ATTITUDES AND SOCIAL COGNITION

Sowing the Seeds of Stereotypes: Spontaneous Inferences About Groups

David L. Hamilton
University of California, Santa Barbara

Jacqueline M. Chen
University of California, Irvine

Deborah M. Ko
Reading Room, Singapore

Lauren Winczewski
University of California, Santa Barbara

Ishani Banerji
Georgetown University

Joel A. Thurston
University of California, Santa Barbara

Although dispositional inferences may be consciously drawn from the trait implications of observed behavior, abundant research has shown that people also spontaneously infer trait dispositions simply in the process of comprehending behavior. These spontaneous trait inferences (STIs) can occur without intention or awareness. All research on STIs has studied STIs based on behaviors of *individual* persons. Yet important aspects of social life occur in *groups*, and people regularly perceive groups engaging in coordinated action. We propose that perceivers make spontaneous trait inferences about groups (STIGs), parallel to the STIs formed about individuals. In 5 experiments we showed that (a) perceivers made STIGs comparable with STIs about individuals (based on the same behaviors), (b) a cognitive load manipulation did not affect the occurrence of STIGs, (c) STIGs occurred for groups varying in entitativity, (d) STIGs influenced perceivers' impression ratings of those groups, and (e) STIG-based group impressions generalized to new group members. These experiments provide the first evidence for STIGs, a process that may contribute to the formation of spontaneous group impressions. Implications for stereotype formation are discussed.

Keywords: spontaneous trait inferences, group perception, stereotype formation, entitativity

When we see another person assist a stranger by carrying some heavy packages, we not only construe the behavior itself to be *helpful* but also infer that the person is in fact a *helpful person*. We have moved from observed act to inferred disposition (Jones & Davis, 1965). In doing so, we have not only comprehended the

meaning of the behavior (*helpful*) but also have formed an impression of the actor (*helpful*). Why would we make this inference? Social interaction depends on the ability of each person to anticipate the behavior of the other in order to effectively coordinate behaviors between them. Understanding the dispositional qualities of others aids our ability to anticipate future behaviors and is therefore highly functional in adapting to a complex social environment.

Obviously these inferences can be the product of a conscious, deliberative process. In addition, however, there is now considerable evidence that *spontaneous trait inferences* (STIs) occur regularly, without intention, and perhaps even without awareness, as we process behavioral information from the social world (Carlston & Skowronski, 1994; McCarthy & Skowronski, 2011a; Todorov & Uleman, 2002, 2003; Uleman, 1999; Uleman, Blader, & Todorov, 2005; Uleman, Newman, & Moskowitz, 1996; Winter & Uleman, 1984). Those inferred attributes may be corrected and changed later in light of new information (e.g., Gilbert, Pelham, & Krull, 1988), but they are made quickly and spontaneously as information is initially processed.

The accumulated evidence supports the view that traits are spontaneously inferred during the encoding of behavioral information, that those traits become inferred properties of the actors

(above and beyond their ability to capture the meaning of their actions), and that this is the result of an inference about the actor and not simply an association based on contiguity (see Uleman, Saribay, & Gonzalez, 2008). The fact that evidence for these STIs has been obtained in research using several different paradigms highlights the robustness of this phenomenon.

Interestingly, *all* of the research on STIs has investigated this process as perceivers encode and comprehend behavioral information about *individual* stimulus persons. Yet many important aspects of social life occur in *groups*. Just as we perceive individuals engage in various actions, we also regularly perceive groups engage in behaviors. For example, consider the following group actions. "The sorority members took the children to the zoo." "The striking union members protested at the factory gates." Do we spontaneously make group-level dispositional inferences about these groups, parallel to what we do with individual target persons? Do we quickly and without intention infer that the sorority members are *kind*, that the demonstrators are *aggressive*? Surprisingly, there have been no studies investigating whether people make such inferences spontaneously as they process information about group behaviors. We propose that they do, and we refer to this process as *spontaneous trait inferences about groups* (STIGs).

Such inferences would seem to be of crucial importance in understanding social perception. In the first experiment demonstrating STIs (Winter & Uleman, 1984), participants were told that they were taking part in a memory study and that their task was to study the sentences carefully because they would be tested on them later. Despite this focus on memory, the results suggested that participants made trait inferences about the actors in those sentences. It was this possibility—that people would spontaneously infer traits from behaviors in a task having nothing to do with perceiving persons—that immediately captured the attention of researchers. It suggested that STIs could contribute to the emerging *impression* of an individual. In parallel manner, STIGs would lay the foundation for developing a group impression or *stereotype* of the target group. Therefore, we propose that perceivers form STIGs from groups' behaviors and that this process is a new mechanism by which stereotypes may form. Our research investigated this possibility.

Spontaneous Inferences in Group Contexts

To our knowledge, only two published articles (Crawford, Sherman, & Hamilton, 2002; Otten & Moskowitz, 2000) have investigated the role of spontaneous inferences in the development of group impressions. Otten and Moskowitz (2000) demonstrated a spontaneous in-group bias in the minimal group paradigm. They used a probe reaction-time procedure (Uleman, Hon, Roman, & Moskowitz, 1996), in which trait inference is revealed in slower responses to correctly indicate that a probe word (a trait implied by behavior) was not in a stimulus sentence. They found that response times were significantly longer when positive probe traits followed sentences describing in-group members (compared with out-group members) performing behaviors that implied those traits. These results demonstrate a spontaneously formed in-group bias.

Adapting the savings-in-relearning inference paradigm (Carlston & Skowronski, 1994), Crawford et al. (2002) studied how behavioral information about individual members of a group is integrated into a global group representation, and how, once formed, this impression is applied to other group members. Participants read about behaviors performed by members of two different groups, A and B, and all behaviors implied specific personality traits. Behaviors by all Group A members implied one of two traits (lazy or intelligent) and behaviors by all Group B members implied one of two other traits (aggressive or honest). In addition, the two groups were characterized in a way that made both of them appear to be high or low in entitativity, the degree to which an aggregate of individuals constitutes a group (Hamilton, Chen, & Way, 2011; Hamilton, Sherman, & Castelli, 2002; Hamilton, Sherman, & Rodgers, 2004). In a later phase, each group member was presented again, this time paired with a trait word rather than a behavior. In some cases, the trait was the one implied by the behavior originally performed by that target member (e.g., "lazy"); in other cases, it was a trait that was implied by the behavior of other members of the same group, but that did not match the behavior of this particular individual (e.g., "intelligent"). The key measure was the ease with which participants learned these member–trait pairings. The first case (an individual paired with a behavior-implying trait) was a *trait inference pairing*, as the trait matched the inference from that individual's previous behavior. The other case (an individual paired with a trait implied by a different group member's behavior) was referred to as a *trait transference pairing*. In transference pairings, the trait did not match the inference from that individual's previous behavior and therefore learning such pairs would be facilitated only if the traits inferred from the behaviors of some group members had been spontaneously transferred or generalized to all group members. Crawford et al. found that participants made STIs about the group members, regardless of the group being high or low in entitativity. Trait transference, however, occurred only for high-entitativity groups.

Crawford et al.'s (2002) findings are important for two reasons. First, their results documented an important role of spontaneous inferences in group impression formation. Traits spontaneously inferred about one person were transferred to other members of the same group, which would lay the groundwork for the formation of an overall group impression. Second, the spontaneous transference results (for high-entitativity groups) have important implications for stereotyping. Through such transference, the group members become interchangeable in the sense that the inferred attributes of any member of a highly entitative group can become associated with all members of that group. This, then, is a mechanism for spontaneous overgeneralization of traits to group members. Such overgeneralization is an important foundation for stereotyping (Allport, 1954).

Although these studies (Crawford et al., 2002; Otten & Moskowitz, 2000) were the first to examine the role of spontaneous inferences in group impression formation, their focus was on the implications of the actions of *individual* group members on the perceiver's overall impression of the group. Therefore, they do not address the question we posed earlier: Do perceivers make spontaneous inferences about groups from *group* behaviors (i.e., behaviors performed by the group as a unit)? We predict that people do make STIGs.

Theoretical Context: Implications of STIGs for Stereotype Formation

Although the parallel between STIs and STIGs may seem straightforward, the theoretical implications of STIGs are unique from STIs because of the difference in target. Specifically, the proposed STIG process has important theoretical implications for group perception and stereotype formation. It is useful to consider these ideas in a historical context.

For many years, it was assumed that stereotypic beliefs are formed as a consequence of first-hand intergroup experiences with group members, or are acquired second-hand through social learning and socialization (Brigham, 1971; Hamilton, 1976; Hamilton, Stroessner, & Driscoll, 1994). The initial conceptions formed may be enhanced when accompanied by a history of conflict between the groups, intergroup feelings of relative deprivation (Crosby, 1976; Runciman, 1966), or competition for scarce resources (Sherif, 1966). Moreover, these perceived differences may be sustained and perpetuated by cognitive biases (Hamilton & Sherman, 1994; Hamilton, Sherman, & Ruvolo, 1990) and by system-justifying beliefs (Jost & Banaji, 1994). Even when recognizing that stereotypes are gross overgeneralizations of actual group differences (Allport, 1954), it was nevertheless assumed that there was some "kernel of truth" on which those differences are based and exaggerated.

The necessity for this kernel of truth was challenged by research conducted in the 1970s, which introduced new cognitive and motivational mechanisms that could create perceptions of intergroup differences that were not necessarily based on actual differences. Research using both the minimal group paradigm (Tajfel, 1970; Tajfel, Billig, Bundy, & Flament, 1971) and the illusory correlation paradigm (Hamilton & Gifford, 1976) demonstrated differentiation between groups for which there was no informational basis. Participants in those studies were given information about groups that they used to make group judgments. In both cases, the results showed that biased processing generated perceptions of group differences that were not justified by the information provided.

Although the minimal group and illusory correlation paradigms were quite different from each other in many respects, they had one important element in common: The research established that intergroup differentiation and stereotype formation could emerge due to properties of information processing, in the absence of any actual differences between groups or history of intergroup conflict. This new development did not in any way challenge the important roles of intergroup conflict and social learning as elements on which stereotypes and prejudice are based. It did, however, show that these elements are not necessary precursors to perceiving intergroup differences and established that motivational and cognitive biases can themselves produce those same outcomes.

The idea that group impressions might form through STIGs extends this tradition and suggests a new cognitive process by which stereotypic concepts may form. The proposed process begins with observation of group behavior, which then moves spontaneously "from acts to dispositions," in this case an inferred disposition characterizing the group. This inference occurs without conscious intention as a part of behavior encoding and comprehension. This concept is represented in memory and becomes the initial group impression. Once established in memory, that con-

ception can be embellished, sustained, and perpetuated by the same cognitive and motivational forces that shape, elaborate, and maintain other cognitive representations. Thus, the initial STIGs—spontaneously inferred group attributes—may lay the foundation of a newly formed stereotype.

As with the research that expanded on traditional conceptions of stereotypes in the 1970s, the notion that stereotypic concepts can emanate spontaneously from inferences formed during the encoding and comprehension of group behaviors would represent a new process by which these group conceptions can form. In this case, there are some important aspects of the underlying process that distinguish it from previous accounts of stereotype formation. First, research on spontaneous inferences has shown that they occur without intention and the perceiver may not even be aware that these inferences are being made. Thus, a STIG-based group impression would constitute the foundation for stereotype formation that occurs without intention. This feature separates it from most other accounts of stereotype formation. Second, in both minimal group and illusory correlation research, the participants are fully aware that their task involves making judgments about groups. In contrast, most spontaneous inference studies (including all reported here) go to great lengths to avoid participants having that knowledge. The studies are typically introduced as experiments investigating memory for verbal information, with no mention of using the stimulus information to form impressions or make social judgments. Thus, the implication is that STIG-based stereotypic concepts could be formed under conditions in which the perceiver is oblivious to the social perception implications of the material presented. Third, some have argued that forming stereotype concepts relies on a contrast between two (or more) groups (e.g., in-group vs. out-group, or two different target groups). For example, minimal group and illusory correlation studies present information about two (or more) groups, suggesting an intergroup context and perhaps inducing intergroup comparison processes. In contrast, in the paradigm used in the present research, the instructions and the task given to participants did not include any suggestion of group contrast or comparison processes.

Thus, the possibility that perceivers make STIGs based on group behaviors has implications for, and suggests new ideas about, the nature of stereotyping and group perception. These questions and implications could not have been generated from the STI literature.

Before we can fully explore these possibilities and their ramifications, it is first necessary to establish that perceivers do in fact make spontaneous inferences about groups based on group behavior. The goal of the present research is to present evidence that STIGs do occur spontaneously and to explore some of their properties and parameters.

We present five experiments designed to answer several questions about STIGs. Do people make spontaneous inferences about groups (STIGs) as readily as they do about individuals (STIs), or is one type of inference more likely to occur, or to occur more strongly, than the other? To what extent do STIGs manifest the characteristics of a spontaneous process, occurring even when cognitive resources are limited? Are perceivers more inclined to make STIGs about some types of groups than about other types of groups? To what extent do STIGs guide the impressions formed of these groups? Can a group impression based on STIGs influence perceptions of a newly encountered group member, just as group

stereotypes generalize to group members? Our research sought to address these questions.

