Running head: SCHEMA-DRIVEN FALSE MEMORIES

Schema-Driven Source Misattribution Errors:

Remembering the Expected from a Witnessed Event

Heather M. Kleider

Georgia State University

Kathy Pezdek

Claremont Graduate University

Stephen D. Goldinger and Alice Kirk

Arizona State University

Correspondence:
Heather M. Kleider
Department of Psychology,
Georgia State University
Atlanta, GA 30030

phone: (404) 651-1623
e-mail: **hkleider@gsu.edu**

Abstract

When recollection is difficult, people may use schematic processing to enhance memory. Two experiments showed that a delay between witnessing and recalling a visual sequence increases schematic processing, resulting in stereotypic memory errors. Participants watched a slide show of a man and a woman performing stereotype-consistent and stereotype-inconsistent actions, followed by an immediate or delayed memory test. Over a two-day delay, stereotype-inconsistent actions were increasingly misremembered as having been performed by the stereotype-consistent actor (Experiment 1). All source errors increased, regardless of stereotype consistency, when the wrong actor was suggested. When we merely suggested that "someone" performed an action (Experiment 2), only stereotype-consistent source errors were increased. Although visual scenes are typically well remembered, these results suggest that when memory fades, reliance on schemata increases, leading to increased stereotypic memory errors.

Schema-Driven Source Misattribution Errors: Remembering the Expected from a Witnessed Event

In a classic study, Allport and Postman (1945) conducted a version of the "telephone game," wherein participants were shown a picture of a White man on a bus, holding a switch blade and talking with a Black man. The first participant studied the picture described it to another person, beginning a chain of other participants who described the picture to each other in succession. After the description of the picture had been passed down the chain, the final participants often reported that the Black man was holding the switch blade. At least one falsely stated that he "brandished it in a threatening manner." This finding suggested that people use stereotypes when interpreting and later recollecting information. More recently, Plant and Peruche (2005) reported that, in a computer simulation, police officers were more likely to mistakenly shoot an unarmed Black suspect than an unarmed White suspect. This suggests that, when people make quick decisions, stereotypic knowledge is readily accessed and may guide action. In an effort to understand how this type of judgment error occurs, recent research has focused on identifying the cognitive mechanisms that lead to both racial and gender stereotyping, and how such stereotyping may impact recollection (Devine, 1989; Payne, 2001; Sherman & Bessenoff, 1999). The present study examined some of the conditions wherein stereotypic expectations are likely to influence what people remember about witnessed actions or events.

Consider the classic example of a forensic eyewitness, in this case, a person who witnesses a robbery involving (at least) two perpetrators. When the eyewitness recalls the crime details, he must rely on event memory to determine what occurred, but also source memory, to delineate which suspect performed which action. When making such source attributions, people may rely on systematic, effortful processes or on relatively effortless, heuristic processes (Chaiken, Liberman & Eagly, 1989; Johnson, Hashtroudi & Lindsay, 1993; Johnson & Raye, 1981).

Systematic source monitoring involves an intentional, effortful attempt to verify that remembered information stems from a particular source. Conversely, heuristic (or "rule of thumb") processing is relatively less effortful and may occur without intention (Johnson et al., 1993; Mather, Johnson & De Leonardis, 1999). Feelings of familiarity, derived from pre-existing schemata, are frequently the basis for heuristic source monitoring judgments (Banaji & Greenwald, 1995; Hicks & Cockman, 2003; Johnson et al. 1993).

Considerable research has focused on *schematic processing*, a heuristic approach to memory encoding and retrieval that affects wide-ranging behaviors, such as stereotyping (O'Sullivan & Durso, 1984), eyewitness testimony (Greenberg, Westcott & Baliey, 1998) and mock-jury decision making (Holst & Pezdek, 1992; Smith & Studebaker, 1996). Schemata are organized collections of information stored in long-term memory that are quickly accessible and flexible in application (Hastie, 1981). Schemata help in the interpretation and retrieval of information and can be applied to individuals or events (event schemata, also known as a *script*).

There is some debate regarding whether schema-consistent or schema-inconsistent information is better remembered from a visual scene. Many researchers have reported that schema-inconsistent information has a memorial advantage (e.g., Pezdek, Whetstone, Reynolds, Askari & Dougherty, 1989), suggesting that inconsistent information attracts extra attention during encoding, relative to consistent information, and is thus better recalled. However, schematic processing also results in schema-consistent false memories (Brewer & Treyens, 1981) and judgment errors. For example, Jones (2003) presented mock-jurors with cases describing Black defendants charged either with "race-congruent" crimes (e.g. auto theft) or "race-incongruent" crimes (e.g. embezzlement). Jones reported that, given race-congruent crimes: (A) verdicts and attributions were more negative, (B) more limited memory searches were conducted, and (C) more confirmatory evidence was sought. Jones suggested that Black

defendants were more likely to be evaluated schematically, based on the race-congruence of alleged crimes, than by effortful consideration of evidence. Extending this idea from criminal trials to sentencing, Eberhardt, Davies, Purdie-Vaughns and Johnson (2006) collected prototypicality ratings to photographs of 44 Black men, all of whom were convicted of murdering White victims. Eberhardt et al. found that, when the convicts appeared more "prototypically Black," they were approximately twice as likely to receive the death penalty. This finding suggests that jurors are influenced by schematic processing, even in conditions that entail strict instructions on evidence weighting.

In the present research, we assessed whether schematic processing affects memory for actions witnessed in photographs. Our experiments included two factors – stereotypicality of action and retention interval – that may affect the saliency or quality of encoded memories, and therefore change peoples' susceptibility to schematic processing. Prior research shows that when cognitive capacity is reduced (e.g., with a memory load or secondary task), people use less effortful, schematic recollection strategies to remember information (Sherman, Macrae, & Bodenhausen, 2000). For example, Sherman and Bessenoff (1999) examined source memory for stereotype-consistent and –inconsistent actions, described in written lists. When participants performed a dual-task during the source memory test (conducted 24 hours after learning), stereotype-consistent misattribution errors were increased. In a conceptually related study, Payne (2001) had people view Black and White faces as primes, then quickly identify objects as either tools or guns. When people made these decisions under time pressure, priming with Black faces selectively increased "gun" decisions to tools. Payne suggested that race cues and time pressure combine to increase stereotypic errors, noting that these conditions are often present for eyewitnesses to real crimes.

When event details fade from memory over time, people unconsciously use schematic

processes to complete (or embellish) those faded memories. For example, Greenberg et al. (1998) reported that, after a one-week delay, participants were more likely to falsely recognize central details omitted from a robbery sequence. Similarly, Neuschatz, Lampinen, Preston, Hawkins and Toglia (2002) reported that delayed recollection of a videotaped lecture (that included schema-consistent and -inconsistent actions) resulted in more false alarms to schema-consistent actions. Together, these results suggest that, when episodic memory fades, people increasingly rely on schematic memory, leading to schema-consistent errors. Given such results, an important question is whether people have a *subjective experience* of schematic interpretation, or whether they have feelings of "true memory." Lampinen, Faries, Neuschatz and Toglia (2000; also Neuschatz, et al., 2002) used the *remember/know* procedure (Tulving, 1985) to test peoples' memory for stereotype-consistent and inconsistent actions from a story. People were more likely to experience "remembering" details (with associated thoughts and emotions) about schema-inconsistent actions (although see Tuckey & Brewer, 2003, for qualifications).

