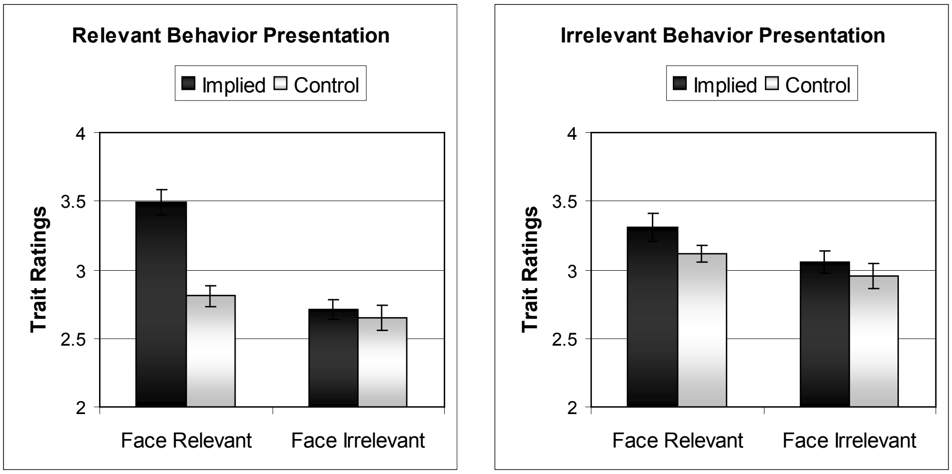

FIGURE 9. Mean ratings of trait applicability, from 1 ("Certain he/she ISN'T") to 3 ("Not Sure") to 5 ("Certain he/she IS"), as a function of face-trait pairing, face relevance, and stated relevance of behavior (Experiment 6). Standard error bars are shown.

In the stated irrelevant condition, there was a main effect for pairing, $F(1, 27) = 4.62$, $p = .041$, $\eta^2 = .079$, and no interaction between relevance and pairing, $F < 1$. Again, this reflects typical trait transference, using trait applicability ratings.

There were the usual main effects for pairing, $F(1, 54) = 28.61$, $p < .001$, $\eta^2 = .346$, and for face relevance $F(1, 54) = 32.47$, $p < .001$, $\eta^2 = .375$. Additionally, there was a main effect for stated relevance, $F(1, 54) = 6.62$, $p = .013$, $\eta^2 = .109$, indicating that trait ratings were higher for stimuli pairings said to be irrelevant ($M = 3.11$, $SE = .058$) than for those said to be relevant ($M = 2.92$, $SE = .048$). Again, this may be due to the relative difficulty of the task when stimuli were said to be irrelevant, leading participants to be less certain of trait applicability in that condition. The face relevance by pairing interaction was significant, $F(1, 54) = 15.37$, $p < .001$, $\eta^2 = .222$, reflecting the smaller difference in trait ratings between implied and control pairings for other (irrelevant) versus same (relevant) faces. There were also significant interactions between face relevance and stated relevance, $F(1, 54) = 4.79$, $p = .033$, $\eta^2 = .082$, and between stated relevance and pairing, $F(1, 54) = 5.33$, $p = .025$, $\eta^2 = .090$, which are not theoretically relevant beyond dissecting the aforementioned three-way interaction.

Replicating Experiments 4 and 5, this experiment showed that STT effects were eliminated, but only when behaviors were said to be relevant to the targets presented. Using a similar irrelevance instruction manipulation to the one used in Experiments 1, 2, and 3, this experiment demonstrated not only that STT effects occurred only when relevant targets were absent, but that traits were just as likely to be associated with one inappropriate face on screen as with another. This rules out the possibility that exclusive trait-face binding in the stated relevant condition was due simply to spatial proximity of corresponding faces and sentences (in which case traits would not have been associated inappropriately with the other faces on screen in the stated irrelevant condition), and most importantly, it provides further support for the notion that traits will be associated with any present target, unless the relevant target is present, in which case exclusive attributive binding will occur.

GENERAL DISCUSSION

Experiments 1 and 2 found that STT effects were weaker than STI, but still occurred, when presented targets were said to be irrelevant to their corresponding behaviors. Experiment 3 found that these STT effects could be eliminated when behaviors were presented immediately following their irrelevant corresponding targets, suggesting that STT requires that some sort of target must at least be present at the same time as the inferred trait. Experiments 4, 5, and 6 found that STT to irrelevant other-face targets was eliminated when relevant targets were simultaneously present. Further, Experiment 6 found, using the same paradigm as Experiment 5, that STT effects could be reinstated by instructing participants that behaviors are irrelevant to the faces with which they are presented. This suggests that STT can occur, but only when appropriate targets for trait attribution are absent. Experiment 6 contributes two novel and important pieces to the STT puzzle. First, it shows that STT can occur under conditions that at least superficially resemble those under which STT was consistently eliminated in other STT elimination studies (Crawford et al., 2007). Second, by doing so, Experiment 6 supports the idea that different associational and/or perceptual processes can be engaged when multiple faces and behaviors are presented but are simply described as irrelevant to each other. The findings demonstrate that the mere presence of a second (but still irrelevant) target face not only does not inevitably eliminate STT, but also does not guarantee that STT will only happen to the spatially proximal target face (as would be predicted by an attentional process that always links spatially proximal face-behavior pairs).