Overview of the Current Research

We propose that perceivers make STIGs when groups engage in trait-implying behaviors. To test our hypotheses, we used the *false recognition paradigm* (Todorov & Uleman, 2002), which occurs in two phases: a learning phase and a recognition phase. In the learning phase, participants are shown a series of stimuli, each consisting of a picture of a stimulus person and a sentence describing a behavior performed by that individual. Participants are asked to memorize the information for a memory test to occur later. In the recognition phase, participants are given a recognition test in which each item consists of a face paired with a trait word. The faces are the same ones presented earlier. The participants' task is to indicate whether or not the trait word was presented in the behavior-descriptive sentence that had described the person shown in the photo. On the critical trials, the trait word *did not occur* in the sentence previously paired with the face, but it is a trait *implied by* that behavior (a match trial). On other trials, the trait is one that was implied by a different person's behavior (a mismatch trial).

The logic of the method is as follows. If the trait presented in the recognition phase is strongly implied by that person's behavior, and hence could have been spontaneously inferred by the participant while encoding the stimulus information, then it should be more difficult to make that "No" judgment. Saying "Yes" would constitute a false recognition. The signature evidence that spontaneous trait inference has occurred is a significantly greater number of false recognitions on match than on mismatch trials.

In fact, Todorov and Uleman (2002, 2003, 2004; Goren & Todorov, 2009) demonstrated exactly that effect in a series of experiments. Their results reflect the fact that the implied trait was spontaneously inferred by the participant during the learning phase. Furthermore, their studies documented that STIs occurred during initial encoding and were uniquely associated with the specific target person. Our research adapted this paradigm to investigate STIs about group targets.

In Experiment 1, we determined whether perceivers have the same propensity to make STIGs as STIs. In Experiment 2, we tested the efficiency of the STIG process by manipulating perceivers' cognitive load at encoding. In Experiment 3, we investigated whether groups' level of entitativity affects STIG formation. In Experiment 4, we tested for a downstream consequence of making STIGs, namely, whether these inferences carry over and influence participants' ratings of the groups. In Experiment 5, we tested whether a group impression based on STIGs generalizes to perceptions of a new group member.

Experiment 1: Comparing STIs and STIGs

The purposes of Experiment 1 were (a) to test our hypothesis that perceivers make STIGs as they encode information about group actions, and (b) to compare the frequency and strength of these group inferences to STIs drawn about individuals.

We predicted that we would find evidence that participants make both STIs and STIGs. Predictions regarding the relative frequency and strength of STIs and STIGs present an interesting

challenge, in that there are good evidentiary grounds to support conflicting predictions. On the one hand, as noted earlier, there is considerable evidence arguing that STIs are made quickly and efficiently, as a part of the process of comprehending the behavioral information. They occur when presentation of information is fast paced, and when participants are simultaneously performing a second task. Also, these inferences are not dependent on recall of the behavioral items, and they are specifically linked to the actor (Carlston & Skowronski, 1994; Todorov & Uleman, 2002, 2003, 2004). This evidence suggests that spontaneous inferences occur as part of behavior comprehension and would occur for both individual and group targets.

On the other hand, evidence from other literatures suggests viable reasons that STIs would be more prevalent than STIGs. First, people typically interact more frequently and consistently with individuals than with groups. They therefore observe individual behavior more often than they observe group behavior, providing more opportunities to make STIs than STIGs. If they make inferences about individuals with greater frequency, it could, in turn, result in STIs becoming a more routinized process than that of making STIGs. Second, Hamilton and Sherman (1996) reviewed a considerable amount of evidence showing differences in the outcomes of information processing and impressions formed of individual and group targets, even when targets are presented the same information and task instructions. Hamilton and Sherman proposed that there are some fundamental differences in the way information about individual versus group targets is processed. Specifically, they argued that perceivers assume greater unity and consistency in individual than group targets, a difference that might lead to more frequent inferences about individual than about group targets. Since Hamilton and Sherman's analysis was published, a large literature has accumulated documenting differences in processing information about persons versus groups (see Hamilton, Sherman, Way, & Percy, 2014). Both of these considerations provide a basis for expecting that evidence for STIs (i.e., more false recognitions on match than mismatch trials) would be more prevalent than for STIGs.

Given these alternative bases for anticipating different outcomes, we did not make a prediction regarding the similarity or difference in STI versus STIG results. Experiment 1 provides the first opportunity to obtain evidence testing these competing possibilities.

Development and Pretesting of Stimulus Materials

Generating sentences. To compare spontaneous inferences about persons and groups, the same behaviors must be used in both individual and group target conditions, and therefore must be equally appropriate as individual actions and as group actions. In addition, evaluation of the behaviors must be comparable when applied to individuals and to groups. Finally, each behavior must imply a particular trait. We developed an extensive list of behavior-descriptive items and pretested them with regard to these criteria.

To develop stimulus sentences, we began with a list of 31 different traits and, for each one, generated two to four sentences describing behaviors reflecting that trait. This process produced a list of 89 potential stimulus sentences. We then wrote two versions of each sentence, one describing a person and the other describing

a group performing the behavior (e.g., "This individual makes donations yearly to Hospice"; "This group makes donations yearly to Hospice").

Pretesting. Sixty-nine students in an upper-division psychology course participated in the pretesting of these sentences. Participants were given either the individual target or group target version of the stimulus sentences. For each sentence, participants were asked to rate how desirable the behavior was on a 9-point Likert scale (1 = *very negative*, 9 = *very positive*). Next they were asked whether or not the behavior could be performed by a person (in the individual condition) or by a group (in the group condition; 1 = "yes," 2 = "no"). Finally, participants were asked to list the first three trait attributes that came to mind when reading each sentence.

Pretest analyses and results. Our first objective was to find sentences that did not differ by target condition (individual or group) or valence. To determine the appropriateness of behaviors for describing individuals and groups, we examined participants' yes–no responses. All behaviors that had more than one response indicating that participants could not imagine the behavior being performed by a person or by a group were discarded. Thus, our list was restricted to behaviors that all (or all but one) participants regarded as behaviors that could be enacted by both a person and a group.

Next we examined the valence ratings of the behaviors. The average valence ratings of the behaviors when performed by person or group were nearly identical: $M = 4.98$, $SD = 2.04$ for individuals, $M = 4.90$, $SD = 2.35$ for groups. In addition, the correlation between ratings of the person and group versions of the behaviors was calculated. Different samples of participants rated the individual and the group versions of the sentences. Therefore, we determined the mean rating of each item and correlated those mean values. The resulting correlation was very high, $r(89) = .97$, $p < .001$. These data provided assurance that the valence of the items did not differ when performed by an individual or a group.

Finally, we determined the extent of agreement among participants in the traits they listed as coming to mind when they read a sentence. We made a list of all the traits participants listed for each sentence and determined the percentage of participants that listed each trait or its synonyms (e.g., intelligent, smart, wise) for that behavior. We selected items with the highest consensus.

Based on these three criteria, the final list consisted of 24 stimulus items that were uniformly considered equally applicable to person or group, were equated for valence, and had between 35% and 60% consensus that the behavior implied the trait word, based on participants' freely generated associations. Some examples of behaviors meeting these criteria were:

Worked hard to finish an assignment before a deadline. (ambitious)

Provided food and clothing for the flood victims. (kind)

Heckled a woman speaking on human rights. (rude)

Twenty-four critical sentences were used. Each of these sentences implied a different trait (half implied positive traits, half implied negative traits). Twelve additional sentences were developed as filler sentences. These sentences described a behavior but also contained the trait in the sentence (e.g., "The individual was so dishonest that he claimed credit for someone else's idea"). Half of the filler sentences contained a positive trait and half contained a negative trait.

Method

Participants. Forty-one undergraduate students completed the study for either research credit or $5 reimbursement. Of these, 24 self-identified as White, seven as Asian American, and four as Black. Six participants did not report their ethnicity. Participant gender was not recorded for this experiment.

Stimulus photos. One hundred forty-four neutral male faces were chosen from several databases. All faces had neutral expressions and were presented in front of a white background and in gray scale. In the individual target condition, a photo of one person was shown in each stimulus frame. In the group target condition, a group was created by displaying four photos of individual faces onto one frame. All descriptive sentences were presented below the photos.

Design. The experiment had a 2 (target type: individual or group) × 3 (trial type: match, mismatch, and filler) mixed design. Target type was a between-groups factor, whereas the trial type factor was within subjects.[1] We manipulated the target type by showing participants photos of either one or four male individuals with each stimulus sentence. All participants saw each type of sentence in the recognition phase (see Procedure section). The dependent variable was the total number of false recognitions, or "old" responses, given in the recognition phase.

Procedure. Participants entered the lab and were seated in individual cubicles with a computer. They were told that they would be participating in a study investigating people's ability to memorize and remember information. They were then taken through the two phases of the experiment.

In the learning phase, participants were told that their task was to try to memorize the information presented. Participants were then shown 36 photos (either one or four male faces) paired with sentences (either trait-implying or filler), one at a time and in random order. In the group condition, the computer randomly selected and combined the photos into groups of four, so each participant viewed groups of slightly different compositions. The photo–sentence pairings were also randomly selected by the computer and hence differed for each participant. Participants saw each photo–sentence pair for 10 s, with a 2-s gap between stimulus presentations.

In the recognition phase, participants were told that their memory for the stimuli would be tested. They were presented with the 36 photos from the learning task, one at a time and in random order. Each photo was accompanied by a trait word (the probe word). The participants' task was to indicate, as quickly as possible, whether or not they had seen the probe word in the sentence about the person or group shown in the photo in the learning phase.

---

[1] All five experiments reported here included these three types of stimulus sentences. The filler sentences (which included the implied trait word) were not of theoretical interest and were included simply to provide opportunities for participants to correctly respond "Yes" to the trait probe question (i.e., they were not false recognitions). Therefore, in all experiments, data analyses did not include the filler sentences.

Participants were told to press the *old* key ("M") if they believed they had seen the word in the sentence associated with that particular photo in the learning phase, and the *new* key ("X") if they believed that they had not seen the word in the sentence associated with that photo in the learning phase. Of the 24 critical photos, 12 were paired with the trait implied by the sentence that had described that person or group (match trials), and 12 were paired with a trait implied by a sentence that had described another person or group (mismatch trials). The 12 filler sentences were correctly paired with the traits that they had contained in the learning phase. After responding to the 36 trials, participants were debriefed and thanked.

Results

We hypothesized that participants would form STIs about individual targets, replicating the results of many previous studies. We also expected to obtain evidence that participants also formed spontaneous inferences about groups, although as noted earlier, the relative strength of STIs and STIGs is an important question about which past findings suggest alternate possibilities. Evidence for STI and STIG formation is indicated by more false recognitions (i.e., more frequently responding "old" to the probe word in the recognition phase) for match trials than for mismatch trials.

We first calculated the number of false recognitions each participant made for match and mismatch trials in the recognition phase. We then tested for STI and STIG formation by conducting a 2 (target type: individual or group) $\times$ 2 (trial type: match and mismatch) mixed model ANOVA on participants' recognition rates. There was a significant main effect of trial type, $F(1, 39) = 50.06$, $p < .001$, $\eta_p^2 = .56$. As predicted, participants made more false recognitions on match trials ($M = 6.17$, $SD = 2.38$) than on mismatch trials ($M = 3.37$, $SD = 2.22$), and this occurred for both individual and group targets. Therefore, evidence for both STIs and STIGs was obtained (see Figure 1). The main effect for target type approached, but did not achieve, significance, $F(1, 39) = 2.52$, $p = .12$, $\eta_p^2 = .06$, with slightly higher false recognitions for individual than for group targets. There was no Target Type $\times$ Trial Type interaction, $F(1, 39) = 0.01$, $p = .93$, $\eta_p^2 = .00$, indicating that participants made STIs and STIGs with the same frequency. Two paired samples t tests confirmed that participants

in the individual condition, $t(21) = -6.02$, $p < .001$, and participants in the group condition, $t(18) = -4.26$, $p < .001$, made more false recognitions on the match trials ($M_i = 6.59$, $SD_i = 2.15$; $M_g = 5.68$, $SD_g = 2.58$) than on the mismatch trials ($M_i = 3.82$, $SD_i = 2.32$; $M_g = 2.84$, $SD_g = 2.04$).

Discussion

The results of Experiment 1 are important for several reasons. First, the fact that participants made STIs for individual targets replicates the results of many past STI studies. Second, in the group target condition, participants also made significantly more false recognitions in the match than in the mismatch condition. This finding is important in that it provides the first documentation that people make STIGs as they process information about group behavior. Third, the magnitude of these spontaneous inference effects did not differ for the individual and group target conditions. This result is consistent with the view that spontaneous inferences are made early in the encoding process. In extending results showing spontaneous inferences from individual to group targets, our findings provide further evidence that these spontaneous processes are quite robust.

It is important to note that our participants were given memory instructions and told to remember the information they would read, with no suggestion that they should form impressions of the target persons or groups. Therefore the inferences they have drawn were not the result of intentional impression processes. Rather, and paralleling the STIs made about individual targets, they occurred without clear intention during a memory task in which no mention was made of group perception or impressions. Yet these inferences can form the beginnings of a group impression, which, if it were to develop further over time, could become a stereotypic conception of the group.

