"I Think He Had A Tattoo On His Neck": How Co-Witness Discussions About A Perpetrator's Description Can Affect Eyewitness Identification Decisions.

Mitchell L. Eisen

California State University, Los Angeles

Fiona Gabbert

Goldsmiths University, London

Rebecca Ting Joseph Williams

California State University, Los Angeles

Acknowledgements: We would like to acknowledge Patricia Arrendondo, Taly Barak, Cynthia Perez, Lindsay Wandry, Amber Ritter, and Orbelin Montes for there work in collecting the data.

All correspondence reading this paper should be directed to Professor Mitchell Eisen, meisen@calstatela.edu.

Abstract

This experiment was designed to examine the effect of misinformation imparted through co-witness discussions on memory reports and line-up decisions obtained after varied retention intervals. Two-hundred and eighty-nine participants viewed a simulated car-jacking and then heard co-witnesses describe their memory for the event. Confederate accounts included three plausible and three implausible pieces of misinformation. Memory for the event was assessed after five-minutes, 50-minutes, two-days, or one-week. In addition to examining free-recall memory, we also looked at how misinformation about the perpetrator's appearance affected recognition memory by obtaining identifications from culprit-present and absent lineups. One of the confederates falsely described the perpetrator having a tattoo on his neck, and one lineup filler had this feature. Results revealed that mistaken identifications of the tattooed filler increased significantly at the longer retention intervals, while recall for the misinformation decreased at the longer intervals. Also, as expected, plausible misinformation was recalled more often than implausible.

In the early morning hours of a summer night in Southern California, Brenda J. and her friend were coming back from a night of bar hopping and were confronted by two men who demanded her car keys. After a brief struggle, one of the assailants took her purse and fled. The police interviewed the witnesses separately. Brenda recalled that the assailant had some type of undefined tattoos on his face, but her friend remembered him having the letters tattooed on his head. In the days after the event, the two friends talked more about the shared experience, and searched social media together, looking for pictures of tattoos similar to what they remembered. After reviewing dozens of pictures together, the two women arrived at the shared conclusion that the culprit had letters tattooed on his face (a merging of the two witnesses' memory reports), and the women believed they may have even found a picture of the culprit. The witnesses then went back to the police with this new information, and the police immediately developed a suspect named Richard Torres who had letters tattooed on his face, similar to what the victims reported. The police then put Torres' picture in a sequential line-up with five other individuals, none of whom had tattoos on their face, and administered it to both witnesses using double-blind procedures. Not surprisingly, both witnesses picked Torres from the line-up, and each noted the importance of the letter tattoos in making their decision (*People v. Richard Torres*, San Diego Superior Court; February, 2015, CD256364).

Surveys of real eyewitnesses to crimes reveal that co-witnesses like Brenda J. and her friend frequently talk together about their shared experiences (Paterson & Kemp, 2006; Skagerberg & Wright, 2008). Moreover, because witnesses' descriptions of perpetrators tend to vary after a crime (Gabbert & Brown, 2015), it is reasonable to assume that witnesses who participate in these post-event discussions are frequently exposed to inaccurate information about a perpetrator's appearance. A growing body of empirical research shows that, when people talk about their memories, they can influence each other such that their subsequent individual memory reports become similar to one another, as demonstrated in the Torres case (for reviews, see

Gabbert & Hope, 2013; Wright, Memon, Skagerberg, & Gabbert, 2009). This phenomenon is typically referred to as *memory conformity* (Wright, Self, & Justice, 2000).

The majority of memory conformity research has focused on examining how misleading information encountered during co-witness discussions is subsequently reported in an interview or written statement (Gabbert & Hope, 2013). However, the extent to which co-witness discussion can influence subsequent facial recognition decisions is relatively under-researched. Since recall and recognition rely on different underlying psychological processes (e.g., Schooler, 2002; Wells, 1985), it cannot be assumed that both are susceptible to memory conformity effects to the same extent. Zajac and Henderson's (2009) study is one of the few that has examined how co-witness discussions could directly influence identifications from lineups. Here, participants viewed a video of a staged theft, after which half of the participants were misinformed by a confederate that the thief's accomplice had blue eyes (when they were actually brown). Misinformed participants were significantly more likely than controls to describe the accomplice as having blue eyes, and were twice as likely to misidentify a blue-eyed suspect from a culprit-absent lineup.

Zajac and Henderson's (2009) findings can potentially be explained by recoding, or retrieval-based, interference, whereby faces, when described, are translated from visual to verbal information (see also Alogna et al., 2014; Meissner, Brigham & Kelly, 2001). If the verbal memory contains inaccurate information, and is relied upon during the identification test, then the participant will be less likely to correctly identify the target. For example, Meissner, Sporer, and Susa (2008) found that errors in descriptions can interfere with a witness's later ability to make an accurate identification decision. Unfortunately, because Zajac and Henderson only used target-absent lineups, they could not examine how this misinformation affected the witnesses' ability to recognize the target. The current experiment addresses this issue by looking at the how suggested misinformation encountered from co-witness discussions could lead witnesses to select a specific suspect who possesses a suggested trait in both culprit-present and culprit-absent conditions.