The present research replicates and extends previous research demonstrating that STI are bound to actors' faces (Carlston & Skowronski, 1994; Carlston et al., 1995; Mae et al., 1999; Skowronski et al., 1998; Todorov & Uleman, 2002, 2003, 2004). These studies have shown that the binding process is independent of explicit memory for the behaviors and cognitive resources during encoding of faces and behaviors, suggesting a fairly automatic process. One of the puzzling findings in this literature is that STI are also associated with completely irrelevant faces; that is, STT effects (Mae et al., 1999). This finding casts some doubt on the hypothesis that binding of STI to actor faces reflects person attribution processes. However, binding of STI to nonactor, irrelevant faces is weaker than binding of STI to actor faces. This has been shown in the savings paradigm and in trait judgment tasks (Mae et al., 1999; Skowronski et al., 1998), and was confirmed in a meta-analysis (Todorov & Uleman, 2004). In the current experiments, we provided convergent evidence for this phenomenon using the false recognition paradigm. Experiments 1, 2, and 6 provided evidence for STT, while across all six experiments, binding of STI to nonactor faces was either eliminated or was found to be weaker than binding of STI to actor faces.

Interestingly, and building on what little evidence there is that STT effects can be eliminated (Carlston & Skowronski, 2005; Crawford et al., 2007), four of our current experiments showed that STT is prevented whenever a relevant target is present or when an irrelevant target is absent in perceptual or temporal proximity.

It is remarkable, as well as worrisome, that STI can occasionally be attached to the "wrong" face. This finding poses interesting theoretical questions about the nature of the processes of binding of STI to actor faces and nonactor faces. Skowronski et al. (1998) suggested that STT "is a simple associational phenomenon devoid of the attributional activity presumably involved in spontaneous trait inference" (p. 847). The current findings support this view but constrain it with the notion that associations leading to STT are prevented when conditions are ripe for STI.

The current findings are consistent with the possibility that the process leading to binding of STI to actor (STI effects) and nonactor faces (STT effects) is like an automatic switch, allowing only exclusive binding of traits to relevant targets (STI) when those targets are present, but allowing other simple associations (STT effects) when relevant targets are absent. Another way to think about this is to conceptualize inferred personality traits as "free floating" entities that "stick" to any available target(s), unless appropriate targets are detected (or inappropriate targets unavailable), in which case the inferred traits bind only to the appropriate target, to the exclusion of other available stimuli.

This account does not necessarily imply serial, independent processes. In principle, the postulated processes can be implemented in parallel, connectionist models (e.g., Van Overwalle & Labiouse, 2004). In particular, models with learning rules that posit competition between different inputs and internal nodes predict that the stronger links or associations that are formed between certain stimuli (e.g., relevant face-trait pairs) can essentially block the formation of other links between one of those stimuli (e.g., the trait) and another stimulus (e.g., an irrelevant face). This can explain our face relevance findings. When participants view a behavior in the presence of an actor, the trait that they infer links strongly to the face of the actor, due to the legitimate correspondence between those stimuli. Any potential link between the inferred trait and another face is then preempted by the strong existing link. In the absence of a legitimate correspondence between face and behavior (e.g., in the stated irrelevance condition), one can assume that the proximal face-trait link is weak, and therefore the inappropriate link with another face is not blocked (allowing for other-face STT). Furthermore, in experiments in which only a single irrelevant face is presented (Experiments 1 and 2 here), there is nothing to block the proximal face-trait link. In Experiment 3, nothing blocks the irrelevant face-trait link, except that the face and trait are not present at the same time for that link to occur in the first place.

An interesting direction for future research may be to determine the limits of trait transference when relevant targets are missing—that is, how strongly, and to how many and what kinds of irrelevant targets can a trait bind in a given situation? Conversely, what are the limits of trait *inference* when relevant targets are actually present? That is, how strongly, and to how many and what kinds of relevant targets can a trait bind? For example, referring to Brown and Bassili's (2002) studies, if participants are told that an inanimate object such as a computer is attributed some behavior (e.g., it is bogged down by programs, implying "sluggish"), will the trait they infer from the behavior bind exclusively to the object, or is an actual human target necessary for exclusive trait-target binding to occur, meaning that a trait inferred about a computer will also be associated with other irrelevant targets (e.g., staplers, humans, etc.)? That is, will inanimate object STT occur even when the relevant target is present? Supporting this possibility is recent functional imaging evidence that suggests different neural processes underlying the association of trait attributes with people versus objects. Unlike with objects, person impressions invoke activity in medial prefrontal cortex (Mitchell, Macrae, & Banaji, 2005), a region that is also activated when people spontaneously retrieve trait knowledge about other people (Todorov, Gobbini, Evans, & Haxby, 2007).