Ever since Winter and Uleman's (1984) classic article, evidence has been accumulating that people routinely make STIs as they comprehend the behaviors they learn about persons (see Uleman et al., 2008). It is plausible that similar processes would be engaged when perceivers learn about the behavior of groups. Although this may seem like a straightforward step, it is not necessarily a given. As described earlier, Hamilton and Sherman (1996) argued that cognitive processing may be engaged to differing degrees as a function of whether the target is a person or group, due to differing inherent assumptions perceivers make about these targets. The present finding further enlightens that process. That is, when presented behavioral information, perceivers comprehend that behavior and, in doing so, make spontaneous inferences about the actor. This happens for both individual and group targets: The strong main effect difference between match and mismatch trial types in the total absence of an interaction of target type with trial type suggests that the same processes are engaged in both cases. Therefore, the individual and group target conditions of Experiment 1 manifested the same results. However, having made such inferences, the *use* of that new knowledge may differ for individual and group targets as a function of the (perceived or assumed) unity of the target. Differences in perceived unity may engage other processes to different degrees, which may in turn generate different outcomes for individual and group targets (Hamilton & Sherman, 1996; Hamilton et al., 2014). Research further investigating
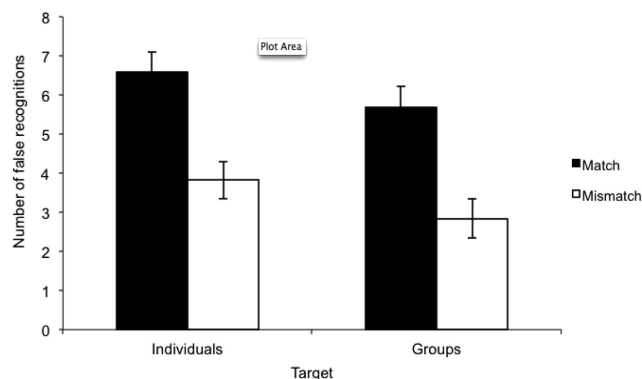

Figure 1. Mean number of false recognitions for individual and group targets on match and mismatch trials (Experiment 1). Error bars represent standard error.

conditions that may influence the relative strengths of STIs and STIGs would be a valuable direction for future work.

Experiment 2: STIGs and Cognitive Resources

One feature that differentiates deliberative, systematic processes from spontaneous, highly routinized processes is their susceptibility to interference and their need for cognitive resources. Highly deliberative processing is resource-consuming and therefore can be disrupted when additional task demands are placed on the cognitive processing system. In contrast, spontaneous processes are well-developed routines that are highly efficient, that is, they can occur without conscious thought and are not disrupted by other simultaneous tasks. Based on this difference, Experiment 2 was designed to provide an additional test of the spontaneity of STIGs.

The efficiency of spontaneous inferences can be tested by using a cognitive load manipulation, comparing the extent to which inferences are formed under high- versus low-load conditions. In the false recognition paradigm used in our studies, the key question is whether the difference in frequency of false recognitions between match and mismatch trials—the indication that spontaneous inferences were made—is eliminated (or significantly diminished) by cognitive load.

In past research testing the efficiency of STIs (using individual targets), several strategies have been used to make the task more taxing and thereby create conditions that might interfere with the ongoing processes engaged in encoding information. Some studies, using different paradigms and a variety of load manipulations, have found little, if any, effect of cognitive load manipulations on STI formation. For example, one study compared a self-paced condition with a faster paced presentation condition (Todorov & Uleman, 2003). Although the faster pace lowered overall accuracy, it did not affect the difference in false recognitions between implied and nonimplied traits, suggesting that traits were linked to actors very quickly. Increasing the number of trials (Todorov & Uleman, 2002) and including a week delay between the learning and recognition phases (Todorov & Uleman, 2004) did not eliminate the formation of STIs. Having participants count the number of nouns in the stimulus sentences while performing the task reduced the magnitude of the STI effect, but again the difference between match and mismatch conditions was still significant in both load and no load conditions (Todorov & Uleman, 2003).

Perhaps the most frequently used method of manipulating cognitive load is to give participants in the high-load condition a multidigit number that they must retain while they read the behavior-descriptive sentences; their performance is then compared with that of participants in no-load or low-load conditions. Some of these studies have found that cognitive load manipulated in this way had no effect on STI formation (Crawford, Skowronski, Stiff, & Scherer, 2007, Experiment 3; Todd, Molden, Ham, & Vonk, 2011; Winter, Uleman, & Cunniff, 1985; but also see Uleman, Newman, & Winter, 1992). Others have found that although the load diminished the magnitude of the effect, specific comparisons between match and mismatch conditions were still significant under load as well as no load (Todorov & Uleman, 2003; Wells, Skowronski, Crawford, Scherer, & Carlston, 2011, Experiment 1). The results of all of these studies suggest that STI formation (as indicated by the difference in false recognitions between match and mismatch conditions) is not influenced by cognitive load manipulations. In contrast, a few studies have found that a cognitive load can eliminate the STI effect (e.g., suggesting that STIs do require cognitive resources; Crawford et al., 2007; Uleman et al., 1992). Thus, although the results of these studies are not entirely consistent, the preponderance of evidence is that cognitive load manipulations have had little or no effect on STIs, suggesting that they are a highly efficient process.

However, this issue has never been studied in spontaneous inferences about groups. Experiment 2 was designed to provide the first test of the efficiency of STIGs. Participants' spontaneous inferences about groups were assessed after they were asked to complete a cognitively demanding task or a nondemanding task. Based on the majority of previous research examining the efficiency of STIs, we hypothesized that participants would form STIGs in both high- and low-load conditions.

Method

Participants. Sixty-nine undergraduates (50 females) participated in exchange for $5. The mean age was 20.33 years ($SD =$ 4.79). Thirty participants self-identified as White, 19 as Asian, 13 as Latino, two as Black, one as American Indian, and four participants did not specify their race.

Materials and procedure. The procedure mirrored that of Experiment 1, with a few exceptions. Upon arriving to the lab, participants were randomly assigned to either the high-load ($n =$ 34) or low-load ($n =$ 35) condition. Participants were told that the experiment was about memory and how we remember both verbal and numerical information. Participants were seated at computers in individual cubicles. All participants first completed the learning phase, in which they were presented with 36 separate sets of four photographs depicting Caucasian male faces with neutral expressions. Participants were told that each set of four faces represented a different group, and each group was accompanied by a sentence describing a behavior performed by the group. Of the 36 trials, 24 were considered "critical" trials and contained behaviors that implied trait characteristics. Twelve control trials presented behaviors in which the trait was made explicit (trait explicit "fillers") and are not included in our analyses. Each group appeared for a period of 8 s before automatically advancing to the next group.

Cognitive load was manipulated during the learning phase. Before the 36 trials began, participants in the high-load condition were presented with a randomly generated seven-digit number and were asked to remember the number to the best of their ability while they read about a set of six groups and their behaviors. After seeing and reading about six groups, participants were asked to enter the seven-digit number that they had seen earlier. Participants were then presented with a new seven-digit number and were again asked to remember the number while reading about six new groups. This pattern recurred every six trials. Across the 36 trials, participants were instructed to remember and correctly recall six different seven-digit numbers as they advanced through the learning phase. Participants in the low-load condition saw the same sets of faces, in random order, but were instead asked to remember a series of two-digit numbers.

The recognition phase was the same as that of Experiment 1. All participants were presented, in random order, the same 36 sets of four faces. For each trial, participants were asked to indicate whether a given trait word had appeared in the sentence about that

group's behavior. On 12 of the trials, this probe word was the trait implied by the behavior of that group (match trials); on 12 trials, the probe word was a trait implied by the behavior of a different group (mismatch trials); and on 12 trials, the probe word was a trait that had actually appeared in the sentence about that group's behavior (fillers trials).

At the end of the recognition phase, all participants were asked to rate the difficulty of remembering the seven-digit or two-digit numbers they had to retain while reading the group behaviors. These ratings were made on a 7-point Likert scale (1 = *extremely easy*, 7 = *extremely difficult*). Participants then provided demographic information and were fully debriefed upon study completion.

Design. We employed a 2 (load condition: high, low) × 2 (trial type: match, mismatch) mixed model design, in which the latter factor was within subjects. As in Experiment 1, the dependent variable was the number of false recognitions during the recognition phase.

Results

Cognitive load manipulation check. Participants in the high-cognitive-load condition reported greater difficulty remembering the seven-digit numbers (M = 3.15, SD = 1.73) than did participants who were asked to remember two-digit numbers (M = 1.71, SD = .96), t(67) = 4.28, p < .001. Thus, the cognitive load manipulation was effective at rendering a task that participants considered more demanding in the high-load condition than in the low-load condition.

A second way to assess the relative degree of cognitive load was to compute participants' accuracy in recalling the numbers they were asked to remember during the learning phase. We reasoned that participants in the high-load condition should experience greater difficulty in remembering the numbers, and thus would evidence more inaccuracy in recalling their seven-digit numbers. Any deviation from the correct number was considered an "error," and the number of errors could range from zero to six. Participants in the high-load condition made significantly more errors (M = 1.68, SD = 1.51) than did participants in the low-load condition (M = .71, SD = .99), t(67) = 3.14, p = .003, suggesting that it was more difficult for high-load-condition participants to remember seven digits than it was for our low-load participants to remember two digits.

Thus, both ways of assessing the cognitive load manipulation testified to its effectiveness.

Influence of cognitive load on STIGs. The primary question of interest in this study was the degree to which cognitive load would interfere with encoding of information and subsequent formation of STIGs. Specifically, we tested the prediction that participants in both high-load and low-load conditions would make more false recognitions on match than on mismatch trials. If the load manipulation did not alter this difference, this result would provide evidence for the efficiency of the process underlying STIGs.

To test our hypothesis, we first summed each participant's number of false recognitions separately for both the match and mismatch trials. As in Experiment 1, a false recognition was operationalized as the erroneous belief that the trait probe word presented in the recognition phase had appeared in the sentence

about that same group when it was presented in the learning phase. For match and mismatch trials, the correct answer was always "No."

The 2 × 2 mixed model ANOVA replicated the significant main effect of trial type seen in Experiment 1, F(1, 67) = 41.26, p < .001, η_p^2 = .39. Participants made significantly more false recognitions on the match trials (M = 5.12, SD = 2.97) than on the mismatch trials (M = 3.51, SD = 2.56), t(68) = 6.21, p < .001. Overall false recognitions did not significantly differ between subjects in the high-load and low-load conditions, F(1, 67) = 2.15, p = .15, η_p^2 = .03. There was, however, a significant interaction of load condition and trial type, F(1, 67) = 4.94, p = .03, η_p^2 = .07 (see Figure 2). This interaction is primarily due to the high number of false recognitions on match trials in the high-load condition (M = 5.85, SD = 3.23) compared with the low-load condition (M = 4.40, SD = 2.53), t(76) = 2.08, p = .04). Of theoretical importance for our purposes, post hoc analyses showed that the number of false recognitions was significantly greater on match than on mismatch trials in both the high-load (M_{match} = 5.83, SD = 3.23, versus $M_{mismatch}$ = 3.68, SD = 2.73, p < .001, t[67] = 2.18, p < .001) and low-load (M_{match} = 4.40, SD = 2.53, versus $M_{mismatch}$ = 3.34, SD = 2.41, p < .001, t[67] = 1.06, p < .01]) conditions.

Discussion

The results of Experiment 2 provide further evidence for the interpretation that STIGs occur as a highly spontaneous process by which traits of groups are inferred from their actions. The difference in false recognitions between match and mismatch trials was not diminished by cognitive load.

One unexpected finding was the very high number of false recognitions in the high-load/match condition, resulting in a greater difference in match versus mismatch trials under high load than low load. We have no explanation for this result, other than to speculate that people may rely on highly routinized processes (e.g., STIGs) under high load, whereas people under low load, with fewer cognitive demands, may have encoded more irrelevant information (e.g., targets' attractiveness, idiosyncratic features) that then diluted the influence of STIGs on false recognitions in the memory phase. We believe this outcome first needs to be replicated and, if reliable, to be considered in future research.

Previous research, investigating properties of STIs based on behaviors of individual targets, has generally (with some exceptions) shown that STIs occur even when cognitive resources were
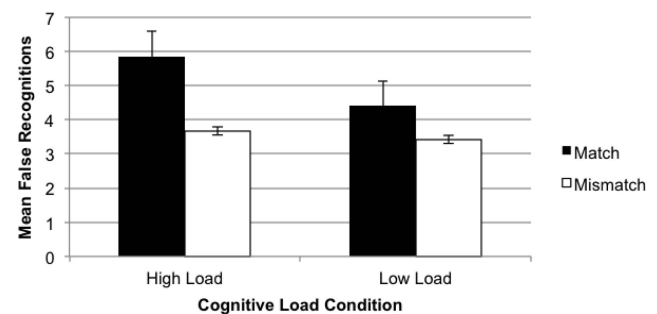

Figure 2. Mean number of false recognitions by load condition (Experiment 2). Error bars represent standard error.

constrained due to cognitive load manipulations. Our results, extending that work to spontaneous inferences about groups, are consistent with those findings. Like STIs, STIGs appear to reflect a highly efficient process that is not substantially disrupted by simultaneously performing a second cognitive task. Specifically, the difference between match and mismatch trials was significant in both load and no-load conditions. Again, the STIG effect appears to be robust.