The current study also seeks to extend Zajac and Henderson's (2009) work by examining the effect of the misinformation on witness performance after varied retention intervals, ranging from five-minutes to a week. The retention interval in the Zajac and Henderson experiments was quite brief, whereas lineup identifications in actual cases are often made following much longer delays. Previous research has demonstrated the importance of accounting for longer retention intervals when studying eyewitness memory (Deffenbacher, Bornstein, McGorty, & Penrod, 2008). This work has shown that descriptions of perpetrators provided after longer delays tend to be less complete (Van Koppen & Lochun, 1997), and has also revealed a linear decline in the correct identifications of previously seen faces after longer delays (Shapiro & Penrod, 1986). Moreover, recent research has shown that even very modest alterations in the retention-interval between verbal descriptions and identifications can lead to significant differences in verbal overshadowing effects (Alogna et al., 2014). Since face memory tends to decline over time, we hypothesized that the misinformation encountered during co-witness discussions related to the culprit's appearance would have a greater effect on witnesses' lineup decisions made at longer retention intervals, as memory for the perpetrator's face fades.

Many factors can influence whether a suggested detail is accepted and later reported, for example, not detecting the discrepancy and accepting it without question (Blank, 1998; Loftus, 2005), detecting it but then making a source error (Horry, Colton, & Williamson, 2014), and detecting the discrepancy but reporting it anyway due to believing it is likely to be correct (Allan & Gabbert, 2013). Since discrepancies in plausible post-event misinformation are less likely to be detected, plausible suggestions are more likely to be believed than non-plausible items (Berntsen & Rubin, 2007; Pezdek & Blandon-Gitlin, 2009). Furthermore, people are more likely to accept misleading post-event details that concern schema-relevant items (i.e., plausible), in comparison to low expectancy items that are more salient (Meade & Roediger, 2002; Walther, et.al., 2002). Scoboria et al. (2012) explains that suggested events that are detected (remembered) and judged to be implausible are often evaluated and dismissed quite rapidly. However, Scoboria (2012) also

argued that, although the information that is judged to be personally implausible is *less* likely to lead to memory change, the threshold for this effect may be lower than previously thought, such that even suggestions that are judged to be minimally plausible can lead to memory change under some circumstances.

The Current Study

The current study examined how suggested information about the perpetrator's appearance (i.e., having a tattoo on his neck) affected identifications made from photo arrays after varied retention intervals ranging from five-minutes to one-week. It was predicted that participants who heard the suggestion that the culprit had a tattoo on his neck would be more likely to misidentify the filler who possessed this feature from a lineup, particularly at longer retention intervals, when memory for the perpetrator's face has faded. The current study also examined participants' free recall memory for plausible and implausible suggested misinformation assessed after varied retention intervals ranging from five-minutes to one-week. Consistent with previous research, we predicted that witnesses would be less likely to be influenced by misinformation that is considered to be implausible.

Method

Participants

Participants were 289 introductory psychology students at a university in Southern California (69.6% female). The age of the participants ranged from 18 to 63 (M = 20.1, SD = 3.73). The racial background of participants varied, with 53.3% identifying themselves as Latino, 23.5% Asian, 6.6% Anglo/Caucasian, and 5.9% African American. The remaining 9.0% of participants identified themselves as being mixed or other.

Procedure

Sessions involved five co-witnesses: three participants and two confederates. The five co-witnesses were seated around a table and in front of a 30" television monitor. Participants first completed a consent form where they were informed this was a study about memory. They then

viewed a video of a simulated car-jacking. Participants were instructed to pay close attention to the details of the crime simulation, and were told that they would be tested about their memory for the event and the people involved.

The video. The carjack video lasted one minute and seven seconds. The first 50 seconds of the video showed a Latino Male in his 20s with a shaved head on his cell phone in the forefront of the screen (the perpetrator) watching two women walking towards their cars. The man on his cell phone was approximately 25 feet from the camera, and he was standing by an SUV to give witnesses a clear reference point for his height and weight. His exposed neck was clearly visible from all angles, as he nervously shifted his position relative to the camera while watching the women. The two women hugged each other, and one of them got into her car and drove away. The second woman (the victim) then started to enter her car. As she opened her car door, the perpetrator pulled a gun from his waistband, ran towards her, demanded her keys, and stole her car.

Sharing statements after the event: Conveying the misinformation. After viewing the video, participants were instructed to treat the event as if they were a witness to an actual crime in a police investigation, and were asked to provide eyewitness statements that would assist the police in catching the culprit. Each participant was asked to state aloud what they remembered from the video, and were asked to speak one at a time so the experimenter could write down everything they said verbatim. They were specifically asked to make sure to include all details regarding the perpetrator's appearance such as height, weight, age, race, hairstyle, dress, as well as other defining features such as scars, tattoos, and jewelry that would assist the police in capturing the culprit. Participants were instructed not to comment on the statements made by the others in the group or to interrupt others while they were talking. This was done to ensure that the confederates' memory reports could be communicated without interruption, and to make sure that the misinformation provided within these reports was not publicly challenged or contradicted.