Although efficiency is conferred from using schematic memory processes, there are also disadvantages, especially in situations that require accurate, factual recall (such as eyewitness testimony). In the present study, we examined the effects of schematic processing on action and source memory for a series of slides. Both immediate and delayed testing were compared. As noted earlier, Sherman and Bessenoff (1999) found that schemata influenced source memory (especially when participants were under cognitive load). Their study, however, used linguistic stimuli, which may not directly translate to visual stimuli, which are more typical for eyewitness memory. Recognition memory is generally superior for pictures, relative to words or sentences, a well-established phenomenon known as the *picture superiority effect* (see McBride & Dosher, 2002, for a review). Potential explanations for this effect is that pictures have more distinct features, or that pictures are more closely linked to meanings, with either possibility giving them

a processing advantage (Nelson, 1979; Pezdek & Chen, 1982). The vividness of contextual detail in pictures results in richer memory traces and enhanced conceptual recollection (Serge, 1995). Given these facts, picture memory provides a stringent test for the potential role of schematic processing in memory modification.

Experiment 1

Experiment 1 was motivated by Sherman and Bessenoff's (1999) study, wherein people studied a list of actions, attributed either to a priest or a skinhead. In a later source-memory test, participants under cognitive load committed increased stereotype-consistent source errors, relative to participants without an extra cognitive load. In Experiment 1, we used a modified version of Sherman and Bessenoff's procedure, testing memory for photographs showing stereotype-consistent and stereotype-inconsistent actions by a handyman and a homeowner. We then tested memory for actions and source. In addition, following "old" judgments, participants made remember/know decisions. Rather than manipulate cognitive load, we contrasted immediate versus delayed testing, with a general hypothesis that stereotype-consistent source errors would increase as retention interval increased. Following Frost (2000; see Nemeth & Belli, 2006), we also anticipated that "remember" decisions may selectively increase for stereotype-consistent errors as the retention interval increased. By using action photographs in Experiment 1, rather than written descriptions, we intended to elicit relatively high overall recall, thus providing a strong and meaningful test of this hypothesis.

Method

Participants and Design. Ninety introductory psychology students at Arizona State University (41 males and 49 females) participated for course credit. Experiment 1 had a 2 x 2 mixed factorial design with retention interval manipulated between-subjects. Testing occurred either immediately or after two days, with 46 and 44 participants in each condition, respectively.

Presentation condition (stereotype-consistent versus inconsistent) was manipulated within-subjects.

Materials. The stimuli presented during learning included 129 digitized photographs, presented in sequence on a computer screen. They showed a woman and a handyman performing various actions, first in the kitchen and then in the backyard of a home. The man's attire (and most of his actions, e.g., arriving at the door with tools) clearly indicated that he was a handyman working at the house. The woman's casual attire (and most of her actions) indicated that she lived in the house. In the opening scene, the woman, carrying a small child, answered the front door for the handyman. The two then proceeded to the kitchen where she pointed to the sink; he looked at the drain and began working. The target sequence continued with the man performing actions that were stereotype-consistent for a handyman (e.g., tightening hinges on a cabinet, using a wrench on a faucet, sanding a pipe). In similar fashion, the woman performed stereotype-consistent actions for a homemaker (e.g., mixing cake batter, folding baby clothes, sweeping the floor). However, the sequence also included stereotype-inconsistent actions by both people. For example, the handyman opened a can of frosting, and the homemaker hung a tool belt on a chair. Figure 1 shows examples of two presented actions, showing stereotype-consistent and stereotype-inconsistent version of each.

Figure 1

Each action was shown as a series of three sequential photographs with a beginning, middle and end for the action, ensuring that it was unambiguously depicted. The "kitchen" segment included 24 target actions (12 stereotype-consistent and 12 stereotype-inconsistent) and six role-neutral filler actions (e.g., drinking a glass of water). To control for action salience, each of the 24 target actions was photographed with both the handyman and the homemaker as the actors

(see Figure 1). Across participants, the stereotype-consistent and inconsistent actions were completely counterbalanced: Two versions of the sequence were made; all participants saw all target actions, but the depicted actors were evenly divided between versions. After the critical "kitchen" sequence, all participants saw the same "backyard" sequence (with no break separating sequences). This included 13 filler actions that were all stereotype-consistent, and was intended to reinforce the stereotypic roles of both actors (the homemaker played with children; the handyman fixed play equipment). Both actors were then shown returning to the house, where the handyman wrote a bill and handed it to the woman. Next, they walked to the front door and both waved "goodbye" as the handyman departed.

The validity of stereotype-consistent and inconsistent actions was determined in a pilot study using 49 student volunteers. They were presented two identical lists of 35 actions (one list for the handyman; one list for the homemaker) and rated each on a five-point scale, with one denoting "*highly expected"* and five denoting "not at all expected." The actions selected for inclusion in the experiments (Appendix, set A) were those that received the lowest mean ratings for one actor and also the highest mean ratings for the other actor; these means ranged from 1.25 to 4.75.

Procedure. The experiment consisted of a presentation phase, followed by a distraction task and a test phase, all presented on Dell PC computers with 17-inch CRT monitors. Students participated in groups ranging from two to eight, with all volunteers seated in separate carrels. Prior to the presentation of photographs, instructions were both spoken aloud and shown on the computer. Participants were instructed that they would see a series of pictures showing two people working in a kitchen: Laura, a homemaker, and Scott, a handyman. Participants were explicitly advised to pay attention to the sequence, as their memory would be tested afterward. The presentation sequence included 129 pictures, each presented for 3 seconds, for a total of 6

minutes and 45 seconds viewing time. A 20-minute distraction phase task immediately followed the presentation phase, consisting of a face learning/recognition test (using all novel faces), an object categorization task, and number/letter discrimination task.

Following the distraction phase, the 10-minute test phase took place either immediately or after two days. Participants in the delayed condition were reminded to return for the memory test and were asked not to discuss the experiment with other participants. In the test phase, 48 written action statements were presented, one at a time, on the computer screen. All 24 target actions from the original sequence (e.g., *stirred cake batter*) were presented (see Appendix, set A). Another 24 new action statements (e.g. *used a muffin tin*) were presented as foils; these were thematically consistent with the originally studied actions (see Appendix, set B). Note that sets A and B were not counterbalanced; we only photographed the actions in set A. However, half of the new action statements were also considered typical of a handyman; the other half were considered typical of a homemaker. A written (rather than a pictorial) memory test was used to avoid ceiling effects, as recognition memory for pictures is usually quite strong, even after a delay (e.g., Nickerson, 1968).

During the test, participants first indicated (via keypress) whether each action statement was *new* or *old*. Next, for any action deemed *old*, participants indicated: (1) which actor (Scott or Laura) had performed the action, and (2) they made a *remember/know* judgment (Tulving, 1985). Participants were instructed to respond *"remember"* if they had conscious recollection of seeing the chosen actor performing the action. They were instructed to respond *"know"* when the action was familiar but they did not have a conscious (or vivid) recollection of the original actor.