Experiment 3: STIGs and Perceived Group Entitativity

The finding that people spontaneously infer group characteristics as they encode group behavior raises new questions and provides new opportunities for investigation. When do STIGs occur and when do they not occur? Are they more likely to occur from behavioral information about some groups, or types of groups, than others? Experiment 3 was designed to determine whether one central property of groups, *entitativity*, influences STIG formation.

During the last 15 years, there has been a great deal of research investigating the perception of entitativity (or perceived "groupness") in groups (Hamilton et al., 2002; Sherman, Hamilton, & Lewis, 1999; Yzerbyt, Judd, & Corneille, 2004). Both the antecedent conditions that lead to the perception of group entitativity and the consequences that follow from such perceptions have been studied (see Hamilton et al., 2011, for a review). Hamilton and Sherman (1996) proposed that people process information about high-entitativity groups in the same way they process information about individual persons, that is, assuming consistency across time and inferring underlying attributes to a degree not manifested for low-entitativity targets. This proposition was supported by subsequent research (Hamilton et al., 2014; McConnell, Sherman, & Hamilton, 1994, 1997; Susskind, Maurer, Thakkar, Hamilton, & Sherman, 1999). In light of these findings, one plausible hypothesis would be that people would be more likely to make STIGs about groups that are high, compared with low, in entitativity.

However, another line of reasoning, based in part on past STI research, suggests an alternative hypothesis. STIs are made as a part of comprehending behavior. They occur spontaneously, without intention; they simply happen (Uleman et al., 2005, 2008). If that is the case, then STIGs may occur as an inherent aspect of processing behavioral information, regardless of the properties of the target group. If so, then STIGs may occur routinely for groups of any kind, whether high or low in entitativity.

Experiment 3 tested the viability of these alternative hypotheses.

Method

Participants. Fifty-one undergraduate participants (M_{age} = 19.16 years, SD = 1.36; 36 female) took part in the study for course credit. There were 33 Whites, nine Asian Americans, and nine Latinos.

Design. The experiment was a 2 (entitativity: high or low) × 2 (trial type: match, mismatch) mixed design, with the latter factor being within subjects. The dependent variable was the total number of false recognitions given in the recognition phase.

Procedure. All participants learned about groups of persons and read descriptions of behaviors they had performed. The same group photos, sentences, and computer program used in the previous studies were also used in Experiment 3.

Experimental sessions were run with one to six participants per session in a computer lab. Participants were told that they would be engaged in a study on memory. Prior to the learning phase, participants received either the high- or low-entitativity induction about the groups they would be learning about (see below). After the learning phase, participants completed the recognition phase. Upon completion, an entitativity manipulation check was administered. Participants were then debriefed and thanked.

Entitativity manipulation. Most of the participants either had the experience of living in the campus dormitories or were currently living in the dormitories. Although first-year room assignments were generally made by the administration, dorm arrangements after the first year of college were typically made by students themselves. The widespread perception among students was that people living in the same suite are close-knit friends who share similarities and do a lot together, whereas such assumptions would not be made about others living in separate rooms on different floors in the same large dormitory building. Our manipulation of entitativity took advantage of these perceptions.

Participants were randomly assigned to either the high- or the low-entitativity condition. At the beginning of the experiment, after the initial instructions explained that participants would read about groups of persons and a behavior the group had performed, participants in the high-entitativity condition were given the following information:

> In the first part, the learning phase, you will be shown information about different groups of people living in a dormitory at another university. Each group consists of dormitory suitemates. In each case, you will see a set of four pictures of the men who are all part of the same dorm suite. You will also see a sentence describing the group. Because they live in the same suite, all of the members of the group know each other well. They are similar to each other, they share similar interests and goals, and they spend a lot of time together. Each set of four men are from a different suite. Your task is to look at the suitemates and to read and remember the sentences describing them.

Participants in the low-entitativity condition were given the following information.

> In the first part, the learning phase, you will be shown information about different groups of people living in a dormitory at another university. In each case, you will see a set of four pictures of men who are all residents in the same dormitory, but they live on different floors of the dormitory. You will also see a sentence describing the group. Because they live on different floors, these members see each other occasionally but do not know each other well. They are only moderately similar to each other, do not always share the same interests and goals, and sometimes see each other in passing or at dormitory functions. Each set of four men are from a different dormitory. Your task is to look at the dormitory residents and to read and remember the sentences describing them.

We assessed the success of this manipulation after the recognition phase of the experiment. Participants were asked to rate the collection of groups they had seen on scales assessing the perceived similarity, cohesiveness, feelings of inclusion, feelings of importance of the group, and unity of the target groups (α = .81). Each participant's ratings on these scales were averaged to form an index that was used as a manipulation check.

Results

Manipulation check.   To determine whether the entitativity manipulation had influenced participants' perceptions of the groups, we conducted an independent samples t test comparing the high- and low-entitativity conditions on responses to the manipulation check. Participants did in fact perceive groups to be more entitative in the high-entitativity condition ($M = 5.63$, $SD = 1.27$) than in the low-entitativity condition ($M = 4.21$, $SD = 1.05$), $t(49) = -4.35$, $p < .001$.

Influence of entitativity on STIGs.   We had recognized two possible outcomes of the entitativity manipulation, with differing theoretical implications. Past research has documented that differences in perceived group entitativity can influence numerous downstream outcomes. In particular, if high-entitativity groups are perceived as more like individual targets (Hamilton & Sherman, 1996; Hamilton et al., 2014), then participants should make more false recognitions for high- than for low-entitativity groups. Alternatively, if STIGs, like STIs, are made spontaneously as a part of comprehending behavior, they may occur for all groups, regardless of entitativity. In that case, the entitativity manipulation would not affect the difference in false recognitions between match and mismatch trials.

To test these alternative possibilities, we conducted a 2 (entitativity: high or low) × 2 (trial type: match and mismatch) mixed model ANOVA on the number of false recognitions made in the recognition phase. Results are shown in Figure 3. There was a significant main effect of trial type, $F(1, 49) = 83.71$, $p < .001$, $\eta_p^2 = .63$. Participants made more false recognitions on the match trials ($M = 4.80$, $SD = 2.66$) than the mismatch ones ($M = 1.67$, $SD = 1.95$). This result documents that participants did in fact make STIGs, as predicted. The main effect of the entitativity manipulation was not significant, $F(1, 49) = 0.62$, $p > .41$, $\eta_p^2 = .01$. Importantly, the effect of trial type was not moderated by entitativity condition, $F(1, 49) = 0.543$, $p > .47$, $\eta_p^2 = .01$. As in the previous experiments, two paired samples t tests confirmed that participants made more false recognitions on match trials ($M_{low} = 4.44$, $SD_{low} = 2.76$; $M_{high} = 5.15$, $SD_{high} = 2.57$) than on mismatch trials ($M_{low} = 1.56$, $SD_{low} = 1.96$; $M_{high} = 1.77$, $SD_{high} = 1.97$) in both the low-entitativity condition, $t(24) = 6.77$, $p < .001$, and the high-entitativity condition, $t(25) = 6.36$, $p <$

Figure 3.   Mean number of false recognitions for low-entitativity and high-entitativity groups on match and mismatch trials (Experiment 3). Error bars represent standard error.

.001. Thus, participants made STIGs to the same extent as they processed behavioral information about high- and low-entitativity groups.

Discussion

Having obtained evidence in Experiments 1 and 2 that STIGs occur when people encode information about groups' actions, the purpose of Experiment 3 was to learn more about the group properties that increase or decrease the likelihood that STIGs will occur. It seemed quite plausible that perceivers would be more likely to make spontaneous inferences about some groups than about others, and the accumulated literature on entitativity suggested this variable as a likely candidate for identifying groups for which STIGs would be more or less likely. Specifically, if perceivers process information about highly entitative groups in a manner similar to the way they process information about individual persons (Hamilton & Sherman, 1996; McConnell et al., 1994, 1997), then it is reasonable to assume that STIGs would be more likely to be made for high- than for low-entitativity groups. Our results, however, do not lend support to this expectation. Our manipulation of entitativity was successful (as evidenced on the manipulation check), but it did not produce differences in the frequency of STIGs.[2]

It is possible that, while producing significance on the manipulation check measure, the manipulation may nevertheless not have been sufficiently strong. Although four White males who simply live in the same dormitory are perceived as a less entitative group than four males who share a suite, the fact that they are all living in a dormitory may be enough to convey some level of perceived groupness (college students of similar age, pursuing a college degree, common goals, etc.). Perhaps a manipulation using groups even lower in entitativity (for example, what Lickel et al., 2000, called *nongroups* or *loose associations*) would provide a better test.

There are, of course, many factors and properties of groups that could influence the likelihood and ease of making STIGs. Given the recent literature on group perception, entitativity seemed like a particularly viable candidate for such a variable. Further research examining the antecedents of STIGs might use alternative manipulations of entitativity as well as explore the effects of other group properties (e.g., size, group composition) on STIGs.

There was, however, another equally plausible possibility. Research on STIs has repeatedly documented that people make inferences about individual target persons spontaneously and without intention. They occur as a part of comprehending behavior, and in doing so the inference spontaneously moves from act to disposition. If this were also true of STIGs, then they would occur in processing behavioral information about all groups. The fact that the entitativity manipulation did not significantly interact with
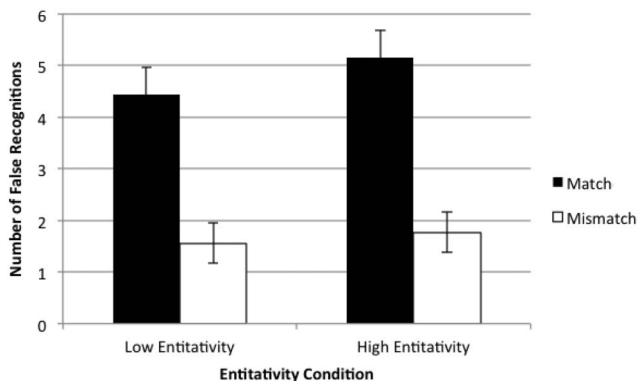
---

[2] Prior to Experiment 3, we conducted another study testing the same hypothesis. The experiment was very similar to Experiment 3 except it used a different manipulation of entitativity, in which participants were induced to think about groups in general as either high or low in entitativity. The results were exactly comparable with those reported in Experiment 3: The manipulation of entitativity produced significant differences on manipulation check measures but no differences in STIG formation. Because of the similarity of these studies and their results, we report only one of the experiments here.

match versus mismatch conditions is consistent with this interpretation and also reinforces the findings of Experiments 1 and 2.

Together, the evidence from Experiments 2 and 3 indicates that STIGs are quite resilient, in that they occurred regardless of cognitive load and of the entitativity of the group described. Given this robustness, a further question concerns the extent and nature of downstream effects that these spontaneous inferences have on group perceptions. Experiments 4 and 5 investigated this question.

Experiment 4: STIGs and Group Perceptions

The amount of research on STIs has increased considerably in recent years (see Uleman et al., 2008 for a review). However, until recently, there has been surprisingly little research directed at identifying and understanding downstream effects of making STIs. That is, given that people make STIs spontaneously, what implications and consequences follow from that fact? From the beginning, one of the reasons for interest in STIs has been the underlying assumption that those inferences would lay the foundation for impressions of the target persons, impressions that may initially emerge without the perceiver's intention or awareness but might then guide future processing. Thus, one implication would be that those STIs ought to have ripple effects on other aspects of the emerging impressions. Recent research addressing this question indicates that STIs do influence other judgments of target persons (e.g., Carlston & Skowronski, 2005; Crawford et al., 2007; McCarthy & Skowronski, 2011b) but not always (Skowronski, Carlston, Mae, & Crawford, 1998).

The present article presents the first studies of spontaneous inferences about groups based on group behavior. Thus far, we have shown that STIGs do occur in processing information about group behavior, that they occur in learning about groups that are both high and low in entitativity, and that they occur even when perceivers' cognitive resources are at least partially constrained by performing another task simultaneously. Thus, STIGs have the properties of being a highly efficient process. In Experiment 4, we extend this work further to seek evidence that making STIGs has implications for other processes. Specifically, just as STIs have implications for emerging person impressions, STIGs may have implications for emerging group impressions and, if so, may lay the groundwork for stereotype development.

In Experiment 4, we sought to provide the first evidence relevant to that question. As in the earlier studies, participants were shown a series of stimulus groups, each consisting of four men whose faces were shown along with a description of a behavior performed by the group. All stimulus groups were presented in the first phase. In the second phase, rather than assessing false recognitions, participants were asked to rate each group on several trait scales. The traits included the trait implied by that group's behavior, a trait implied by another group's behavior, and two additional traits not implied by the group's behavior but equated on likability with the implied trait. These latter traits were included to permit a test for halo effects. That is, if a group performed a desirable behavior, it could lead to inferences of other desirable traits in general. However, our prediction was that, although halo effects may occur, STIG effects on perceptions would be more specific and would primarily influence ratings on the implied trait, significantly more than ratings on the mismatch traits or the traits of equal likability.