Participants were seated at the table and spoke in turns from left to right. The confederates were always positioned on either end of the group (in the first and last chairs). This seating arrangement ensured that the confederates would always give their accounts first and last. The confederates memorized a script describing a fairly comprehensive account of the event that included 14 pieces of correct information relating to the perpetrator's description and actions, and the victim's description and actions. The script also included three key errors for each confederate; thus, participants were exposed to six items of misinformation in total (here forward, *the misinformation*). All statements made by the participants and confederates in these sessions were recorded verbatim.

The misinformation. The six misinformation items were designed to be related to plausible or implausible aspects of the event. Five of the six misinformation items were related to descriptive features of the perpetrator's appearance that would have clear forensic relevance in an investigation: height, race, hairstyle, clothing, and the presence of tattoos. The sixth piece of misinformation was related to whether the perpetrator was smoking a cigarette and snuffed it out at the scene, which might provide a way to obtain physical trace evidence.

The three plausible misinformation items included the following suggestions: (1) the perpetrator was over six feet tall, possibly 6'4" or 6'5" (when he was actually 5'10"); (2) he had on long pants (when he was actually wearing shorts); and (3) he had a tattoo on his neck (when he did not). The three implausible misinformation items included the following suggestions: (1) the perpetrator had short hair – an inch to an inch and a half long (when he was actually bald); (2) the culprit was African-American (when he was actually Latino); and (3) he flicked a cigarette away (when he actually was not smoking and was on a cell phone the entire time).

The counterbalancing of confederate statements. The presentation of the misinformation by the confederates was counterbalanced across groups by changing the positioning of confederates at the table, so they alternated in presenting their memory reports first and last. One confederate reported recalling the three plausible errors, and the other reported the

three implausible errors. This allowed us to counterbalance the presentation of the plausible and implausible information. Multiple research assistants rotated roles as confederates.

The experimenter recorded all witness statements verbatim to ensure the misinformation was properly reported, and also to document how often participants repeated the misinformation during this initial group recall session. All analyses controlled for the order of the presentation of the misinformation (if the misinformation in question came before or after each participant spoke). Also, all analyses controlled for whether the misinformation was repeated by the participant and/or anyone else in their small group during this initial co-witness session.

Retention intervals. Participants were randomly assigned to one of four experimental groups, where the delay between encountering the misinformation and reporting memories of the event was either five minutes, 50 minutes, 48 hours, or one week. A fifth (control) group did not encounter any misinformation, and was tested after a five-minute delay. Memory reports were obtained using a free recall format, followed by specific open-ended prompts. Specifically, participants were first asked to report everything they remembered about the carjack video, including everything they mentioned when providing verbal statements right after the event in the small group sessions, as well as anything else they may have remembered that they did not recall and/or report earlier. The open-ended prompts then encouraged participants to describe specific elements of the perpetrator's appearance, including his race, type and color of clothes, color and length of hair, height, defining features (any scars, tattoos, jewelry, etc.), and actions. These non-suggestive open-ended prompts were related to the six-misinformation items (clothes, height, tattoos, race, hair style, and actions). Again, participants were asked to treat this as though they were witnesses to an actual police investigation.

Identifying the culprit from the photo array. After completing the memory questionnaire, the participants were administered either a culprit present or absent six-person photographic lineup. As noted earlier, one of the confederates had suggested that the culprit had a tattoo on his neck during the small group co-witness session, and one of the fillers in the lineup

had a tattoo photoshopped onto his neck. Before viewing the lineup, participants were read the pre-lineup admonishment used by the Los Angeles Police Department.

> I am going to show you a group of photographs. This group may or may not contain a picture of the person who committed the crime now being investigated. Keep in mind that hairstyles, beards, and mustaches may be easily changed. Also keep in mind that photographs may not always depict the true complexion of the person; it may be lighter or darker than shown in the photo. Pay no attention to markings or numbers that may appear on the photos or any other differences in the type of style of photographs.

After being read the admonition, participants were asked if they recognized the carjacker from the video in the lineup, and, if so, to indicate who they thought he was.

Lineup construction. A pool of 20 pictures were selected depicting possible fillers who were judged to be similar to the culprit in terms of age, race, hairstyle, and general appearance. These pictures were shown to a group of 40 participants who judged each picture on how similar it was to the culprit. The top six most similar pictures were used to construct the lineup. Lineup fairness was assessed using a mock lineup method; *Tredoux's E* = 1.40.

Twelve versions of the lineup were created: six for culprit-present (CP) and six for culprit-absent (CA), with each member rotated to a different position of the six-person photographic lineup. For the culprit-absent lineup, the target was replaced with a picture of his brother, who was designated as the nominated suspect. Pilot testing revealed that, of the six pictures, the brother's picture was judged to be most similar to the target, and the filler with the tattoo was judged to be third most similar of the six. The 12 lineups were administered in a counterbalanced manner to ensure random assignment; alternating between culprit presence and absence, and also rotating the positioning of the pictures within each condition. The lineups were administered on a computer faced away from the experimenter so they were blind to the target's position.