An important theoretical question that remains to be clarified is whether STI and STT effects represent two distinct processes, or only one, or perhaps even more than two. In this paper, we are positing two processes and a single mechanism wherein a trait will bind exclusively or indiscriminately depending on the presence or absence of a relevant target, respectively, and that the mechanism is an automatic selector between attributional and associational processes. Another possibility is that STI and STT are the end result of a single, broader attributional process. Whether this alternative possibility is true or more than a mere semantic distinction remains to be determined. More importantly, it would be useful to discern the true nature of the selector mechanism underlying these effects. The initial results are consistent with the proposed mechanism and help rule out some alternative explanations (e.g., a simple attentional account of STT elimination), but further studies are needed to test this new model more comprehensively.

CONCLUSIONS

People form STI from single behaviors and effortlessly associate these inferences with faces. In some cases, when irrelevant faces are present and relevant targets are missing, these inferences can be associated with the wrong faces (STT). In other cases, when relevant faces are present, even in the presence of irrelevant faces, STT is prevented and only STI occur. This suggests a process by which traits bind exclusively to faces when relevant targets are present, but bind to available irrelevant faces when appropriate targets of attribution are absent. The current set of studies goes beyond existing research by examining novel conditions for the creation and elimination of spontaneous trait inference and transference, thus providing evidence that both supports and strengthens other findings, and helps converge upon a plausible model of the underlying process.

To return to our initial example, describing Jackie in her presence would not result in any trait inferences about Jill. However, describing Jackie in her absence may result in misattributing traits to Jill. One practical upshot of this research is that if you wish to badmouth people, you would be wiser to do it in their presence than behind their back, because in the latter case, the negative traits you imply will end up sticking to you to some degree. In contrast, you ought to say nice things about others behind their back (perhaps good advice in general) so that you benefit from your own generosity.

REFERENCES

Bassili, J. N. (1989). Traits as action categories versus traits as person attributes in social cognition. In J. N. Bassili (Ed.), *Online cognition in person perception* (pp. 61-89). Hillsdale, NJ: Erlbaum.

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). Saving 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. (2005). Linking versus thinking: Evidence for the different associative and attributional bases of spontaneous trait transference and spontaneous trait inference. *Journal of Personality and Social Psychology, 89*, 884-898.

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.

Crawford, M. T., Skowronski, J. J., & Stiff, C. (2007). Limiting the spread of spontaneous trait transference. *Journal of Experimental Social Psychology, 43,* 466-472.

Crawford, M. T., Skowronski, J. J., Stiff, C., & Leonards, U. (2008). Seeing, but not thinking: Limiting the spread of spontaneous trait transference II. *Journal of Experimental Social Psychology, 44,* 840-847.

Gilbert, D. T. (1989). Thinking lightly about others: Automatic components of the social inference process. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 189-211). New York: Guilford.

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

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, 55*, 685-694.

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.

Jones, E. E., & Harris, V. A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology, 3*, 1-24.

Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska symposium on motivation* (Vol. 15, pp. 192-238). Lincoln: University of Nebraska Press.

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.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2005). Forming impressions of people versus inanimate objects: Social-cognitive processing in the medial prefrontal cortex. *NeuroImage, 26*, 251-257.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. Cambridge, UK: Cambridge University Press.

Ross, L. D., Amabile, T. M., & Steinmetz, J. L. (1977). Social roles, social control, and biases in social-perception processes. *Journal of Personality and Social Psychology, 35*, 485-494.

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., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia, 45,* 163-173.

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.

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology, 39*, 549-562.

Todorov, A., & Uleman, J. S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology, 87,* 482-493.

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

Uleman, J. S. (1988). [Trait and gist inference norms for over 300 potential trait-implying sentences]. Unpublished raw data.

Uleman, J. S., Blader S., & Todorov, A. (2005). Implicit impressions. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious.* New York: Oxford University Press.

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.

Van Overwalle, F., & Labiouse, C. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review, 8*, 28-61.

Wyer, R. S., Budesheim, T. L., Lambert, A. J., & Swan, S. (1994). Person memory and judgment: Pragmatic influences on impressions formed in a social context. *Journal of Personality and Social Psychology, 66,* 254-267.

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