Results

Analyses of old and new test actions are presented separately. The analyses of old test

actions were conducted on two measures: (a) action hit rates (regardless of source), and (b) source misattribution rates (percentages of actions correctly recognized but attributed to the wrong actor). The analyses of new test actions were conducted on two measures: (a) false-alarm rates (regardless of source), and (b) source attribution rates to the stereotype-consistent and inconsistent actors. Unless otherwise specified, analyses for hits were conducted using 2 x 2, mixed-model ANOVAs, with the within-subjects factor Presentation Type (consistent, inconsistent), and the between-subjects factor Retention Interval (immediate, delayed). Analyses of false-alarms included the factor Attribution Type, rather than Presentation Type. Simple contrasts (e.g., false-alarm rates in immediate versus delayed tests) were conducted using one-way ANOVAs. All tests assumed a standard (p <.05) significance level.

Table 1

Old actions. Hit rates (shown in Table 1) were higher in the immediate test (M = .88, SD = .11) than the delayed test (M = .82, SD =.15), $F(1, 88) = 6.30$; $MS_e = .03$; $\eta_p^2 = .07$). The main effect of Presentation Type was not significant. There was, however, a reliable interaction of Retention Interval x Presentation Type, $F(1, 88) = 7.06$; $MS_e = .02$; $\eta_p^2 = .07$). As expected, stereotype-inconsistent actions were remembered better than stereotype-consistent actions when tested immediately, $F(1, 45) = 3.97$; $MS_e = .02$; $\eta_p^2 = .04$. In the delayed test, hit rates did not differ for stereotype-consistent and inconsistent actions.

Table 2

The upper half of Table 2 shows source misattribution rates, as a function of Presentation Type and Retention Interval. A significant main effect of Retention Interval was observed, with more errors in the delayed test (M = .23, SD = .16) than in the immediate test (M = .09, SD =

.11), $F(1, 88) = 39.20$; $MS_e = .02$; $\eta_p^2 = .31$. A main effect of Presentation Type was also observed, with more source misattributions to actions performed by the stereotype-inconsistent actor ($M = .21$, $SD = .14$) than the stereotype-consistent actor ($M = .11$, $SD = .13$), $F(1, 88) = 27.26$; $MS_e = .02$; $\eta_p^2 = .24$. A significant interaction also emerged, $F(1, 88) = 9.92$; $MS_e = .02$; $\eta_p^2 = .10$. The effect of Presentation Type was unreliable in the immediate test, but was robust in the delayed test, $F(1, 43) = 22.68$; $MS_e = .02$; $\eta_p^2 = .21$. These findings suggest that, as memories faded over time, people increasingly relied on schematic knowledge in recollection.

Overall, old test items received more "remember" ($M = .55$) than "know" responses ($M = .45$). Of greater interest, the lower half of Table 2 shows the "remember" response rates that were associated with the foregoing source misattribution rates. In the present context, these rates indicate how often participants made source errors, but felt they had vivid recollection of seeing the chosen actor (see Lyle & Johnson, 2006). The main effect of Retention Interval was not reliable, although there was a tendency for more "remember" responses in the delayed test ($M = .08$) than in the immediate test ($M = .05$), $F(1, 88) = 2.27$; $p = .14$. The main effect of Presentation Type was reliable, $F(1, 88) = 23.17$; $MS_e = .02$; $\eta_p^2 = .21$, as "remember" responses were higher for actions seen with stereotype-inconsistent actors, but later attributed to stereotype-consistent actors. The potential interaction did not approach significance. In a sense, these "remember" responses can be interpreted as errors committed with high confidence. When people made errors in a stereotype-consistent direction, they were more likely to experience feelings of true recollection, a tendency that increased, at least numerically, over time.

New actions. For the new actions, we examined false-alarm rates and their associated source attributions (and remember/know decisions). As shown in Table 1, there were fewer false-alarms in the immediate test than in the delayed test, $F(1, 88) = 15.25$; $MS_e = .03$; $\eta_p^2 = .15$.

When people committed false-alarms, they also made source memory decisions and remember/know judgments, as shown in Table 3. In the source misattributions, there was no main effect of Retention Interval, but a large effect of Attribution Type was observed, with more attributions to the stereotype-consistent actor (M = .73, SD = .24) than the inconsistent actor (M = .27, SD = .24), $F(1, 88) = 172.6$; MS_e =.06; η_p^2 = .51. Thus, when people falsely recalled actions, they relied on stereotypes to determine the source. The Retention Interval x Attribution Type interaction was not reliable.

Table 3

As shown in the bottom half of Table 3, the proportions of "remember" responses were generally consistent with the source misattribution rates. A main effect of Retention Interval was observed, with more "remember" judgments in the immediate test (M = .27, SD = .12) than the delayed test (M = .16, SD = .06), $F(1, 88) = 19.41$; MS_e = .09; η_p^2 = .22. A main effect of Attribution Type was also observed, with more "remember" judgments to stereotype-consistent actors (M = .32, SD = .24) than inconsistent actors (M = .12, SD = .09), $F(1, 88) = 35.18$; MS_e = .09; η_p^2 = .26. There was also a reliable interaction, $F(1, 88) = 28.40$; MS_e = .09; η_p^2 = .23, with many "remember" judgments for consistent attributions in the immediate test, and few for inconsistent attributions in the delayed test (with the other two conditions intermediate). Taken together, the source errors and "remember" judgments suggest that schematic knowledge is used to help generate source-memory decisions, and that it also affects the phenomenological quality of those decisions, making them feel more like true memories.

Discussion

The results of Experiment 1 were relatively straight-forward, showing effects of schematic processing that increased over time. When photographic actions were studied, and memory for

those “old” actions was tested immediately, people were slightly (5%) more accurate to actions that were associated with stereotype-inconsistent actors, perhaps reflecting a salience difference between expected and unexpected events. Given a delayed test, this pattern reversed (producing a reliable interaction), such that actions associated with consistent actors were better recalled. Overall hit and false-alarm rates showed that memory faded over time, as any theory would predict. Of greater interest, this decline in memory was associated with an increase in schema-consistent processing. For both old and new actions, participants’ source misattributions were most often stereotype-consistent, a tendency that increased as retention interval increased. The same general pattern was observed in “remember” judgments that accompanied source errors; people were more likely to claim “remembering” to stereotype-consistent errors. This suggests that schematic knowledge may affect both the content and phenomenology of recall.

The results of Experiment 1 support and extend the findings of Sherman and Bessenoff (1999). Their study involved verbal learning, and used a divided-attention task to degrade participants’ episodic memories. In contrast, Experiment 1 involved pictorial learning, and we used delayed testing to degrade participants’ memories. Despite these procedural differences, both experiments revealed schema-consistent modification (or fabrication) of memories. Of particular importance, Experiment 1 suggests that stereotype-consistent errors are possible for event sequences similar to those experienced by eyewitnesses. People made stereotypic source memory errors, despite having vivid visual stimuli, an intentional (stress-free) learning situation, and relatively short delays between learning and recall. Moreover, although the rates of “remember” judgments were fairly modest, significantly more occurred when people made stereotype-consistent errors. This phenomenological quality of the memories suggests that, when people make source misattributions, they will occasionally experience feelings of “true memory.”