Determining plausibility. Ten graduate students and their advisor first made judgments of what they believed were plausible and implausible suggestions that could be used in the study

based on viewing the carjacking video. From these suggestions, three plausible and three implausible suggestions were selected. The process of selecting the items to be used emphasized the importance of the implausible suggestions not being ridiculous or impossible. Rather, the implausible suggestions were designed to be *very unlikely*, but possible. We then obtained likelihood ratings, through pilot testing. Independent ratings were obtained from a sample of 38 participants who viewed the car-jacking video and then rated how likely each misinformation item was on a 1–8 scale; with 1 being not likely at all and 8 being very likely. Here are the ratings for the implausible suggestions: (1) the suggestion that *the culprit was African-American* was judged to be likely or very likely by 18.4% of pilot participants; (2) the suggestion that *the culprit had short hair – about an inch to an inch and a half long* was judged to be likely or very likely by 13.5% of participants; and (3) the suggestion that *the culprit flicked a cigarette* was judged as likely or very likely by 18% of participants. Here are the ratings for the plausible suggestion: (1) the suggestion that *the culprit had a tattoo on his neck* was judged to be likely or very likely by 57.9% of participants; (2) the suggestion that *the perpetrator was over six feet tall, possibly 6'4" or 6'5"* was rated as likely or very likely by 23.7% of participants; and (3) the suggestion that *the culprit had long pants* was judged to be likely by 26.3% participants. Overall, the three plausible misinformation items were judged to be more likely to have been in the video, $M = 3.35$, $SE = .21$, (95% *CI* [2.92 - 3.27]), than the implausible items, $M = 2.67$, $SE = .18$, (95% *CI* [2.31 - 3.01]).

Data from Table 3 shows the frequency that each misinformation item was reported by participants in the small group co-witness sessions (after the suggestions were made by the confederates who spoke first). These data further verify the validity of the distinction between the plausible and implausible item groupings. Close examination of this table shows that all of the plausible items were more likely to be reported than any of the implausible items. When examining how many participants in each small group reported recalling the misinformation items (after the suggestions were made by the confederates who spoke first), the differences between the plausible and implausible items were even larger.

Coding of the participants' memory reports. The features of events method was used to code the participants' free recall memory for the details of the perpetrator and the crime in general (Dickinson & Poole, 2000). A checklist was created that included all of the correct information contained in the confederates' statements and the six pieces of misinformation reported during the small group co-witness session and for their final memory reports given after the varied retention intervals. Pairs of coders read the responses of each participant and then independently recorded which items each participant recalled that had previously been reported by the confederates. Initial rate of agreement across raters was over 98%. Disagreements were resolved through group discussion at a weekly lab meeting.

Results

Identifying the Man with the Tattoo

Rates of choosing the filler with the tattoo on his neck, the culprit (or nominated suspect in the culprit-absent condition), and the next most chosen picture can be found in Table 1. Since more than half of the participants were Latino, ethnicity was dummy coded as Latino vs. others for these analyses to examine potential cross-race effects. Also, data were coded for which confederate spoke first during the group report. In addition, a dummy variable was created for whether the participant stated that they recalled seeing a tattoo during the initial group discussion (initial report of tattoo).

Because of the group nature of the sessions, it was possible that hearing another participant in the group report recalling the tattoo might have an effect on the participants' memory and/or lineup performance. To control for this possibility, data from the initial group discussion were coded to indicate if the participant and/or anyone else in each small group reported recalling the tattoo during the co-witness misinformation sessions.

A hierarchical binary logistic regression was conducted with identification decision as the dependent variable (tattooed filler vs. other), and the following predictors: retention interval (control vs. 5-minutes vs. 50-minutes vs. 2-days vs. 1-week), culprit presence (present vs. absent),

race (Latino vs. other), order that the two confederates spoke (tattoo suggestion before vs. after the participants gave their report), and initial report of the tattoo by the participant during the group co-witness session (yes vs. no). In the first block, all main effects were entered, on the second, two-way interactions were entered, and, on the third, three-way interactions were examined. The predictor model was a significant improvement over the constant-only model, $X^2(5, N = 289) = 13.21$, $p < .02$. Participants were 1.2 times more likely to identify the tattooed filler after one week than the controls ($B = .18$, $SE = .09$, $Wald = 3.95$, $p = .05$, $e^B = 1.20$). When the culprit was not present in the lineup, participants were 60% more likely to identify the filler with the tattoo than when the culprit was present ($B = -.51$, $SE = .25$, $Wald = 4.20$, $p = .04$, $e^B = .60$). No significant effect was found for race ($B = -.43$, $SE = .25$, $Wald = 2.81$, $p=.09$, $e^B = .65$), order of the confederate's report, before-or-after the participants spoke ($B = .15$, $SE = .25$, $Wald = .37$, $p = .55$, $e^B = 1.17$), or whether the participants reported recalling the tattoo during the group co-witness session ($B = .93$, $SE = .95$, $Wald = .95$, $p=.33$, $e^B = 2.53$). See Table 1 for choosing rates across all conditions. When two-way interactions were added, the second block of the analyses showed no significant improvement in the model, $X^2(7, N = 289) = 12.55$, *ns*, and the third block also did not show improvement $X^2(7, N = 289) = 10.57$, *ns*. These analyses were repeated to control for whether *anyone* in each participant's small group co-witness session reported recalling the tattoo. This was accomplished by replacing the predictor variable for whether the participant reported recalling the tattoo during the co-witness session with the variable that indicated whether *the participant and/or anyone else in their small group* reported recalling the tattoo during the co-witness session. This did not result in any changes to the effects described above.