Participants also rated their confidence in each of these trait ratings. We reasoned that they would be more confident in ratings on traits that had been inferred about the group (during the learning phase) than about traits that had not been inferred.

We also included the same manipulation of perceived entitativity used in Experiment 3 to examine the effects of this variable on judgments of the target groups.

Method

Participants. Fifty-eight undergraduates (37 females) volunteered to participate in the experiment in exchange for $5. The mean age was 19.53 years ($SD = 1.18$). The racial breakdown of the sample was 32 Whites, 10 Asians, eight Latinos, two Blacks, two multiracials, one Native American, one Native Pacific Islander, and eight others.

Procedure. The procedure was similar to that of Experiments 1 to 3 and adapted the false recognition paradigm to study participants' ratings of stimulus groups. Again, the experiment was introduced as a memory study in which participants would be presented information and would later be asked about it. As in Experiment 3, participants were randomly assigned to one of two entitativity conditions, and they were told that they would learn about groups that were suitemates (high-entitativity condition) or people living in the same dormitory (low-entitativity condition). Next, participants completed the learning phase, in which they read about 34 groups (two groups were omitted due to experimenter error). Each slide was displayed for 10 s and showed faces of the four group members and a sentence describing a behavior presumably performed by that group. Twenty-three of the groups were described as doing behaviors that implied traits. The other 11 were filler groups whose accompanying sentences contained the trait that would be implied by the behavior of that group.

After completing the learning phase, participants completed the group ratings phase. They viewed each group (four faces) from the learning phase, presented one at a time in random order. Participants rated each group on four traits (see Chen, Banerji, Moons, & Sherman, 2014, for similar method). Consistent with the recognition phase of the false recognition paradigm, one trait was implied by the group's behavior (match trait) and one trait was implied by another group's behavior (mismatch trait). The other two traits were control traits, not implied by the group's behavior but equated with the implied trait on likability (Anderson, 1968). After each trait rating, participants rated their confidence in that rating. All ratings were made on 7-point scales. After the ratings task, participants were asked to rate the entitativity of the groups they had learned about to check the effectiveness of the entitativity manipulation. We used the same five questions as in Experiment 3 ($\alpha = .87$). Participants then provided demographic information, were debriefed, and were thanked for their participation.

Design. The experiment had a 2 (entitativity manipulation: high, low) $\times$ 2 (replication: 1, 2) $\times$ 3 (rating type: match, mismatch, controls) mixed model, with the last factor being within subjects. The replication factor varied the particular traits that were matched versus mismatched in the ratings task.

Results

Entitativity manipulation check. An independent samples t test revealed no significant difference in the perceived entitativity

of groups in the high-entitativity condition ($M = 4.98$, $SD = 1.70$) compared with the low-entitativity condition ($M = 4.78$, $SD = 1.70$), $t(56) = 0.45$, $p = .65$. Two one-sample t tests confirmed that participants' entitativity ratings did not differ from the scale midpoint (5) in either the high- or the low-entitativity condition, $ps > .50$. This failure of the manipulation is surprising, in that in Experiment 3, the same manipulation, evaluated on the same response scale, was highly significant. Perceived entitativity was measured on a 9-point scale. Given this outcome, we collapsed across this factor in all subsequent analyses (Footnote 3 reports the analyses separately by entitativity condition).

Impression ratings of stimulus groups. We hypothesized that participants would form STIGs in the learning phase and that these STIGs would then influence their trait ratings of the groups. Specifically, we predicted that participants would rate groups more highly on the specific trait implied by their behavior (*match*) than on traits implied by other groups' behaviors (*mismatch*) or on other traits equated for likability with the match traits (*controls*).

To test these predictions, we first averaged together participants' ratings on the two control traits for each group (ratings on the two control traits were highly correlated, $r[56] = .90$, $p < .001$). Then we conducted a 2 (replication) × 3 (rating type: match vs. mismatch vs. controls) mixed model ANOVA on participants' trait ratings. The predicted main effect of rating type was highly significant, $F(2, 112) = 35.28$, $p < .001$, $\eta_p^2 = .39$. As shown in Figure 4, participants rated groups more highly on match traits ($M = 5.66$, $SD = 0.85$) compared with mismatch traits ($M = 5.06$, $SD = 0.66$) and control traits ($M = 5.19$, $SD = 0.67$), both $ps < .001$. There was not a significant main effect of replication, $F(1, 56) = .24$, $p = .63$, $\eta_p^2 = .004$.

There was a marginally significant Rating Type × Replication interaction, $F(2, 112) = 3.66$, $p = .06$, $\eta_p^2 = .06$. To determine whether our predictions held in both replication conditions, we conducted pairwise comparisons within each replication. In Replication 1, participants rated groups more highly on match traits ($M = 5.73$, $SD = 0.98$) than on mismatch traits ($M = 4.95$, $SD = 0.67$) and control traits ($M = 5.12$, $SD = 0.76$), both $ps < .001$. In addition, participants rated groups more highly on control traits than on mismatch traits, $p = .048$. In Replication 2, participants also rated groups more highly on match traits ($M = 5.59$, $SD =$

0.68) than on mismatch traits ($M = 5.20$, $SD = 0.64$) and control traits ($M = 5.26$, $SD = 0.67$), both $ps < .01$. However, there was no difference in the extent to which participants rated groups on mismatch traits compared with control traits, $p = .50$. Therefore, in both replication conditions, participants rated the groups significantly higher on traits implied by the groups' behaviors compared with traits implied by different groups' behavior or on evaluatively similar (but not implied) traits. These results provide evidence for the predicted influence of STIGs on perceptions of those groups and for the specificity of those effects.[3]

Confidence in impression ratings of stimulus groups. We also measured the confidence with which participants made their trait judgments. We conducted a 2 (replication) × 3 (rating type: match vs. mismatch vs. controls) mixed model ANOVA on participants' confidence ratings. The only significant effect that emerged was the predicted main effect of rating type, $F(2, 112) = 7.03$, $p = .01$, $\eta_p^2 = .11$. Participants were significantly more confident in their match trait ratings ($M = 4.56$, $SD = 1.68$) than in the control trait ratings ($M = 4.24$, $SD = 1.70$), $p < .001$, and were marginally more confident in their match trait ratings than in the mismatch trait ratings ($M = 4.36$, $SD = 1.60$), $p = .07$. These results are essentially parallel to the results of the trait ratings, though not as pronounced.

Discussion

The results of Experiments 1 to 3 effectively documented that STIGs occur under a variety of experimental conditions. The results of Experiment 4 extend those findings by showing that STIGs have downstream effects on the impressions participants formed of the stimulus groups. Participants made stronger ratings about the target groups on those traits implied by the groups' behaviors than on traits implied by other stimulus groups they had read about. This is an important comparison because both match and mismatch traits had presumably been activated by the behaviors presented during the learning phase. This difference in ratings on match and mismatch traits provides evidence of the specificity of STIGs in their influence on subsequent judgments of the groups.

The behaviors that groups perform vary in valence, and therefore the traits inferred in the STIG process vary in desirability. One
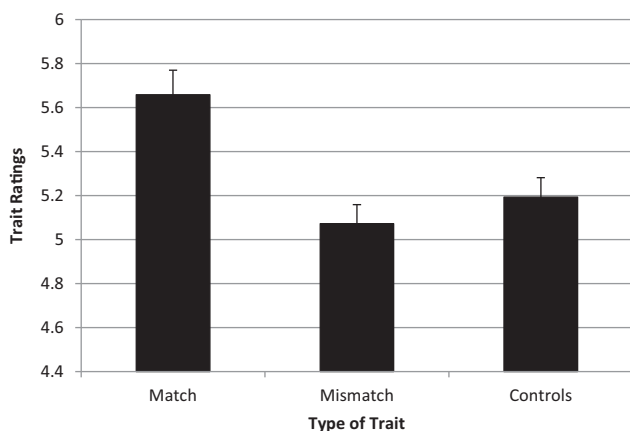

Figure 4. Mean trait ratings of groups by type of trait (Experiment 4). Error bars represent standard error.

---

[3] We also ran the 2 (replication) × 3 (rating type) ANOVA separately by entitativity condition (low vs. high). In the low-entitativity condition ($n = 30$), there was only a significant effect of rating type, $F(1, 56) = 17.48$, $p < .001$, $\eta_p^2 = .38$. Consistent with STIG formation, participants rated the groups higher on match traits ($M = 5.58$, $SD = 0.70$) compared with mismatch ($M = 5.08$, $SD = 0.55$) and control traits ($M = 5.25$, $SD = .57$), both $ps < .001$. They also rated groups higher on control traits than mismatch traits, $p = .04$. In the high-entitativity condition ($n = 28$), there was a significant effect of rating type, $F(2, 52) = 19.79$, $p < .001$, $\eta_p^2 = .43$. Consistent with STIG formation, participants rated groups higher on match traits ($M = 5.75$, $SD = 0.99$) compared with mismatch ($M = 5.05$, $SD = 0.78$) and control traits ($M = 5.13$, $SD = 0.76$). There was no difference in mismatch and control trait ratings. There was also a marginal Replication × Rating Type interaction, $F(2, 52) = 3.79$, $p = .06$, $\eta_p^2 = .13$. Replication 1 displayed the same pattern of means and significance levels as the overall main effect of rating type. In Replication 2, participants rated the groups higher on match traits ($M = 5.59$, $SD = 0.61$) than control traits ($M = 5.15$, $SD = 0.50$), $p = .004$, but not on mismatch traits ($M = 5.24$, $SD = 0.78$), $p = .13$. There was no difference in ratings on control and mismatch traits, $p = .54$. However, this interaction should be interpreted cautiously, as cell size was small (13 and 15 per replication).

possibility is that these inferences are of a general evaluative nature, such that a group that does one highly desirable act (e.g., taking orphans on an afternoon outing to a zoo) would generate an inference not only on the corresponding trait (kind) but also on other similarly evaluative traits. Therefore, we had participants rate the groups on two additional traits that were equated on likability with the behavior-implied trait as a means of testing for such halo effects. Participants' ratings on the match traits were significantly higher than their ratings on these evaluatively equated traits. Moreover, participants had greater confidence in their judgments on traits implied by a group's behavior than on these other traits. These results provide further useful evidence of the specificity of STIGs. They indicate that STIGs are content based and not merely evaluative inferences.

Experiment 5: Generalization of STIG-Based Impressions

When Katz and Braly (1933) conducted their classic study of stereotypes, they presented participants the names of national or ethnic groups (e.g., Germans, Turks, Negroes) and had participants check off, from a list of trait adjectives, those traits that they thought characterized each group. It was, then, a study of inferences within a group perception context. This trait inference paradigm became the primary methodological tool used to study stereotypes for the next several decades (see Brigham, 1971; Hamilton et al., 1994). An obvious weakness of the methodology was that the purpose of the study and the interests of the investigators were evident to the participants. Moreover, the process underlying these inferences was conscious and deliberative. These problems led to an enormous amount of research aimed at finding ways of measuring stereotypes that were less reactive and less transparent (see Olson, 2009; Schneider, 2004, Chapter 2).

Our results document that inferences about groups (STIGs) can arise spontaneously, without conscious intent. Once inferred, these inferred attributes become properties of a newly emerging group impression. If that group impression were to persist, and if it were to be transferred to other members of the group, it would lay the foundation for development of a group stereotype, formed spontaneously and without intention. That is, a group-based concept—the inferences represented in STIGs—would generalize to other group members. Such generalization has been one of the hallmarks of stereotypes since Allport's (1954) classic analysis. The purpose of Experiment 5 was to test the hypothesis that STIG-based traits generalize to another group member.

Earlier we cited research that bears on this question. Crawford et al. (2002) had participants read about individual members of two different groups. Every member of Group A performed a behavior that implied one of two traits (lazy or intelligent), and every member of Group B performed a behavior that implied one of two other traits (aggressive, honest). Crawford et al. determined the extent to which these traits not only were inferred about the actor (STIs) but also became associated with other members of the actor's group. They found that, if the group was high in entitativity, both of the traits implied by behaviors of Group A members became associated with all members of that group, and similarly for Group B members. This study demonstrated generalization of spontaneously inferred traits of *individual* group members (STIs) to other members of the same (but not of the other) group.

The focus of Experiment 5 concerns generalization of a different type, that is, generalization from STIGs based on *group* behaviors to inferences about individual members. If that occurs, it would provide evidence that STIG-based inferences persist and are applied to other members of the group about whom the perceiver has no information beyond group membership.

Our reasoning carried this analysis one step further in an effort to understand the process underlying such generalization. In our paradigm, participants first learn about a number of stimulus groups, including reading about the group's behavior. In the recognition phase, they are asked if a specific word was in the sentence describing that group. A false recognition is an indication that a spontaneous trait inference has been made. Experiments 1 to 3 showed that people make STIGs for match groups more than for mismatch groups. Even for match groups, however, they of course do not make a false recognition in every case. The generalization of a spontaneously inferred trait to a new group member should occur only when the STIG had been made. This implies that ratings of the new member should be higher for traits that had in fact been inferred in the group learning phase. We also tested this hypothesis in Experiment 5.