Experiment 2

It has been frequently demonstrated that memory can be influenced by post-event suggestions (e.g., Loftus, Miller & Burns, 1978; Pezdek, 1977). However, the vulnerability of memory to suggestion depends on various factors, including the similarity between observed and suggested memories (Lindsay, Allen, Chan, & Dahl, 2004). In Experiment 2, we examined how misleading suggestions may combine with schematic knowledge to affect memory. The prior literature is a bit unclear on this issue. For example, Roediger, Meade, and Bergman (2001) had people study visual settings (e.g., a kitchen), and then receive misleading suggestions for objects. False memories were more likely for highly schema-consistent objects (e.g., a toaster) than for less prototypical, but still consistent, objects (e.g., oven mitts). However, with a slightly changed procedure, Nemeth and Belli (2006) reported the opposite pattern. Given these divergent results, it is important to assess the misleading suggestion procedure with our materials.

Experiment 2 extends the results of Experiment 1 by testing if people are more likely to be suggestively influenced by post-event information with increased test delay, and whether they are more likely over time to be suggestively misled by stereotype-consistent than -inconsistent suggestions. Applicable to eyewitness memory situations, Experiment 2 tests whether people are likely to be misled by suggestions that a different actor performed an action, and whether source misattributions are more likely to occur if the suggested actor is stereotype-consistent with the action. In light of the finding in Experiment 1 that increased time delay resulted in stereotype-consistent source memory errors, it follows that in Experiment 2, increased time delay should also increase the acceptance rate to stereotype-consistent post-event suggestions.

In Experiment 2, the photographs and test procedures from Experiment 1 were used again, but a suggestion phase was added, just prior to testing. In the suggestion phase, old actions were presented (in written form) on the computer, but the stated actors for the actions were sometimes

changed, relative to the original photographs. For example, the sequence showed Scott hammering a nail but we later suggested that "Laura hammered a nail." Participants read and visualized five different kinds of suggestions. In a counterbalanced design, these included: (1) Old actions with suggested role-reversals, as in the foregoing example. (2) Old actions with neutral suggestions (e.g., "Someone hammered a nail"). (3) New actions, taken from Experiment 1, with stereotype-consistent suggestions (e.g., "Laura put on an apron"). (4) New actions with stereotype-inconsistent suggestions (e.g., "Scott put on an apron"). (5) New actions with neutral suggestions (e.g., "Someone put on an apron."). All new actions appear in the Appendix, set B.

The test procedure called for participants to respond "old" *only* to actions they remembered seeing in the slide sequence. This approach was meant to simulate actual interview procedures wherein witnesses may be unaware that memories can be influenced by external sources, such as other witnesses' suggestions. Improvements in source accuracy can occur when people are instructed to designate whether remembered information was retrieved from an original encoding event or from post-event information (Zaragoza & Lane, 1994). However, we chose an exclusion task because we believed it was more challenging, and more similar to the situation faced by actual eyewitnesses. Following Experiment 1, we expected that participants would rely on suggested and schematic ideas, especially given a delayed test. Accordingly, they would be susceptible to misleading suggestions, and would be particularly susceptible to suggestions that were stereotype-consistent.

Method

Participants and Design. Eighty-five introductory psychology students (34 males and 51 females) at Arizona State University participated for course credit. The design was a 2 x 2 x 2 mixed factorial: Retention Interval (immediate or two-day delay) was manipulated between-subjects; Presentation Type (stereotype-consistent and inconsistent) and Suggestion Type

(changed actor, neutral actor) manipulated within-subjects. Forty-four subjects were randomly assigned to the immediate test condition; 41 were assigned to the delayed test condition.

Materials and Procedure. The materials and procedure followed those used in Experiment 1, with the addition of a suggestion phase just prior to the memory test. Participants in the immediate condition completed the suggestion and test phases immediately after the 20-minute distraction phase. Participants in the delayed condition returned two days later to complete the suggestion and test phases.

In the suggestion phase, five types of suggestions were used. There were 48 suggested actions; each was a short sentence with a subject, verb and object. Half of the suggestions introduced new actions not in the original sequence (*suggested-only* actions); the other half referred to old actions. Among the suggested-only actions, eight (per participant) were stereotype-consistent (e.g., "Laura used the muffin tin"). Another eight were stereotype-inconsistent (e.g., "Scott used the muffin tin"), and eight were neutral (e.g., "Someone used the muffin tin"). The suggested-only actions were counterbalanced for consistency across participants. The other half of the suggestions dealt with old actions: These either paired the action with the opposite actor, or merely stated the action in neutral fashion. These suggestion types were also counterbalanced for each action across participants.

In the suggestion phase, each statement was presented on the computer for six seconds; participants were asked to visualize the stated action (and actor, if mentioned). They were told that "visualizing information sometimes helps with recall," and that the written actions may or may not have been part of the original slide sequence. Each action statement was presented only once per participant. The test was nearly the same as Experiment 1, with the singular change that participants were asked to denote "old" only those actions they remembered seeing in the

slide sequence; all other actions, including those that were read, were supposed to be classified as "new."

Results

Analyses of old and suggested-only actions are presented separately. The analyses of old actions were conducted on two measures: (a) hit rates (regardless of source), and (b) source misattribution rates (percentages of actions correctly recognized but attributed to the wrong actor). The analyses of suggested-only actions were conducted on two measures: (a) false-alarm rates (regardless of source), and (b) source attribution rates to the stereotype-consistent and inconsistent actors. As in Experiment 1, analyses for hits were conducted using mixed-model ANOVAs, with the within-subjects factors Presentation Type (consistent, inconsistent) and Suggestion Type (inconsistent, neutral), and the between-subjects factor Retention Interval (immediate, delayed). Analyses of false-alarms included the factor Attribution Type, rather than Presentation Type.

Table 4

Old actions. Hit rates (shown in Table 4) were higher in the immediate test (M = .89, SD = .14) than the delayed test (M = .80, SD =.19), $F(1, 83) = 7.33$; $MS_e = .056$; $\eta_p^2 = .08$). Neither the main effect of Presentation Type nor the Retention Interval x Presentation Type interaction were significant. Of course, absolute hit rates provide no information regarding source memory or effects of misleading suggestions. The upper half of Table 5 shows source misattribution rates: The first two rows correspond to old actions that were mentioned in neutral sentences during the suggestion phase (with results further separated by stereotype consistency and retention interval). The next two rows correspond to old actions that were associated with misleading sentences during the suggestion phase.

Table 5

A significant main effect of Retention Interval was observed, $F(1, 83) = 74.8$; $MS_e = .08$; $\eta_p^2 = .47$, with more source misattributions in the delayed test ($M = .38$, SD = .28) than in the immediate test ($M = .12$, SD = .16). There was also a main effect of Presentation Type, $F(1, 83) = 10.16$; $MS_e = .06$; $\eta_p^2 = .11$; more source errors occurred for actions presented as stereotype-inconsistent ($M = .29$, SD = .24) than stereotype-consistent ($M = .21$, SD = .20). A main effect of Suggestion Type was also observed, $F(1, 83) = 16.64$; $MS_e = .04$; $\eta_p^2 = .17$; more source errors occurred following misleading suggestions ($M = .29$, SD = .21) than neutral suggestions ($M = .20$; SD = .23). The three-way interaction (Retention Interval x Presentation Type x Suggestion Type) was also reliable, $F(1, 83) = 4.54$; $MS_e = .04$; $\eta_p^2 = .05$. This reflected differences in error patterns following different suggestion types. Given neutral suggestions, a reliable interaction of Retention Interval x Presentation Type was observed, $F(1, 83) = 12.16$; $MS_e = .04$; $\eta_p^2 = .15$, with a larger consistency effect in the delayed test (as in Experiment 1). Given misleading suggestions, the consistency effect was smaller, and constant, across tests.