A second hierarchical binary logistic regression was conducted using repeated contrasts to examine incremental differences across the four delay conditions. For this analysis, the no-misinformation control group was removed, and only the four different retention interval groups who heard the misinformation were examined (5-minutes vs. 50-minutes vs. 2-days vs. 1-week). As with the previous analyses, identification decision was the dependent variable (tattooed filler

vs. other), with the following predictors: retention interval (control vs. 5-minutes vs. 50-minutes vs. 2-days vs. 1-week), culprit presence (present vs. absent), race (Latino vs. other), order that the two confederates spoke (tattoo suggestion before vs. after the participants gave their report), and participants report of the tattoo during the group co-witness session (yes vs. no). In the first block, all main effects were entered, on the second, two-way interactions were entered, and, on the third, three-way interactions were examined. The predictor model was a significant improvement over the constant-only model, $X^2(7, N = 289) = 15.22$, $p = .03$. When considering differences across the retention intervals, participants were 1.8 times more likely to identify the filler with the tattoo after one week, compared to two days ($B = -.77$, $SE = .40$, $Wald = 3.64$, $p = .057$, $e^B = 2.15$). However, this predicted directional effect was only statistically significant with a one-tailed analysis. When the culprit was not present in the lineup, participants were 52% more likely to identify the filler with the tattoo than when the culprit was present ($B = -.66$, $SE = .27$, $Wald = 5.70$, $p = .02$, $e^B = 52$). Again, no effect was found for race ($B = -.50$, $SE = .28$, $Wald = 3.08$, $p = .08$, $e^B = .61$), the order or the confederate's presentation ($B = -.18$, $SE = .29$, $Wald = .40$, $p = .53$, $e^B = 1.20$) or whether the participant reported recalling the tattoo during the group co-witness session ($B = .87$, $SE = .97$, $Wald = .79$, $p = .37$, $e^B = 2.38$). When two-way interactions were added, the second block of the analyses showed no significant improvement in the model, $X^2(8, N = 289) = 12.87$, $p = .12$. Again, the predictor variable for whether the participant reported recalling the tattoo during the initial group session was replaced with the more inclusive predictor that indicated whether the participant *and/or anyone else in their small group* reported recalling the tattoo during the co-witness session. Again, this did not result in any changes to the effects described above.

Free Recall of the Misinformation

Scales were created for proportion of errors on the plausible and implausible items in the recall sessions conducted at the various retention intervals. Scales were also created for whether participants reported the plausible and implausible items during the initial small group co-witness sessions (after hearing the misinformation). Also, as done with the regression analyses, in order to

control for the possibility that other people in each participant's small group co-witness session reported recalling any of the misinformation items, additional scales were created assessing whether the participant and/or any other people in the small group reported remembering any of the plausible or implausible misinformation items after hearing the misinformation. The distribution of errors for plausible and implausible items can be seen in Tables 2 and 3.

Initial errors on plausible misinformation items made during the group co-witness sessions were correlated with plausible errors in the final recall sessions conducted after the varied retention intervals (r [289] = .41, $p < .001$), but were not related to errors on implausible items made during the initial group co-witness session (r [289] = -.05, *ns*), or implausible errors in the final recall sessions (r [289] = .10, $p = .09$). Initial plausible *group errors* (if the participant and/or anyone in the small group reported the plausible misinformation items during the initial small group co-witness session) were also correlated with plausible errors in the final recall sessions conducted after the varied retention intervals (r [289] = .35, $p < .001$), but were not related to implausible errors in the final recall session (r [289] = .11, $p = .06$). Similarly, implausible errors made during the initial group co-witness sessions were related to implausible errors made during the final recall sessions (r [289] = .27, $p < .001$), but not related to plausible errors made either initially on the group session, or after a delay.

To examine the effect of plausibility and retention interval on the misinformation recalled, a mixed factorial ANOVA was conducted with plausibility as a within-subjects factor (i.e., errors on plausible and implausible items at the final recall sessions), and retention interval (control, 50-min., 2-days, 1-week) and order of the confederates' presentation of the misinformation as between-subject factors. In order to control for initial acceptance of the misinformation during the small group sessons, participant reports of recalling the plausible and implausible misinformation items during the co-witness sessions were also added as between-subjects factors. As predicted, a main effect for plausibility was revealed, with the plausible misinformation being more likely to be included in the memory reports of the participants, $F(1, 275) = 17.08$, $p < .001$, $\eta_p^2 = .06$. Also,

a main effect for retention interval was revealed, $F(4, 275) = 2.61$, $p = 04$, $\eta_p^2 = .04$. Tukey post-hoc tests indicated that, compared to the control group, more misinformation was reported in the 5-minute delay group, *M Diff.* = .05, *SE* = .02, $p = .04$, 95% *CI* = .0009 – .1080, the 50-minute delay group, *M Diff.* = -.07, *SE* = .02, $p = .005$, 95% *CI* = .0150 – .1250, the two-day delay group, *M Diff.* = .05, *SE* = .02, $p = .05$, 95% *CI* = -.0004 – .1012, and the 1-week delay, group *M Diff.* = .05, *SE* = .02, $p = .05$, 95% *CI* = -.0004 – .1012. However, no significant incremental differences were found between the different retention intervals.