Method

Participants. Sixty-five undergraduate students (45 female) completed the study for research credit as part of an introductory psychology course. Of these, 21 self-identified as Asian/Asian American, 19 as White/Caucasian, 18 as Latino/a, three as African American/Black, three as "Other," and one as Multiracial/Mixed.

Stimulus photos. The original stimulus set was expanded from the previous studies to include 180 total male faces. All the faces had neutral expressions and were presented in front of a white background and in gray scale. As in previous studies, groups of four faces were created by displaying four photos of individual faces on a single frame, and the behavior-descriptive sentences were presented below the photos.

Materials and procedure. The procedure mirrored that of Experiment 4, with a few exceptions. Upon arriving at the laboratory, participants were told that the study was about learning and memory, and that they would complete the study on a computer. As in the previous studies, participants first completed a learning phase, in which they saw 36 separate sets of group photographs (created per the procedure outlined in the Stimulus Photo section). The participants were told that each group photograph represented a different group, and each set of photographs was accompanied by a sentence describing a behavior performed by the group. Each set of photographs appeared for a period of 10 s before automatically advancing to the next group. Of the 36 trials, 24 were critical trials and contained behaviors that implied trait characteristics. Twelve filler trials were also presented in which the trait word was explicitly stated. As in the previous studies, the filler sentences were included so that participants had an opportunity to correctly respond "Yes" during the recognition phase. These trials were not of theoretical interest and the data analyses do not include them.

The recognition phase was the same as that used in Experiments 1 to 3. All participants were presented with the same 36 sets of four faces in random order. Each group was accompanied by a trait word, and participants were asked to indicate whether the word they saw had appeared in the sentence paired with the group

photograph in the learning phase. For 12 of the trials, the probe word was the trait implied by the behavior that group had performed (match trials). For 12 of the trials, the probe word was a trait implied by the behavior of a different group (mismatch trials). The remaining 12 trials constituted the filler trails previously discussed.

Following the recognition phase, all participants then completed the generalization phase. During this phase, participants were told that they would be shown the same groups from the previous phases of the study, but that no behavioral or trait information would be presented. Instead, participants were told that they would be shown the picture of another member from the group who had not previously been presented, and that their task was to provide their impression of each new person by rating him on several scales. These scales assessed the same traits as those used in Experiment 4. One of them was the trait that was implied by the group's behavior (match trait), one trait was implied by another group's behavior (mismatch trait), and the other two traits were control traits, not implied by the group's behavior but equated with the implied trait on likability (Anderson, 1968). After each trait rating, participants also rated their confidence in that rating. All ratings were made on 7-point scales. After the ratings task, participants provided demographic information, were debriefed, and were thanked for their participation.

Design. The experiment had a 2 (Condition 1, 2) × 2 (trial type: match, mismatch) × 3 (rating type: match, mismatch, paired adjective) mixed design, with the latter two factors being within subjects. Two replication conditions were run to counterbalance which traits served as match versus mismatch traits during the generalization phase. This was necessary in order to avoid having a trait word appear twice—once with the group whose behavior implied the trait, and once with another group whose behavior had not implied the trait. The dependent variables were (a) the total number of false recognitions made in the recognition phase, (b) trait ratings of new group members during the generalization phase, and (c) rated confidence in the trait ratings. All participants saw each type of sentence in the recognition phase and made each type of rating in the generalization phase.

Results and Discussion

Evidence of STIGs in recognition phase. We first determined whether the participants had made STIGs about the groups presented in the study. This was done by tallying the number of false recognitions made for match and mismatch trials in the recognition phase and conducting a 2 (condition) × 2 (trial type: match and mismatch) mixed model ANOVA on participants' recognition rates. As expected, there was a significant main effect of trial type, $F(1, 60) = 13.15$, $p < .001$, $\eta_p^2 = .18$, with participants making more false recognitions on match trials ($M = 5.33$, $SD = .30$) than on mismatch trials ($M = 4.30$, $SD = .30$). However, there was no interaction with condition, nor was there a main effect for condition. Given the lack of any effect of the counterbalancing conditions, they served as theoretical replications and therefore were combined for further analyses.

Generalization of STIGs to new group members. We hypothesized that STIGs formed about a group in the learning phase would generalize to novel group members and influence trait ratings of these members. We predicted that participants who made

a STIG about a group in the learning phase would rate a new group member more highly on the specific trait implied by the behavior their group had engaged in (*match*) than on a trait implied by the behavior performed by a different group (*mismatch*) or on traits equated for likability with the match trait (*controls*). We also hypothesized that when participants did not make a STIG about a group, there would be no generalization to new group members.

To test these predictions, we first identified the trials on which participants had made a STIG versus the trials on which they had not. Then we calculated participants' average ratings for the three types of trait (i.e., match, mismatch, control) for these two categories. With the data thus sorted, we conducted a 2 (STIG vs. no STIG) × 3 (rating type: match vs. mismatch vs. controls) repeated measure ANOVA on participants' trait ratings. The predicted effect for the type of trait rating was highly significant, $F(2, 60) = 32.35$, $p < .001$, $\eta_p^2 = .52$, and qualified by a significant and predicted interaction, $F(2, 60) = 7.90$, $p < .001$, $\eta_p^2 = .21$.

As shown in Figure 5, participants rated new members more highly on match traits, but only when they had made a STIG about the group in question. When they had made a STIG about a group, participants rated the new member significantly more highly on match traits ($M = 5.37$, $SD = 1.06$) than on mismatch traits ($M = 5.01$, $SD = 1.51$) and control traits ($M = 4.24$, $SD = 1.19$), both ps <. 001. When the participants did not make a STIG, there was no difference between their ratings of new members on the match traits ($M = 5.03$, $SD = .92$) compared with the mismatch traits ($M = 4.97$, $SD = 1.07$), whereas ratings on the control traits were lower ($M = 4.71$, $SD = .93$). Additionally, the difference between participant ratings of new members on match traits was higher when they made a STIG than when they did not make a STIG ($p < .05$).

Confidence in impression ratings of new members. Participants' confidence ratings for their trait judgments were analyzed in a 2 (STIG vs. no STIG) × 3 (rating type: match, mismatch, controls) repeated measures ANOVA. The only significant effect was the main effect for rating type, $F(2, 60) = 31.82$, $p < .001$. Participants had more confidence in their ratings for match ($M = 5.37$, $SD = 1.73$) and mismatch ($M = 5.42$, $SD = 1.92$) traits than
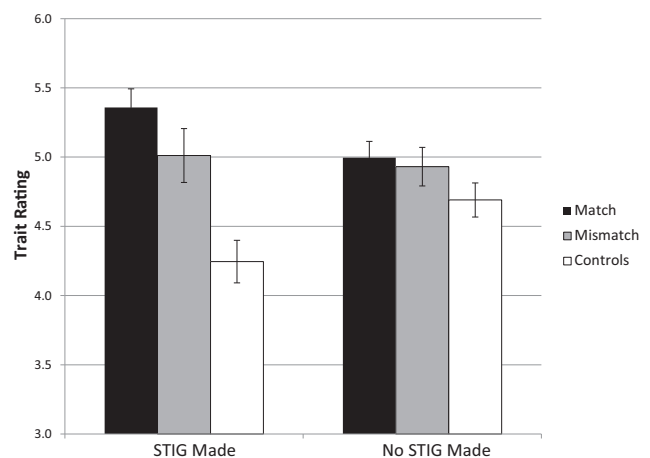
Figure 5. Mean trait ratings of new group members by type of trait on trials in which STIGs were and were not made (Experiment 5). Error bars represent standard error. STIG = spontaneous trait inferences about groups.

for control ($M = 4.83$, $SD = 1.65$) traits. The predicted difference between match and mismatch traits was not significant.

General Discussion

It has been 30 years since Winter and Uleman (1984) first introduced the idea that perceivers make trait inferences from the behaviors of actors spontaneously, without intention and without awareness that they are doing so. Since then, research on the topic has increased enormously, and it has addressed many questions about the processes underlying STIs and the moderators that qualify their occurrence. Given the importance of group perception in social psychology, it is perhaps surprising that all studies of STIs have focused on behaviors of, and inferences about, individual target persons. Our research extends this literature by demonstrating that people make STIGs on the basis of group action. Just as STIs provide the seeds for new person impressions, STIGs provide the seeds for new group impressions that may ultimately develop into group stereotypes.

The five experiments reported here have consistently demonstrated that perceivers make STIGs as they comprehend and process group actions. Experiment 1 provided a direct comparison of STIs and STIGs, using the same behaviors describing either an individual or group target. People made both STIs and STIGs, and the two target conditions did not differ significantly in the frequency of such spontaneous inferences. As discussed earlier, there are reasons one might have expected that STIs would be stronger or more likely than STIGs. Several lines of research have shown differences in processing and using information about individual and group targets (Hamilton & Sherman, 1996; Hamilton et al., 2014). The results of Experiment 1 did not reveal such differences in spontaneous inferences. At this initial stage of behavior comprehension and encoding, processing of both individual and group behavior appears similar in generating spontaneous inferences. These results provided our first indication that STIGs may reflect a basic, routinized, and resilient aspect of encoding group behavior information. Of course, there may be social circumstances in which a difference might appear more clearly, and there may be individual differences in the propensity to make STIs versus STIGs. These are worthwhile topics for future research.

Experiment 2 examined the efficiency of the STIG process by studying performance under cognitive load. The literature on the effects of load on STIs has been somewhat mixed, often finding little or no reduction in STI under load (e.g., Crawford et al., 2007; Todd et al., 2011; Todorov & Uleman, 2003), but sometimes showing that cognitive load reduces the formation of STIs (e.g., Wells et al., 2011). In this first study of STIG formation under load, our results suggest that STIGs are a highly efficient process, as cognitive load did not alter the extent of false recognitions. Additional tests of this relationship, using different paradigms and cognitive load manipulations, would be valuable.

Experiment 3 investigated the relative frequency of making STIGs in processing behavioral information about groups that were high or low in entitativity. Research on perceived group entitativity has documented numerous important differences in the way information about high- versus low-entitativity groups is processed and used (see Hamilton et al., 2002; Hamilton et al., 2011). In fact, McConnell et al. (1994, 1997) showed that groups high in entitativity are perceived as organized units much like

persons, thereby facilitating making inferences about them, compared with low-entitativity groups. If so, then STIGs may occur more readily in processing behavioral information about high- than low-entitativity groups. Our results were not consistent with this reasoning. Although our manipulation of perceived group entitativity in Experiment 3 was significant, it produced no difference in the rate of false recognitions.

Again, there are other ways of defining groups to be compared. For example, Lickel et al. (2000) empirically differentiated several types of groups (e.g., intimacy, task, social categories) that varied in a number of properties (e.g., size, extent of interaction, shared goals and outcomes) as well as in perceived entitativity, and these group types are spontaneously used by perceivers as they encode and store information about group members (Sherman, Castelli, and Hamilton (2002)). A useful avenue for future research would be to compare these types of groups on the extent to which they foster STIGs. Such group comparisons can also be extended into the intergroup domain, determining, for example, the frequency of STIGs for in-groups and out-groups.

The lack of differences in STIGs for high- and low-entitativity groups is, however, meaningful because it conforms to past findings on STIs, showing that they occur early and quickly as information is encoded (Todorov & Uleman, 2002, 2003, 2004). If STIGs (like STIs) are truly spontaneous, then it is quite reasonable that they would occur in processing behavioral information about any group, regardless of its group properties. Our results are consistent with this interpretation.

In all three of these experiments, an independent variable (person vs. group target, cognitive load, group entitativity) did not produce significant outcomes on the dependent measure. These nonsignificant effects may be disconcerting and perhaps would have reached significance with larger sample sizes (although our sample sizes were not unusually small compared with other studies). However, despite this fact, in all three cases, the results are theoretically noteworthy, as other aspects of the findings give them meaning. First, the findings were informative. Experiment 1 showed that STIs of comparable (and statistically significant) magnitude occurred for both individual and group targets. Experiment 2 showed that STIGs occurred significantly whether under cognitive load or not, documenting the efficiency of the process. Experiment 3 showed that STIGs were significant for both high- and low-entitativity groups. None of these results have been reported previously. Second, in each case, the manipulations themselves were effective. In Experiment 1 the manipulation of target (person or group) was inherent in the stimuli. In Experiments 2 and 3, the manipulation checks for cognitive load and entitativity manipulations, respectively, were highly significant. Third, and of greatest theoretical importance, the evidence from these studies documented that STIGs consistently occurred as predicted. That is, the signature evidence of STIG occurrence—the comparison of false recognitions for match and mismatch trials—was significant in every condition of every experiment. The consistent statistical significance of the "false recognition effect" in numerous experimental conditions across these experiments provides ample demonstration of the robustness and resiliency of STIGs as perceivers process information about group targets.