The lower half of Table 5 shows "remember" response rates, organized in the same manner as the source misattribution rates. As in Experiment 1, these data represent the subsets of trials wherein people made source errors, yet felt they "truly remembered" the incorrect actor. The "remember" rates followed the same pattern as the source misattribution rates. A main effect of Retention Interval was observed, $F(1, 83) = 64.0$; $MS_e = .05$; $\eta_p^2 = .48$, with more "remember" judgments in the delayed test ($M = .21$, SD = .13) than in the immediate test ($M = .06$, SD = .08). The main effect of Presentation Type was not reliable, $F(1, 83) = 2.09$, although "remember" judgments were somewhat higher for actions presented as stereotype-inconsistent ($M = .15$, SD = .10) than stereotype-consistent ($M = .12$, SD = .10). A reliable main effect of Suggestion Type

was observed, $F(1, 83) = 7.11$; $MS_e = .05$; $\eta_p^2 = .07$, with more "remember" judgments following misleading suggestions ($M = .17$, SD = 10), relative to neutral suggestions ($M = .10$, SD = .10). Among the potential interactions, only the two-way interaction Retention Interval x Suggestion Type was reliable, $F(1, 83) = 10.88$; $MS_e = .05$; $\eta_p^2 = .09$. Given a delayed test, misleading suggestions led to more erroneous "remember" judgments, relative to neutral suggestions.

Taken together, the source errors and their associated "remember" rates suggest that, when episodic memory fades, misleading suggestions (of a changed actor) can increase overall rates of source misattribution, and also increase the phenomenological experience of "true memory." At least in the source errors, suggestions combined with pre-existing schemata, leading to rather high error rates.

Suggested-only actions. Recall that participants were instructed to respond "old" (and then to make source judgments) only to actions that were truly seen as photographs, not to actions that were merely suggested. The bottom row of Table 4 shows overall false-alarm rates in the immediate and delayed tests, with 19% more false alarms in the delayed test, $F(1, 83) = 19.78$; $MS_e = .13$; $\eta_p^2 = .11$. Of greater interest, Table 6 shows stereotype-consistent source attributions and "remember" rates, with both indices shown as proportions of total false alarms.

Table 6

The source attribution rates were analyzed in a 2 x 3 ANOVA, testing Retention Interval and Suggestion Type. It is evident from Table 6 that no effect of Retention Interval was observed (means for immediate and delayed tests were .70 and .67, respectively). The main effect of Suggestion Type was reliable, $F(1, 83) = 23.31$; $MS_e = .09$; $\eta_p^2 = .17$. As one might expect, people were more likely to make stereotype-consistent attributions following consistent ($M = .79$, $SD = .28$) or neutral suggestions ($M = .78$, $SD = .30$), relative to inconsistent

suggestions (M = .49, SD = 35). Thus, when people falsely recalled suggested-only actions as "seen," they were strongly predisposed to attribute those actions to stereotype-consistent sources, treating neutral and consistent suggestions as equivalent. Inconsistent suggestions, however, reduced participants' tendency toward schema-consistent attributions.

We analyzed "remember" response rates in the same manner. Although there was a tendency toward fewer "remember" judgments following neutral suggestions, the only reliable finding was a main effect of Retention Interval, with increased "remembering" in the delayed test, $F(1, 83) = 19.66$; $MS_e = .04$; $\eta_p^2 = .22$. Among simple contrasts, one finding was reliable: Given stereotype-consistent suggestions, people made 9% more erroneous "remember" judgments in the delayed test, relative to the immediate test, $F(1, 83) = 8.01$; $MS_e = .03$; $\eta_p^2 = .07$. These results suggest that suggestions are more likely to prompt erroneous feeling of memory when more time has passed since original episodic encoding. They also suggest that such false "remembering" is more likely when people are misled in a schema-consistent direction.

With respect to "new" actions, we conducted two contrasts across Experiments 1 and 2, testing whether the added suggestions in Experiment 2 were effective. Because the new actions were identical across experiments, any differences in overall false-alarms or source errors could be attributed to the suggestion phase. We assessed whether suggestions raised the false-alarm and source error rates, entering them into a 2 x 2 x2 mixed-model ANOVA with factors Experiment, Consistency, and Retention Interval. A main effect of Experiment was observed, $F(1, 155) = 39.2$; $MS_e = .04$; $\eta_p^2 = .20$, with more false alarms in Experiment 2 (M = .66, SD = .32) than Experiment 1 (M = .50, SD = .24). A main effect of attribution type also emerged, $F(1, 155) = 97.7$; $MS_e = .11$; $\eta_p^2 = .39$, with more false alarms to stereotype-consistent actors (M =

.76, *SD* = .26) than inconsistent actors (*M* = .38, *SD* = .32). The main effect of Retention Interval was not significant. However, an interaction of Experiment x Consistency was also observed, $F(1, 155) = 9.01$; $MS_e = .11$; $\eta_p^2 = .06$. Stereotype-consistent false alarms occurred at similar rates, irrespective of suggestions, (Experiment 1, *M* = .74, *SD* = .24; Experiment 2, *M* = .79, *SD* = .29, not a reliable difference). In contrast, stereotype-inconsistent false alarms were rare in Experiment 1 (*M* = .26, *SD* = .24), relative to Experiment 2 (*M* = .53, *SD* = .34), $F(1, 163) = 31.0$; $MS_e = .08$; $\eta_p^2 = .16$. The three-way interaction was not reliable. The cross-experiment results suggest that people generally relied on schematic knowledge to make decisions about unfamiliar events, deviating from this pattern only when given stereotype-inconsistent prompts.

Discussion

In Experiment 1, we found that schematic knowledge affected source memory for actions seen in photographs, particularly when a delay was interspersed between encoding and test. Experiment 2 was a logical extension, testing whether added suggestions would alter the observed patterns of false alarms, source errors, and the phenomenological experience of "remembering." For half of the actions that were actually seen, people received neutral suggestions (essentially reminders), stating that "somebody" performed the actions. In the other half, they received directly misleading suggestions, either changing the actor in a stereotype-consistent or inconsistent direction. The results complemented those of Experiment 1: In the immediate test, small trends were observed, wherein misleading suggestions increased source errors by 5% and "remember" responses by 3%, although neither difference was reliable. In the delayed test, however, we replicated Experiment 1, such that neutral suggestions led to more stereoptype-consistent recollection and associated "remembering." Misleading suggestions had strong effects, increasing both source memory and feeling of "remembering" for stereotype-consistent actors.

Results for the suggested-only actions were generally similar. For these new actions, the "misleading" suggestions were less blatant than those for old actions – rather than directly contradict what people saw, the suggestions were entirely new. The results showed that, when participants made false alarms, they were generally disposed to make stereotype-consistent attributions; neutral and stereotype-consistent suggestions had equivalent effects on both source memory and "remember" judgments. Stereotype-inconsistent suggestions, however, led to fewer schema-consistent judgments. Finally, when people received "directional" suggestions, either stereotype-consistent or inconsistent, the suggestions were more likely to affect source memory and phenomenological reports in the delayed test: Relative to the immediate test, suggestion-consistent source attributions increased by 7%, and "remembering" also increased by 7%, in the delayed test.