No effect for the order of the confederates' presentation was revealed, $F(1, 275) = .66$, $p = .42$, $\eta_p^2 = .002$. However, if a participant repeated the misinformation reported by the confederate in the initial small group co-witness session, they were more likely recall the misinformation in the final recall session. This was true for both plausible, $F(1, 275) = 45.49$, $p < .001$, $\eta_p^2 = .14$, and implausible information, $F(1, 275) = 8.1$, $p = .01$, $\eta_p^2 = .03$. No interaction between these variables and delay was found.

Supplemental analyses were conducted after replacing the between-subjects variables of initial plausible and initial implausible errors made by the participants during the group co-witness sessions with a more inclusive variable indicating whether the participant and/or anyone else in the small group reported any of the plausible or implausible items during the initial group session. These analyses revealed the same pattern of results.

Discussion

The current study examined the effect of plausibility and retention interval on memory for misinformation delivered within a co-witness paradigm. We were particularly interested in how overhearing plausible post-event misinformation from a co-witness, related to a distinctive feature of the perpetrator's appearance (i.e., a neck tattoo), would affect witnesses' identification decisions when viewing a lineup where one of the fillers has this suggested feature.

As predicted, the tattooed filler was selected most often at the longer retention intervals when the witnesses' memory for the perpetrator was likely weakest. Table 1 shows that, when the

culprit was absent, participants in the 5-minute delay condition selected the tattooed filler at the same rate as the nominated suspect, and were no more likely to select the tattooed filler then participants in the no-misinformation control condition. However, after 50-minutes, selection of the tattooed filler increased substantially from 28.6% to 47.8%, and continued to increase as the retention interval grew longer. Moreover, this shift towards selecting the tattooed filler at the longer retention intervals was accompanied by a shift away from identifying the nominated suspect, indicating that the post-event information suggesting that the perpetrator had a tattoo on his neck had a greater effect on choosing when memory for the perpetrator's face faded. The misinformation effect was most powerful after 1-week, when identifications of the tattooed filler mushroomed to 60%, while selection of the nominated suspect bottomed out at 7%.

When the actual culprit was present, the same general effect for time delay was observed; however, the effect was only evident at the longest retention interval, when selection of the tattooed filler increased to 44% (up from 32% after five-minutes). Up to that point, the actual culprit was chosen at a rate that was comparable to, or greater than the tattooed filler. Unlike the pattern observed in the culprit-absent condition, where selection of the tattooed filler at the longer delays was associated with a decline in choosing the nominated suspect, identification of the actual culprit remained stable over time, and the shift toward selecting the tattooed filler after one-week was associated with a decrease in choosing the other fillers. Theoretically, those individuals who did not recall his face well had a weak ecphoric experience when viewing the lineup, were more likely to shift to secondary processes and rely on other information independent of recognition memory to make their identification decision. In this case, it appears that they were influenced by their memory for the misinformation that the perpetrator had a tattoo on his neck.

It is worth noting that the tattooed filler was a very good match to the perpetrator (Pilot testing showed that the tattooed filler was selected third most often of the pictures used - tied for third). Table 1 shows that in both the control, and five-minute delay conditions, the tattooed filler was chosen at the same rate as the next most selected photo when the culprit was absent, and

slightly more often than the culprit when he was present. However, despite these initially comparable rates of choosing, results revealed that the predicted increase in false identifications of the tattooed filler over time was clearly substantial, particularly in the culprit-absent conditions, where choosing of the tattooed filler more than doubled from 28.6% after five minutes to 60% after one week.

The real-world implications of these findings are clear; when an outside source suggests to a witness that a perpetrator has a unique feature, this can affect the co-witness's memory for the perpetrator, and, in turn, impact that witness's ability to make accurate identification decisions when viewing subsequent lineups. Although, in many cases, the authorities will make sure that unique features like tattoos are consistent across fillers, this is not always the case.

Memory for the Misinformation

Participants were more likely to misremember plausible misinformation suggested by the confederate co-witness at all retention intervals. Previous research has shown that post-event information is more likely to be accepted when it is plausible, and when it does not contradict, or attempt to transform, a person's memory for the details of the event (Gabbert, Memon, & Wright, 2006). In this study, five of the six misinformation items did, in fact, contradict the elements of the person's appearance and/or actions, and only the suggestion that he had a tattoo was purely additive. Not surprisingly, the tattoo suggestion was most likely to be misremembered at all retention intervals.

Participants' recall for the suggested misinformation was greatest after 50-minutes, and was lowest overall after one week. This effect was clearest for the plausible misinformation. Schwartz and Wright (2012) found that witness conformity effects are strongest when the post-event information is presented closer to the test, simply because the misinformation is most likely to be recalled after a shorter retention interval. Although retention interval effects were not particularly strong in this study, these findings generally fit that pattern.