Experiment 4 extended this work by investigating the extent to which STIGs can influence perceptions of those groups. Participants made stronger ratings, and were more confident in those

judgments, on traits implied by the groups' behaviors than on other previously activated (but not group-relevant) traits or on other evaluatively similar traits. The emerging group impressions generated by STIGs are not generalized evaluative impressions, but rather are content-specific inferences about specific target groups based on their behaviors. STIGs therefore constitute the very beginnings of group perceptions that could grow and develop into stereotypic conceptions of those groups. In that sense, STIGs may sow the seeds of stereotyping.

Experiment 5 provided evidence extending that argument in a meaningful way. One could contend that a STIG is a momentary product of a fast inference process, but has limited staying power and therefore is not likely to influence other aspects of group impressions. In contrast to this view, the results of Experiment 5 document that the group beliefs represented by STIGs in fact are subsequently applied to perceptions of a new group member about whom one has received no information beyond his membership in a target group. Moreover, the group beliefs represented by STIGs formed spontaneously while participants were engaged in a memory task. Thus, the STIG-based beliefs have generalized and have been transferred to a new group member.

Experiment 5 also provided evidence for an important element in the proposed process underlying this generalization effect. Specifically, participants rated the new group member higher on those traits that had been inferred, as evidenced by "Yes" responses to the trait probe in the recognition phase. However, on traits for which the participants had not made a false recognition, there was no difference in ratings of the new member on match versus mismatch traits. Thus, the generalization effect was conditional on having made a STIG about the target group.

A New Perspective on Stereotype Formation

The results of these five experiments provide a solid evidentiary basis for the conceptual innovation offered in the introduction to this article. Specifically, we have shown that inferences about groups, based on group actions, can create group-descriptive concepts in memory that form impressions of the stimulus groups, simply as a result of a spontaneous inference process. These group-descriptive concepts are formed efficiently without intention and without the goal of forming impressions, as participants believed they were participating in a memory experiment and their task was to learn and remember the stimulus information. Nevertheless, these concepts were sufficiently implanted in memory that they not only produced false recognitions of probe words (Experiments 1, 2, 3, and 5) but also influenced perceivers' subsequent ratings of the group (Experiment 4), and were generalized and applied to a new group member (Experiment 5). These data resemble some of the important properties of stereotypes, yet they have emerged from a spontaneous inference task that focuses participants' attention on memory for verbal information.

Considered together, this set of findings strongly implies that the process of making STIGs from group behaviors can constitute the beginnings of stereotype formation. Moreover, this stereotype formation is occurring in the absence of the most commonly cited preconditions on which stereotypes presumably rest. That is, in our experiments, these stereotype-like effects occurred under conditions in which (a) the perceiver's goal was not focused on forming beliefs about the group; (b) the perceiver had no prior knowledge

or preexisting beliefs about the group; (c) there was no interdependence, no competition for scarce resources, and no feelings of relative deprivation between the perceiver's own group and the target group; and (d) it all occurred in a context in which intergroup perceptions and comparisons between groups were not present. As such, the results of our experiments point to a new process by which stereotypes can be formed.

Actual stereotypes, of course, are more full-blown products. However, the idea that STIGs could develop further in this way is not implausible. We know that first impressions, as well as preexisting expectancies, induce a confirmatory bias that preserves and enhances the status quo. We also know that STIs about individuals influence predictions of those individuals' future behavior (McCarthy & Skowronski, 2011b). Thus, an important agenda for future research will be to explore how these "seeds" grow and blossom into more fully developed conceptions of groups (stereotypes). An important element in that endeavor will be determining what variables influence that process.

Remaining Questions of Interpretation

Our research extends past work by studying spontaneous inferences about groups rather than about individual persons. The results of five experiments document the importance of STIGs in group perceptions. As in any new line of work, there are lingering issues of interpretation. In this section, we comment on some of those questions.

Group versus multiple persons. We have argued that STIGs represent trait inferences spontaneously made about groups. The paradigm we use (adapted from Todorov & Uleman, 2002) introduces those groups by showing face photos of four persons who constitute the group, and the question posed to participants is whether a probe word was in the sentence describing that group. It could be, however, that participants are not spontaneously forming a group concept, but instead are simultaneously making parallel STIs about each of the four individuals who are shown, and those four STIs are then combined into a group impression. In both cases the consequence is a group impression that is the result of STIs. The difference is in whether those spontaneous inferences themselves are about individual members or about the group as a unit.

This alternative process is possible, but to be viable, one would have to assume that participants ignored important aspects of the instructions provided to them and instead did something quite different. In our paradigm, it is not the case that individual group members were described as having enacted behaviors that might foster STIs (as, for example, in Crawford et al.'s, 2002, studies). Rather, the behavior was described as a group action performed by the group as a unit. Moreover, the instructions clearly stated that participants would learn about different groups of persons, that each slide showed a different set of four faces, and that each set of four constituted a different group. Later, in the recognition phase, the key dependent measure presented a probe word and asked if that word had been in the sentence "describing that group." It seems unlikely that participants would answer that question by first considering if each of the four group members possessed that trait.

Priming trait concepts. A second question about interpretation concerns the possible role played by the filler items, in

which probe words are presented that in fact were in the stimulus sentence describing the group shown. Could it be that having participants reply "Yes" to these filler items primes trait concepts and/or an impression formation goal? If so, then it may increase the likelihood of saying "Yes" to probe words for other, nonfiller items, thereby increasing the number of false recognitions observed. Again, a priming effect of this type could contribute to the observed results, but seems unlikely as a satisfactory account for several reasons. First, if the filler items encourage trait inferences in this way, the effect would apply to both match and mismatch trials, yet it was the consistent difference between match and mismatch trials in rate of false recognitions that provided the empirical support for our hypotheses. Second, although some probe words on filler trials were trait terms, several others (e.g., "surprised," "soaked," "thirsty") were not the kinds of words that would induce an impression processing goal. For these reasons, we are skeptical about this alternative account of the findings.

Inference versus association. An important question debated in the STI literature concerns the extent to which such effects reflect inferences or associations. As trait inferences, STIs are based on behavior manifested by the actor, and the traits are inferred during behavior encoding to be properties of the person. As associations, STIs are not inferences about the person but are the product of the behavior activating a trait concept in the presence of the person. This is a shallower process in which the trait is not attributed to the person, but instead the two become associated simply by their contiguity. Carlston, Skowronski, and colleagues (Carlston & Skowronski, 2005; Mae, Carlston, & Skowronski, 1999; Skowronski et al., 1998) have studied this distinction in the context of STIs and spontaneous trait transferences (STTs). The theoretical question became whether STIs are really based on an inference process, as many had assumed, or are simply a product of a much simpler associative process.

Research comparing STIs and STTs has been very useful, and from it we have learned several important points: (a) Both STIs and STTs are robust, occurring over time and under a variety of testing conditions; (b) Both of them can occur in response to the same behavioral information (e.g., if Bob describes Ann as having worked hard on a campaign for a local mayoral candidate, the listener not only will infer that Ann is politically active [STI] but also will believe that Bob is politically active as well [STT], despite there being no information about Bob's political activities); (c) STIs and STTs are differentially influenced by other factors, suggesting that they occur as a function of different underlying processes (inference vs. association); and (d) STIs and STTs differ in strength. STIs are consistently stronger effects than are STTs, and in direct comparison tests, STIs may be as much as twice as strong as STTs.

In this article, we have characterized our results in terms of a spontaneous inference process (STIGs). However, we know that STT effects mimic STIs (though they are weaker), and our studies were not designed to provide evidence of the relative contribution of inference versus association processes to the STIG results reported here. The findings from our five experiments establish STIG as a robust effect. An important agenda item for future research is to determine the extent to which these results reflect inferences about the group based on the manifest properties of

their group behaviors or the formation of associations based on contiguity.

New Questions About Spontaneous Inferences

Our research on STIGs also generates numerous new research questions. In fact, an important aspect of the potential contribution of studying STIGs is that it raises questions that have not been—and in some cases could not be—raised in the study of STIs for individual persons. We briefly offer some examples.

Properties of stimulus groups. Several questions of generality of our findings need to be pursued. One question concerns the properties of the stimulus groups. In our studies, the groups have been homogeneous, as shown in photos of four White males. Would the same results be obtained if the groups were composed of four White females? African Americans? Asians? Latinos? We know of no theoretical reason to expect different results in these cases, but these group characteristics have been shown to influence processing in other contexts. Studies comparing different target groups would provide an empirical answer. Extending this line of thought, would STIGs occur as spontaneously and as reliably if the stimulus groups were heterogeneous, for example, if they included a mix of male and female persons or included persons of different races? Would this greater heterogeneity diminish the ease or extent to which STIGs are made? Another potentially important group property is the size of the group. Our stimulus groups have always had four members. That number was arbitrary and chosen for convenience (e.g., availability of photos) more than for any theoretical concerns. However, other research has demonstrated that group size is an important variable that can affect aspects of group perception and group functioning. Studies investigating these effects on STIGs need to be on the agenda for future research.

Generalization. Experiment 5 showed that STIG-based beliefs about a group generalize to a new, previously unencountered member of the group. This demonstration was important because such generalization is a key component of stereotyping in group perceptions. Crawford et al. (2002) have already shown that (in high-entitativity groups) STIs based on behaviors of individual group members can generalize to other group members, resulting in the apparent similarity of members based on these STIs. Experiment 5 documents a different type of generalization. In contrast to STIs, STIGs are based on group actions and the inference drawn applies to the group as a whole. These STIGs influence trait judgments of groups, as shown in Experiment 4, reflecting an emerging group impression of each of the stimulus groups presented. In Experiment 5, those group impressions were applied to a new group member about whom participants knew nothing other than his group membership. This generalization is part of the very essence of stereotyping. It is why, in this article, we have referred to STIGs as sowing the seeds of stereotyping.

Although this generalization has been demonstrated in Experiment 5, more research is needed to further explore this effect. How general is this generalization effect? Any of the group properties discussed in the preceding paragraph presumably could moderate the occurrence and/or the strength of generalization. For example, would STIGs based on race-homogeneous groups generalize to new, other-race members (Chen & Ratliff, 2015; Ratliff & Nosek, 2011)? It could also be that generalization occurs for some types of groups but is less common for others. In addition, some perceivers

may be more prone to generalizing STIG-based impressions than are others as a function of, for example, preexisting attitudes. At present, we do not have answers to these questions, yet they are important issues to investigate in order to determine the parameters on this effect.

More generally, we also believe that the spontaneous generalization effect in Experiment 5 is an excellent example of a question that naturally arises in thinking about STIGs, but seems to have little or no counterpart in thinking about STIs. We know of no study exploring the generalization of STIs; indeed, it is difficult to imagine the context in which it would have real meaning. Generalization to whom? To another member of the target person's group? That would be potentially interesting and informative, although the only difference between that effect and the present finding is whether the initial inference to be generalized is derived from one person's behavior (STI) or from a group's action (STIG). Moreover, we already know that (under some conditions) transfer of an inference (STI) from one group member to other members can happen during the initial encoding phase (Crawford et al., 2002). Moving away from group contexts, one could ask if an STI based on one person's behavior would generalize to another randomly selected person. If so, why? Such an effect would likely occur as a consequence of trait priming, or a halo effect, or one's implicit personality theory. In contrast, generalization of a STIG-based group impression to a new group member seems to carry deeper meaning and broader potential downstream consequences for that target person.

Group identity. The groups presented in our studies were anonymous groups; participants were simply told that the four members shown on each slide was a separate group (in Experiments 3 and 4, they were told the members of each group lived in the same dormitory, but were given no other identifying information). Of course, in everyday life, we do not often encounter such anonymous groups. The groups we observe and interact with differ in many ways, but we typically know what kind of group it is and know (or believe we know) some of its properties. Introducing this information about stimulus groups would increase the real-world correspondence of the research. However, doing so creates new difficulties that would need to be addressed. Specifically, participants obviously will have knowledge of, and beliefs about, such groups, so research will need to include a means of differentiating new spontaneous inferences (STIGs) from inferences based on prior beliefs about the groups. Research using individual target persons has shown that STIs are less likely to be made for behaviors that are inconsistent with the stereotype of the actor's group than for stereotype-consistent behaviors (Ramos, Garcia-Marques, Hamilton, Ferreira, & Van Acker, 2012; Wigboldus, Dijksterhuis, & van Knippenberg, 2003). Investigating the parallel question for inferences based on group behaviors (STIGs) is clearly an important topic for future research.

Multigroup contexts. In our experiments participants learned information providing one behavioral fact about each group and the question was whether a spontaneous inference about the group would be made. An important extension of this research would be to multigroup contexts in which two or more groups are encountered repeatedly in stimulus information and STIGs could be assessed.

One of the most reliable findings in the intergroup perception literature is in-group bias, the strong tendency to have more favorable evaluations of in-groups than out-groups. What effect does one's membership in one group have on spontaneous inferences, particularly on the valence of inferences about in-groups and out-groups? Otten and Moskowitz (2000) studied this question for STIs about individual persons, using a minimal group paradigm, and found evidence consistent with in-group favoritism. Given our findings that perceivers make spontaneous inferences from group behaviors, it becomes important to determine whether STIGs about in-groups and out-groups manifest this in-group bias as well. Moreover, whereas our studies have focused on trait inferences, emphasizing the content of emerging group impressions, other work (Otten & Moskowitz, 2000; Schneid, Carlston, & Skowronski, 2015; Schneid, Crawford, Skowronski, Irwin, & Carlston, 2015) has shown that perceivers make spontaneous evaluative inferences. The relative importance of evaluative and descriptive spontaneous inferences, and their interplay, in group perceptions remains fertile ground for future research.