General Discussion

The present results verified the importance of person schemata in memory for visual action sequences, findings that are consistent with prior studies involving verbal stimuli (e.g., Sherman & Bessenoff, 1999). By extending these results to photographed actions, we hope to relate more directly to eyewitness memory, testing the influences of schematic processing in memory for witnessed events. In Experiment 1, we examined memory, source memory, and feelings of "remembering" for actions that were presented in stereotype-consistent or inconsistent photographs. For actions that were actually seen, source errors in the immediate memory test were fairly rare, with a tendency for more errors in favor of stereotype-consistent actors. When people received the memory test after a two-day delay, the stereotype consistency effect grew stronger. The same pattern occurred in "remember" judgments. When people committed false alarms to new actions, their source attributions and "remember" judgments were strongly influenced by stereotypes, with little change over retention intervals. Taken together, the results

of Experiment 1 suggest that, even with encoding of rich visual stimuli, people tend to make stereotype-consistent source errors. This tendency increases over time, as original memories presumably fade. Moreover, as retention intervals increase, such false source attributions are increasingly experienced as "true memories," at least according to participants' *remember/know* classifications.

In Experiment 2, we extended Experiment 1, testing whether reliance on schematic knowledge interacts with memory suggestibility. For actions that were actually seen, a suggestion phase either restated the actions in neutral contexts (e.g., *"Someone hammered a nail"*), or in misleading contexts that changed the original actors. Given an immediate memory test, people were slightly more likely to commit stereotype-consistent source errors (and to report stereotype-consistent "remembering"), regardless of suggestions. In the delayed test, people were equally likely to make suggestion-consistent source errors, without any clear influence of stereotype consistency (although all trends still favored consistent attributions). Given neutral suggestions, error patterns replicated Experiment 1, with stereotype-consistent source errors increasing over retention intervals. Thus, when original memories faded, people were most influenced by the most recent information available, which were the suggestions. But, when those suggestions were neutral with respect to sources (suggesting only "someone"), people again relied upon schemata. Indeed, for new actions, neutral and stereotype-consistent suggestions were functionally equivalent, indicating that stereotype-consistency is the default mode for source attributions.

Across both experiments, participants were more likely to indicate "remembering" stereotype-consistent sources, although this effect was suppressed when people were given misleading suggestions supporting the inconsistent actors. In the literature, several models have been proposed to account for the recollective experience of false memories. For example,

Lampinen, Meier, Arnal, and Leding (2005) proposed a *content borrowing* model wherein experiential details are "borrowed" from actual events to provide corroboration of false memories (see also Henkel & Franklin, 1998, for a reality-monitoring theory). According to Lampinen et al., such borrowing occurs when activated familiar events systematically bias search processes for episodic details. If such details can be recalled, they intensify feelings of familiarity. An alternative hypothesis is that people "overweigh" schematic information (Mitchell, Johnson & Mather, 2003). According to this view, if stereotyping is indeed automatic, as some studies suggest (e.g., Macrae, Bodenhausen, Schloerscheidt, & Milne, 1999; Wegner & Bargh, 1998), then the fluency with which stereotypic information comes to mind may be interpreted as a cue to recollection. This application of the *fluency heuristic* (Whittlesea & Leboe, 2000) would systematically give greater memorial weight to stereotype-consistent source attributions. Either account would easily predict the present results, as schema-consistent false source attributions were associated with higher levels of phenomenological remembering. Such feelings of memory could arise from borrowing features of encoded photographs or from general availability of schematic knowledge.

Brainerd, Wright, Reyna, and Mojardin (2001) described a process wherein *fuzzy trace theory* (Reyna & Brainerd, 1995) can create false recollective experiences. In the present context, we are interested in how participants falsely "remember" seeing stereotype-consistent actors performing actions, especially in trials when they had actually seen the opposite. According to fuzzy trace theory, episodic memory traces vary on a continuum from gist to verbatim. Gist traces represent the general meaning of items; verbatim traces represent more specific and detailed information. The experience of recognition memory can arise either from retrieval of a verbatim trace or retrieval of gist traces that closely match the test item. In similar fashion, false memory can arise by retrieving gist representations that closely match foils. A

false experience of recollection occurs when the underlying gist trace is quite strong and is an especially good match for the presented foil. Sometimes people will "borrow" features from a similar perceived event, importing those features into a false event. This makes the embellished false memory feel very similar to bona fide memories (Lampinen et al., 2005; Lyle & Johnson; 2006). Given the results of the present study, we suggest that such feature-borrowing is more prevalent and/or convincing when schematic knowledge is activated. Thus, the presentation of written action statements (e.g., stirring cake batter) activates prior traces, which will often be laden with schema-consistent features (e.g., "Mom in the kitchen"). The more readily available such schemata are, the easier it may be to assemble features into convincing false memories.

Interpreting our results according to fuzzy trace theory, participants given an immediate memory test should be more likely to retrieve verbatim traces when presented with action statements. Therefore, as we observed, they should fairly accurate in both action and source memory. However, given a delayed test, verbatim traces should either decay or become less available, leading to greater reliance on gist traces. This reliance on gist traces should make people more susceptible to highly available, schematic knowledge. If so, we would expect inflated stereotype consistency effects in a delayed test, along with increased experiences of "remembering."

Although pictorial stimuli are typically well remembered (McBride & Dosher, 2002), the present results suggest that, when episodic memories for photographs fade, people will still rely on schemata to interpret test targets and foils. Although unexpected information in visual scenes is typically remembered with greater detail, relative to expected information (Pezdek et al., 1989, Neuschatz et al., 2002), we consistently found more source memory errors were higher for actions performed by schema-inconsistent actors. This observation has important implications for eyewitness memory. In a controlled memory study such as the present one, conditions are

highly favorable for learning, relative to conditions in typical eyewitness situations. In typical crimes, there are frequently multiple sources of distraction, brief exposure times, and high stress levels (Morgan, Hazlett & Doran, 2004; Valentine, Pickering & Darling, 2003). Given that people relied upon schemata in the present, nearly ideal conditions suggests that schema-consistent source errors may be especially prevalent for eyewitnesses to real crimes. These findings extend the observations of Holst and Pezdek (1992; also Greenberg, Westcott, & Bailey, 1998), who reported that eyewitnesses possess scripts for typical crimes. Because witnesses activate these scripts when recalling details of specific crimes, schema-consistent memory distortions can occur.

People naturally rely on schemata in everyday life to organize and remember ideas and events. However, memorial accuracy is occasionally paramount, as when a person witnesses a crime. In such occasions, schematic retrieval strategies are likely to be detrimental, and may systematically harm members of frequently stereotyped groups. The present results suggest that such stereotypic errors may arise quite easily, especially when memory is assessed after the passage of time.

References

Allport, G.W. & Postman, L. (1945). The basic psychology of rumor. *Transactions of the New York Academy of Sciences, 11,* 61-81.

Banaji, M.R., & Greenwald, A.G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology, 68,* 181-198.