Recognition Versus Recall: The Differential Effect of the Retention Interval

Although free recall for the post-event misinformation provided by the co-witnesses dropped off after the 50-minute retention interval, the effect of the suggested information on lineup performance increased after this delay, and was highest after one week. At the longer retention intervals, participants were far more likely to choose the filler with the tattoo than to report remembering that the perpetrator had a tattoo. For example, after one-week, only 13% of participants reported recalling the tattoo, whereas 60% of participants in the culprit-absent condition chose the tattooed filler. This shows that most of the participants who selected the tattooed filler after one week did not show clear independent evidence of consciously recalling the tattoo. A similar pattern was revealed in the culprit-present condition. As noted earlier, the increased effect of the tattoo suggestion on lineup performance at the longer retention intervals appeared to be associated with an inevitable decline in participants' memory for the perpetrator's face.

Summary

The current findings show that both recall and recognition are vulnerable to suggestions from others. When this occurs in the context of a forensic investigation, consistent statements obtained from witnesses might be seized upon as valuable corroborative evidence from independent witnesses, when, in fact, the evidence might be contaminated if one mistaken witnesses has discussed their erroneous memories with others. This is especially true when ecphoric experience cannot be relied upon, and individuals rely on secondary deliberative processes to inform their identification decision. In actual cases, lineups are frequently conducted after much longer delays than used in this study. These data indicate that, if witnesses encounter erroneous information about a suspect's appearance after the event, but before the identification procedure, then their lineup decisions can be affected by this misinformation, even when the actual perpetrator is present in the display, and when they show no evidence of consciously recalling the misinformation.

References

Allen, R., & Gabbert, F. (2013). Exogenous social identity cues differentially affect the dynamic tracking of individual target faces. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 39(6), 1982-1989.

Alogna, V. K., Attaya, M. K., Aucoin, P., Bahnik, S., Birch, S., Birt, A. R., Zwaan, R. A. (2014). Registered replication report: Schooler & Engstler-Schooler (1990). *Perspectives on Psychological Science, 9,* 556–578.

Berntsen, D., & Rubin, D. C. (2007). When a trauma becomes a key to identity: Enhanced integration of trauma memories predicts posttraumatic stress disorder symptoms. *Applied Cognitive Psychology, 21*(4), 417-431. doi:10.1002/acp.1290

Deffenbacher, K. A., Bornstein, B. H., McGorty, E. K., & Penrod, S. D. (2008). Forgetting the once-seen face: Estimating the strength of an eyewitness's memory representation. *Journal of Experimental Psychology: Applied, 14*(2), 139-150. doi:10.1037/1076-898X.14.2.139

Dickinson, J. J., & Poole, D. A. (2000). Efficient coding of eyewitness narratives: A comparison of syntactic unit and word count procedures. *Behavior Research Methods, Instruments, & Computers, 32*(4), 537-545. doi:10.3758/BF03200826

Gabbert, F. & Brown, C. (2015). Interviewing for face identification. In T. Valentine, & J. P. Davis (Eds.), *Forensic Facial Identification: Theory and Practice of Identification from Eyewitnesses, Composites and CCTV* (pp. 17-42). Chichester, England: Wiley-Blackwell.

Gabbert, F., & Hope, L. (2013). Suggestibility and memory conformity. In A. M. Ridley, F. Gabbert, & D. J. La Rooy (Eds.), *Suggestibility in Legal Contexts: Psychological Research and Forensic Implications.* (pp. 63-84). London, England: Wiley-Blackwell.

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps towards influence during a discussion. *Psychonomic Bulletin & Review, 13*, 480-485. doi:10.3758/BF03193873

Horry, R., Colton, L. M., & Williamson, P. (2014). Confidence-accuracy resolution in the

misinformation paradigm is influenced by the availability of source cues. *Acta Psychologica, 151,* 164-173. doi:10.1016/j.actpsy.2014.06.006

Hyman, I. J., & Loftus, E. F. (2002). False childhood memories and eyewitness memory errors. In M. L. Eisen, J. A. Quas, & G. S. Goodman (Eds.), *Memory and suggestibility in the forensic interview* (pp. 63-84). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361-366. doi:10.1101/lm.94705

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(1), 19-31. doi:10.1037/0278-7393.4.1.19

Meade, M. L., & Roediger, H. L. (2002). Explorations in the social contagion of memory. *Memory & Cognition, 30*(7), 995-1009. doi:10.3758/BF03194318

Meissner, C., Brigham, J., & Kelley, C. (2001). The influence of retrieval processes in verbal overshadowing. *Memory & Cognition, 29*(1), 176-186. doi:10.3758/BF03195751

Meissner, C. A., & Memon, A. (2002). Verbal overshadowing: A special issue exploring theoretical and applied issues. *Applied Cognitive Psychology, 16*(8), 869-872. doi:10.1002/acp.928

Meissner, C. A., Sporer, S. L., & Susa, K. J. (2008). A theoretical review and meta-analysis of the description-identification relationship in memory for faces. *European Journal of Cognitive Psychology, 20*(3), 414-455.