Cultural differences. Finally, it seems plausible that there may be cultural differences in the propensity to make STIs and STIGs. Specifically, one way that East Asian and European American cultures differ is in the "unit of analysis" in social perception and cognition (e.g., Menon, Morris, Chiu, & Hong, 1999; Spencer-Rodgers, Williams, Hamilton, Peng, & Wang, 2007; Zárate, Uleman, & Voils, 2001). For Westerners, the individual person is seen as the locus of causation, the unit of organization, whereas for East Asians, the group as a unit plays that role to a much greater extent. This cultural difference may suggest that, as a consequence of living in these cultural contexts, parallel differences in habitual spontaneous inference processes may develop. Specifically, European Americans may be more inclined to make STIs than STIGs, whereas for East Asians, making STIGs may be a more natural process than making STIs (see Na & Kitayama, 2011). Again, we know of no research specifically pursuing that difference.

Conclusion

Our research has extended work on STIs in new and meaningful ways and has provided ample directions for future work. Our results document that STIGs are a robust and resilient aspect of processing information about groups. Just as STI research has provided important evidence of spontaneous processes that can contribute to the initial formation of an impression of a person, research on STIGs may reveal the role of spontaneous processes in sowing the seeds of stereotypes by spontaneously planting inferred dispositions of groups simply as a part of comprehending group behaviors.

References

Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison Wesley.

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology, 9,* 272–279. http://dx.doi.org/10.1037/h0025907

Brigham, J. C. (1971). Ethnic stereotypes. *Psychological Bulletin, 76,* 15–38. http://dx.doi.org/10.1037/h0031446

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. http://dx.doi.org/10.1037/0022-3514.66.5.840

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.

Chen, J. M., Banerji, I., Moons, W. G., & Sherman, J. W. (2014). Spontaneous social role inferences. *Journal of Experimental Social Psychology, 55,* 146–153. http://dx.doi.org/10.1016/j.jesp.2014.07.003

Chen, J. M., & Ratliff, K. A. (2015). Implicit attitude generalization from Black to Black–White biracial group members. *Social Psychological & Personality Science, 6,* 544–550. http://dx.doi.org/10.1177/1948550614567686

Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology, 83,* 1076–1094. http://dx.doi.org/10.1037/0022-3514.83.5.1076

Crawford, M. T., Skowronski, J. J., Stiff, C., & Scherer, C. R. (2007). Interfering with inferential, but not associative, processes underlying spontaneous trait inference. *Personality and Social Psychology Bulletin, 33,* 677–690. http://dx.doi.org/10.1177/0146167206298567

Crosby, F. (1976). A model of egotistical relative deprivation. *Psychological Review, 83,* 85–113. http://dx.doi.org/10.1037/0033-295X.83.2.85

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology, 54,* 733–740. http://dx.doi.org/10.1037/0022-3514.54.5.733

Goren, A., & Todorov, A. (2009). Two faces are better than one: Eliminating false trait associations with faces. *Social Cognition, 27,* 222–248. http://dx.doi.org/10.1521/soco.2009.27.2.222

Hamilton, D. L. (1976). Cognitive biases in the perception of social groups. In J. S. Carroll & J. W. Payne (Eds.), *Cognition and social behavior* (pp. 81–93). Hillsdale, NJ: Erlbaum.

Hamilton, D. L., Chen, J. M., & Way, N. (2011). Dynamic aspects of entitativity: From group perception to social interaction. In R. M. Kramer, G. Leonardelli, & R. Livingston (Eds.), *Social cognition, social identity, and intergroup relations: A Festschrift in honor of Marilynn B. Brewer* (pp. 27–52). New York, NY: Psychology Press.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12,* 392–407.

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.

Hamilton, D. L., & Sherman, S. J. (1996). Perceiving persons and groups. *Psychological Review, 103,* 336–355. http://dx.doi.org/10.1037/0033-295X.103.2.336

Hamilton, D. L., Sherman, S. J., & Castelli, L. (2002). A group by any other name—The role of entitativity in group perception. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 12, pp. 139–166). Chichester, UK: Wiley. http://dx.doi.org/10.1080/14792772143000049

Hamilton, D. L., Sherman, S. J., & Rodgers, J. (2004). Perceiving the groupness of groups: Entitativity, homogeneity, essentialism, and stereotypes. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 39–60). Philadelphia, PA: Psychology Press.

Hamilton, D. L., Sherman, S. J., & Ruvolo, C. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues, 46,* 35–60. http://dx.doi.org/10.1111/j.1540-4560.1990.tb01922.x

Hamilton, D. L., Sherman, S. J., Way, N., & Percy, E. (2014). Convergence and divergence in perceptions of persons and groups. In M. Mikulincer & P. R. Shaver (Eds.), J. F. Dovidio, & J. A. Simpson (Assoc. Eds.), *APA handbook of personality and social psychology: Vol. 2. Group processes* (pp. 229–261). Washington, DC: American Psychological Association.

Hamilton, D. L., Stroessner, S. J., & Driscoll, D. M. (1994). Social cognition and the study of stereotyping. In P. G. Devine, D. L. Hamilton, & T. M. Ostrom (Eds.), *Social cognition: Impact on social psychology* (pp. 291–321). San Diego, CA: Academic Press.

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, NY: Academic Press.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology, 33,* 1–27.

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology, 28,* 280–290. http://dx.doi.org/10.1037/h0074049

Lickel, B., Hamilton, D. L., Wieczorkowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology, 78,* 223–246. http://dx.doi.org/10.1037/0022-3514.78.2.223

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

McCarthy, R. J., & Skowronski, J. J. (2011a). The interplay of controlled and automatic processing in the expression of spontaneously inferred traits: A PDP analysis. *Journal of Personality and Social Psychology, 100,* 229–240. http://dx.doi.org/10.1037/a0021991

McCarthy, R. J., & Skowronski, J. J. (2011b). What will Phil do next? Spontaneously inferred traits influence predictions of behavior. *Journal of Personality and Social Psychology, 47,* 321–332.

McConnell, A. R., Sherman, S. J., & Hamilton, D. L. (1994). On-line and memory-based aspects of individual and group target judgments. *Journal of Personality and Social Psychology, 67,* 173–185. http://dx.doi.org/10.1037/0022-3514.67.2.173

McConnell, A. R., Sherman, S. J., & Hamilton, D. L. (1997). Target entitativity: Implications for information processing about individual and group targets. *Journal of Personality and Social Psychology, 72,* 750–762. http://dx.doi.org/10.1037/0022-3514.72.4.750

Menon, T., Morris, M. W., Chiu, C. Y., & Hong, Y. Y. (1999). Culture and the construal of agency: Attribution to individual versus group dispositions. *Journal of Personality and Social Psychology, 76,* 701–717. http://dx.doi.org/10.1037/0022-3514.76.5.701

Na, J., & Kitayama, S. (2011). Spontaneous trait inference is culture-specific: Behavioral and neural evidence. *Psychological Science, 22,* 1025–1032. http://dx.doi.org/10.1177/0956797611414727

Olson, M. A. (2009). Measures of prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, sand discrimination* (pp. 367–386). New York, NY: Psychology Press.

Otten, S., & Moskowitz, G. B. (2000). Evidence for implicit evaluative in-group bias: Affect-biased spontaneous trait inference in a minimal group paradigm. *Journal of Experimental Social Psychology, 36,* 77–89. http://dx.doi.org/10.1006/jesp.1999.1399

Ramos, T., Garcia-Marques, L., Hamilton, D. L., Ferreira, M. B., & Van Acker, K. (2012). What I infer depends on who you are: The influence of stereotypes on trait and situational spontaneous inferences. *Journal of Experimental Social Psychology, 48,* 1247–1256. http://dx.doi.org/10.1016/j.jesp.2012.05.009

Ratliff, K. A., & Nosek, B. A. (2011). Negativity and outgroup biases in attitude formation and transfer. *Personality and Social Psychology Bulletin, 37,* 1692–1703. http://dx.doi.org/10.1177/0146167211420168

Runciman, W. G. (1966). *Relative deprivation and social justice.* London, UK: Routledge and Kegan Paul.

Schneid, E. D., Carlston, D. E., & Skowronski, J. J. (2015). Spontaneous evaluative inferences and their relationship to spontaneous trait infer-

ences. *Journal of Personality and Social Psychology, 108,* 681–696. http://dx.doi.org/10.1037/a0039118

Schneid, E. D., Crawford, M. T., Skowronski, J. J., Irwin, L. M., & Carlston, D. E. (2015). Thinking about other persons. *Social Psychology, 46,* 24–35. http://dx.doi.org/10.1027/1864-9335/a000218

Schneider, D. J. (2004). *The psychology of stereotyping.* New York, NY: Guilford Press.

Sherif, M. (1966). *Group conflict and cooperation: Their social psychology.* London, UK: Routledge and Keegan Paul.

Sherman, S. J., Castelli, L., & Hamilton, D. L. (2002). The spontaneous use of a group typology as an organizing principle in memory. *Journal of Personality and Social Psychology, 82,* 328–342. http://dx.doi.org/10.1037/0022-3514.82.3.328

Sherman, S. J., Hamilton, D. L., & Lewis, A. C. (1999). Perceived entitativity and the social identity value of group memberships. In D. Abrams & M. Hogg (Eds.), *Social identity and social cognition* (pp. 80–110). Oxford, UK: Blackwell.

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. http://dx.doi.org/10.1037/0022-3514.74.4.837

Spencer-Rodgers, J., Williams, M. J., Hamilton, D. L., Peng, K., & Wang, L. (2007). Culture and group perception: Dispositional and stereotypic inferences about novel and national groups. *Journal of Personality and Social Psychology, 93,* 525–543. http://dx.doi.org/10.1037/0022-3514.93.4.525

Susskind, J., Maurer, K., Thakkar, V., Hamilton, D. L., & Sherman, J. W. (1999). Perceiving individuals and groups: Expectancies, dispositional inferences, and causal attributions. *Journal of Personality and Social Psychology, 76,* 181–191. http://dx.doi.org/10.1037/0022-3514.76.2.181

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American, 223,* 96–102. http://dx.doi.org/10.1038/scientificamerican1170-96

Tajfel, H., Billig, M. G., Bundy, R. E., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology, 1,* 149–178. http://dx.doi.org/10.1002/ejsp.2420010202

Todd, A. R., Molden, D. C., Ham, J., & Vonk, R. (2011). The automatic and co-occurring activation of multiple social inferences. *Journal of Experimental Social Psychology, 47,* 37–49. http://dx.doi.org/10.1016/j.jesp.2010.08.006

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. http://dx.doi.org/10.1037/0022-3514.83.5.1051

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology, 39,* 549–562. http://dx.doi.org/10.1016/S0022-1031(03)00059-3

Todorov, A., & Uleman, J. S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology, 87,* 482–493. http://dx.doi.org/10.1037/0022-3514.87.4.482

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

Uleman, J. S., Blader, S. L., & Todorov, A. (2005). Implicit impressions. In R. R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The new unconscious* (pp. 362–392). New York, NY: Oxford University Press.

Uleman, J. S., Hon, A., Roman, R., & Moskowitz, G. B. (1996). On-line evidence for spontaneous trait inferences at encoding. *Personality and Social Psychology Bulletin, 22,* 377–394. http://dx.doi.org/10.1177/0146167296224005

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

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: An International Journal, 1,* 77–90. http://dx.doi.org/10.1016/1053-8100(92)90049-G

Uleman, J. S., Saribay, S. A., & Gonzalez, C. M. (2008). Spontaneous inferences, implicit impressions, and implicit theories. *Annual Review of Psychology, 59,* 329–360. http://dx.doi.org/10.1146/annurev.psych.59.103006.093707

Wells, B. M., Skowronski, J. J., Crawford, M. T., Scherer, C. R., & Carlston, D. E. (2011). Inference making and linking both require thinking: Spontaneous trait inference and spontaneous trait transference both rely on working memory capacity. *Journal of Experimental Social Psychology, 47,* 1116–1126. http://dx.doi.org/10.1016/j.jesp.2011.05.013

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. http://dx.doi.org/10.1037/0022-3514.84.3.470

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. http://dx.doi.org/10.1037/0022-3514.47.2.237

Winter, L., Uleman, J. S., & Cunniff, C. (1985). How automatic are social judgments? *Journal of Personality and Social Psychology, 49,* 904–917. http://dx.doi.org/10.1037/0022-3514.49.4.904

Yzerbyt, V. Y., Judd, C. M., & Corneille, O. (2004). *The psychology of group perception: Perceived variability, entitativity, and essentialism.* London, UK: Psychology Press.

Zárate, M. A., Uleman, J. S., & Voils, C. I. (2001). Effects of culture and processing goals on the activation and binding of trait concepts. *Social Cognition, 19,* 295–323. http://dx.doi.org/10.1521/soco.19.3.295.21469