Bower, G.H., Black, J.B., & Turner, T.J. (1979). Scripts in memory for text. *Cognitive Psychology, 11,* 177-220.

Brainerd, C.J., Wright, R., Reyna, V.F., & Mojardin, A.H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 307-327.

Brewer, W.F. & Treyens, J.C. (1981). Role of schemata in memory for places. *Cognitive Psychology,* 13, 207-230.

Chaiken, S., Liberman, A., & Eagly, A.H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. In J. Uleman & J. Bargh (Eds.), *Unintended Thought* (pp. 212-252). New York: Gilford Press.

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

Eberhardt, J.L., Davies, P.G., Purdie-Vaughns, V.J., & Johnson, S.L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science, 17,* 383-386.

Frost, P. (2000). The quality of false memory over time: Is memory for misinformation "remembered" or "known?" *Psychonomic Bulletin & Review, 7,* 531-536.

Greenberg, M.S., Westcott, D.R., Bailey, S.E. (1998). When believing is seeing: The effect of scripts on eyewitness memory. *Law and Human Behavior, 22*: 685-694.

Hastie, R. (1981). Schematic principles in human memory. In E. Higgins E., C. Herman, & M. Zanna (Eds), *Social Cognition: The Ontario Symposium (Vol.1),* pp. 39-88. Hillsdale, NJ: Erlbaum.

Heit, E. (1993). Modeling the effects of expectations on recognition memory. *Psychological Science, 4,* 244-252.

Henkel, L.A. & Franklin, N. (1998). Reality monitoring of physically similar and conceptually related objects. *Memory & Cognition, 26,* 659-673.

Hicks, J.L. & Cockman, D.W. (2003). The effect of general knowledge on source memory and decision processes. *Journal of Memory and Language, 48,* 489-501.

Holst, V.F., & Pezdek, K. (1992). Scripts for typical crimes and their effects on memory for eyewitness testimony. *Applied Cognitive Psychology, 6,* 573-587.

Johnson, M.K., Hashtroudi, S., & Lindsay, D.S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3-28.

Johnson, M.K., & Raye, C.L. (1981). Reality monitoring. *Psychological Review, 88,* 67-85.

Johnson, W.A., Hawley, K.J. (1994). Perceptual inhibition of expected inputs: The key that opens closed minds. *Psychonomic Bulletin & Review, 1,* 56-72.

Jones, C.S. (2003). The effects of racially stereotypical crimes on juror decision-making and information-processing strategies. *Basic and Applied Social Psychology, 25,* 1-13.

Lampinen, J.M., Faries, J.M., Neuschatz, J.S. & Toglia, M.P. (2000). Recollections of things

schematic: The influence of scripts on recollective experience. *Applied Cognitive Psychology, 14*, 543-554.

Lampinen, J.M., Meier, C.R., Arnal, J.D., & Leding, J.K. (2005). Compelling untruths: Content borrowing and vivid false memories. *Journal of Experimental Psychology*, 31, 954-963.

Lindsay, D.S., Allen, B.P., Chan, J., & Dahl, L.C. (2004). Eyewitness suggestibility and source similarity: Intrusions of details from one event into memory reports of another event. *Journal of Memory and Language, 50*, 96-111.

Loftus, E.F., Miller, D.G. & Burns, H.J. (1978). Semantic integration of verbal information into visual memory. *Journal of Experimental Psychology: Human Learning & Memory, 4*, 19-31.

Lyle, K.B., & Johnson, M.K. (2006). Importing perceived features into false memories. *Memory 14*, 197-213.

Mather, M., & Johnson, M.K. (2003). Affective review and schema reliance in memory in older and younger adults. *American Journal of Psychology, 116*, 169-189.

Mather, M., Johnson, M.K. & De Leonardis, D.M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology, 16,* 437-458.

Macrae, C.N., Bodenhausen, G.V., Schloerscheidt, A.M., & Milne, A.B. (1999). Tales of the unexpected: Executive function and person perception. *Journal of Personality and Social Psychology, 76*, 200-213.

Macrae, C.N., Schloerscheidt, A.M., Bodenhausen, G.V., & Milne, A.B. (2002). Creating memory illusions: Expectancy-based processing and the generation of false memories.

Memory, 10, 63-80.

McBride, D.M., Dosher, B.A. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition: An International Journal, 11,* 423-460.

Mitchell, K.J., Johnson, M.K., & Mather, M. (2003). Source monitoring and suggestibility to misinformation: Adult age-related differences. *Applied Cognitive Psychology, 17,* 107-119.

Morgan, C.A. III, Hazlett, G., Doran, A. (2004). Accuracy of eyewitness memory for persons encountered during exposure to highly intense stress. *International Journal of Law and Psychiatry, 27,* 265-279.

Nakamura, G.V., Graesser, A.C., Zimmerman, J.A., & Riha, J. (1985). Script processing in a natural situation. *Memory & Cognition, 13,* 140-144.

Nelson, D.L. (1979). Remembering pictures and words: Appearance, significance, and name. In L. Cermak & F. Craik (Eds.), *Levels of Processing in Human Memory* (pp. 45-76). Hillsdale, NJ: Erlbaum.

Nemeth, R.J. & Belli, R. F. (2006). The influence of schematic knowledge on contradictory versus additive misinformation: False memory for typical and atypical items. *Applied Cognitive Psychology, 20,* 563-573.

Neuschatz, J.S., Lampinen, J.M., Preston, E.L., Hawkins, E.R., & Toglia, M.P. (2002). The effect of memory schemata on memory and the phenomenological experience of naturalistic situations. *Applied Cognitive Psychology, 16,* 687-708.

Nickerson, R.S. (1968). A note on long-term recognition memory for pictorial material.

Psychonomic Science, 11, 58.

O'Sullivan, C.S., Durso, F.T. (1984). Effect of schema-incongruent information on memory for stereotypical attributes. *Journal of Personality and Social Psychology, 47*, 55-70.

Payne, B.K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 81*, 181-192.

Pezdek, K. (1977). Cross-modality semantic integration of sentence and picture memory. *Journal of Experimental Psychology: Human Learning and Memory, 3*, 515-524.

Pezdek, K., & Chen, H-C. (1982). Developmental differences in the role of detail in picture recognition memory. *Journal of Experimental Child Psychology, 33*, 207-215.

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N. & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15*, 587-595.

Plant, E.A., & Peruche, B. (2005). The consequences of race for police officers' responses to criminal suspects. *Psychological Science*, *16*. 180-183.

Reyna, V.F., & Brainerd, C.J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1-75.

Roediger, H.L., III, Meade, M.L., & Bergman, E.T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8*, 365-371.

Serge, N. (1995). The picture-superiority effect in category-association tests. *Psychological Research, 58*, 218-224.

Sherman, J.W., & Bessenoff, G.R. (1999). Stereotypes as source-monitoring cues: On the

interaction between episodic and semantic memory. *Psychological Science, 10,* 106-110.

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

Smith, V.L., & Studebaker, C.A. (1996). What do you expect? The influence of people's prior knowledge of crime categories on fact-finding. *Law and Human Behavior, 20,* 517-532.

Tuckey, M.R., & Brewer, N. (2003). The influence of schemas, stimulus ambiguity, and interview schedule on eyewitness memory over time. *Journal of Experimental Psychology: Applied, 9,* 101-118.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist, 26,* 1-12.