Paterson, H. M., & Kemp, R. I. (2006). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime & Law, 12*(2), 181-191. doi:10.1080/10683160512331316334

People v. Richard Torres, San Diego Superior Court (February, 2015, CD256364)

Pezdek, K., & Blandon-Gitlin, I. (2009). Planting false memories for childhood sexual abuse only happens to emotionally disturbed people… not me or my friends. *Applied Cognitive Psychology, 23*(2), 162-169. doi:10.1002/acp.1466

Schooler, J. W. (2002). Verbalization produces a transfer inappropriate processing shift. *Applied Cognitive Psychology, 16*(8), 989-997. doi:10.1002/acp.930

Schwartz, S. L., & Wright, D. B. (2012). Memory conformity for new and old items with immediate and delayed testing. *Applied Cognitive Psychology, 26*(4), 508-515. doi:10.1002/acp.2820

Scoboria, A., Mazzoni, G., Jarry, J., & Shapero*, D. (2012). Implausibility inhibits but does not eliminate false autobiographical beliefs. *Canadian Journal of Experimental Psychology.* doi.org/10.1037/a0030017.

Shapiro, P. N., & Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*(2), 139-156. doi:10.1037/0033-2909.100.2.139

Skagerberg, E. M., & Wright, D. B. (2008). The prevalence of co-witnesses and co-witness discussions in real life. *Psychology Crime and Law, 14*(6), 513-521. doi:10.1080/10683160801948980

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition, 14*(4), 329-338. doi:10.3758/BF03202511

Underwood, J., & Pezdek, K. (1998). Memory suggestibility as an example of the sleeper effect. *Psychonomic Bulletin & Review, 5*(3), 449-453. doi:10.3758/BF03208820

van Koppen, P. J., & Lochun, S. K. (1997). Portraying perpetrators: The validity of offender descriptions by witnesses. *Law and Human Behavior, 21*(6), 661-685. doi:10.1023/A:1024812831576

Walther, E., Bless, H., Strack, F., Rackstraw, P., Wagner, D., & Werth, L. (2002). Conformity effects in memory as a function of group size, dissenters and uncertainty. *Applied Cognitive Psychology, 16*(7), 793-810. doi:10.1002/acp.828

Wells, G. L. (1985). Verbal descriptions of faces from memory: Are they diagnostic of

identification accuracy? *Journal of Applied Psychology, 70*(4), 619-626. doi:10.1037/0021-9010.70.4.619

Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. *Current Directions in Psychological Science, 18*(3), 174-178. doi:10.1111/j.1467-8721.2009.01631.x

Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology, 91*(2), 189-202. doi:10.1348/000712600161781

Zajac, R., & Henderson, N. (2009). Don't it make my brown eyes blue: co-witness misinformation about a target's appearance can impair target-absent line-up performance. *Memory, 17*(3), 266-278. doi:10.1080/09658210802623950

Table 1. Percentage of participants who selected the tattooed filler, the nominated suspect/actual culprit, and the next most selected photo for culprit absent and culprit present conditions across all retention intervals.

	Culprit Absent			Culprit Present		
	Nominated Suspect	Tattooed Filler	Next Most Chosen Pic	Actual Culprit	Tattooed Filler	Next Most Chosen Pic
Control	20.8%	29.2%	29.2%	26.9%	30.8%	11.5%
5-minutes	28.6%	28.6%	28.6%	25.0%	32.1%	10.7%
50-minutes	8.7%	47.8%	21.7%	26.9%	30.8%	15.4%
2-days	12.1%	48.5%	21.2%	34.4%	22.9%	14.3%
1-week	6.7%	60.0%	16.7%	34.4%	43.8%	6.3%

Table 2. Percentage of participants who reported recalling each misinformation item at the various retention intervals.

	Plausible Misinformation Items			Implausible Misinformation Items		
	Tattoo on Neck	Over 6' Tall	Long Pants	Short Hair	African-American	Cigarette Flicked
Misinformation Reported During the Small Group Co-witness Session						
Participant*	3.0%	3.6%	4.3%	2.4%	2.4%	0.8%
Group**	7.2%	8.4%	12.0%	3.0%	4.1%	2.4%
Misinformation Reported During the Final Recall Session						
Control	0%	2.0%	3.7%	1.0%	0%	0%
5-min.	16.7%	11.1%	4.8%	9.3%	1.9%	0%
50-min.	24.0%	10.0%	6.0%	4.0%	8.0%	0%
2-days	19.1%	11.8%	1.5%	8.8%	2.9%	0%
1-week	14.5%	8.1%	4.8%	9.7%	3.3%	0%

* Initial reports given by the participant during the small group co-witness session, when the confederate presenting that misinformation spoke first.
**Initial reports by the participant and/or anyone else in each small group co-witness session when the confederate presenting that misinformation spoke first.

Table 3. Information reported across retention intervals

	Effect Size		Misinformation Recalled					
	Plausible vs. Implausible		Plausible MI			Implausible MI		
	N	Cohen's *d*	M	SD	95% CI	M	SD	95% CI
Control	56	.17	.08	(.27)	.01 - .16	.04	(.20)	-.02 - .10
5-minutes	50	.42	.32	(.54)	.18 - .47	.13	(.33)	.04 - .21
50-minutes	71	.55	.42	(.67)	.23 - .61	.12	(.39)	.01 - .23
2-days	62	.43	.31	(.58)	.17 - .45	.11	(.32)	.04 - .19
1-Week	50	.32	.29	(.46)	.17 - .41	.15	(.40)	.04 - .25