Valentine, T., Pickering, A., & Darling, S. (2003). Characteristics of eyewitness identification that predict the outcome of real lineups. *Applied Cognitive Psychology, 17,* 969-993.

Wegner, D.M., & Bargh, J.A. (1998). Control and automaticity in social life. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), Handbook of social psychology (4th ed, Vol. 1, pp. 446-496). New York: McGraw-Hill.

Whittlesea, B.W.A., & Leboe, J. (2000). The heuristic basis of remembering and classification: Fluency, generation and resemblance. *Journal of Experimental Psychology: General, 129,* 84-106.

Zaragoza, M.S. & Lane, S.M. (1998). Processing resources and eyewitness suggestibility. *Legal and Criminological Psychology, 3,* 305-320.

Author Note

This research was supported by NRSA post doctoral grant F32-MH63521-01 to Heather M. Kleider, conducted under the supervision of Kathy Pezdek at Claremont Graduate University. We thank Iris Blandon-Gitlin, Karen Mitchell and an anonymous reviewer for helpful comments, and Adrianna Orozco for help in data collection.

Table 1. Hit rates (and standard deviations) in Experiment 1, as a function of presentation type and retention interval. False-alarm rates (and standard deviations) in Experiment 1, as a function of retention interval.

	Retention Interval	
	Immediate Test	Delayed Test
Hits:		
Stereotype Consistent	.86 (.13)	.84 (.14)
Stereotype Inconsistent	.91 (.11)	.81 (.15)
False Alarms:	.20 (.16)	.34 (.17)

Table 2. Source misattribution and associated "remember" response rates (with standard deviations) for old actions in Experiment 1, as a function of presentation type and retention interval.

Source Misattributions:

Actor Seen / Chosen	Immediate Test	Delayed Test
Consistent / Inconsistent	.07 (.11)	.14 (.15)
Inconsistent / Consistent	.11 (.11)	.30 (.16)

"Remember" Responses:

Actor Seen / Chosen	Immediate Test	Delayed Test
Consistent / Inconsistent	.03 (.03)	.04 (.05)
Inconsistent / Consistent	.07 (.08)	.11 (.09)

Notes: (1) Source misattributions are expressed as proportions of hits in each condition, as shown in Table 1. (2) "Remember" responses are also expressed as proportions of hits per condition, and are thus unaffected by base-rate differences in source misattributions across conditions.

Table 3. Source misattribution rates and associated "remember" response rates (with standard deviations) for new actions in Experiment 1, as a function of retention interval.

Source Misattributions:

Actor Chosen	Immediate Test	Delayed Test
Stereotype Consistent	.72 (.28)	.75 (.19)
Stereotype Inconsistent	.28 (.28)	.25 (.19)

"Remember" Responses:

Actor Chosen	Immediate Test	Delayed Test
Stereotype Consistent	.37 (.11)	.26 (.06)
Stereotype Inconsistent	.17 (.12)	.06 (.06)

Notes: (1) Source misattributions are expressed as proportions of false-alarms in each condition, as shown in Table 1. (2) "Remember" responses are also expressed as proportions of false-alarms per condition, and are thus unaffected by base-rate differences in source errors across conditions.

Table 4. Hit rates (and standard deviations) in Experiment 2, as a function of presentation type and retention interval.

	Retention Interval	
	Immediate Test	Delayed Test
Hits:		
Stereotype Consistent	.89 (.12)	.80 (.19)
Stereotype Inconsistent	.88 (.15)	.80 (.19)
False Alarms:	.29 (.23)	.48 (.25)

Table 5. Source misattribution and associated "remember" response rates (with standard deviations) for old actions in Experiment 2, as a function of suggestion type, presentation type, and retention interval.

Source Misattributions		
Neutral Suggestions:		
Actor Seen / Chosen	Immediate Test	Delayed Test
Consistent / Inconsistent	.07 (.13)	.22 (.25)
Inconsistent / Consistent	.11 (.15)	.41 (.31)
Misleading Suggestions:		
Actor Seen / Chosen	Immediate Test	Delayed Test
Consistent / Inconsistent	.11 (.17)	.42 (.27)
Inconsistent / Consistent	.17 (.20)	.46 (.29)
"Remember" Responses		
Neutral Suggestions:		
Actor Seen / Chosen	Immediate Test	Delayed Test
Consistent / Inconsistent	.04 (.07)	.13 (.14)
Inconsistent / Consistent	.06 (.06)	.17 (.13)
Misleading Suggestions:		
Actor Seen / Chosen	Immediate Test	Delayed Test
Consistent / Inconsistent	.05 (.08)	.26 (.11)
Inconsistent / Consistent	.08 (.09)	.27 (.12)

Notes: (1) Source misattributions are expressed as proportions of hits in each condition, as shown in Table 1. (2) "Remember" responses are also expressed as proportions of hits per condition, and are thus unaffected by base-rate differences in source misattributions across conditions.

Table 6. Source misattribution and "remember" response rates (with standard deviations) for suggested-only actions in Experiment 2, as a function of suggestion type, attribution type, and retention interval. Only stereotype-consistent attributions are shown.

	Immediate Test	Delayed Test
Stereotype-Consistent Suggestions		
Consistent Attributions:	.75 (.33)	.83 (.23)
"Remember" Responses:	.14 (.12)	.23 (.18)
Neutral Suggestions		
Consistent Attributions:	.84 (.30)	.72 (.30)
"Remember" Responses:	.11 (.11)	.15 (.19)
Stereotype-Inconsistent Suggestions		
Consistent Attributions:	.52 (.36)	.46 (.34)
"Remember" Responses:	.13 (.13)	.18 (.17)

Notes: Source attribution and "remember" rates are expressed as proportions of overall false-alarms in the immediate and delayed tests.

Appendix: Test Action Statements Presented in Experiments 1 and 2

Set A	Set B
sprays pan with oil	washes a bowl with scrubber
opens cake mix	polishes a bowl
cracks egg into a bowl	gets a muffin tin
puts on work gloves	turns on a mixer
uses a hammer	pours oil into a measuring cup
uses a level on the kitchen counter	stirs batter with a spoon
puts c-clamp on kitchen counter	folds dinner napkins
extends hose to back of fridge	pours drink into child's cup
reads in cookbook	washes a window
takes child's cup from fridge	makes a sandwich
pours cake batter into pan	puts on an apron
opens can of frosting	takes eggs out of refrigerator
plugs drill into wall	twists pipe apart
uses tape measure on window	measures the sink with tape measure
sweeps the floor	uses a socket set
hangs tool belt on a chair	hands over a business card
folds baby clothes	locks the tool box
uses a wrench on the faucet	wipes grease from hands on a work rag
uses sandpaper on a pipe	uses oil on the cabinet hinge
takes measuring cups from drawer	removes cover from kitchen electrical outlet
puts cake into oven	drills a hole into the counter
tightens hinge on cabinet	gets a small saw from tool box
uses silicone/latex on the sink	uses plumber's "snake" in the drain
gets baby bibs from a drawer	sands cabinet door

Figure Caption

Figure 1. Examples of photographs showing stereotype-consistent and stereotype-inconsistent version of two presented actions. Photographs were shown in color during the experiments.

Figure 1: