Can fabricated evidence induce false eyewitness testimony?

Kimberley A. Wade, Sarah L. Green, and Robert A. Nash
University of Warwick, Coventry, United Kingdom

Word count: 3620 words (main text)

Corresponding author:
Kimberley A. Wade
Department of Psychology
University of Warwick
Coventry, CV4 7AL
United Kingdom
Direct: +44 24 7657 5680  Fax: +44 24 76572 4225
Email: K.A.Wade@warwick.ac.uk
Abstract
False information can influence people’s beliefs and memories. But can fabricated evidence induce individuals to accuse another person of doing something they never did? We examined whether exposure to a fabricated video could produce false eyewitness testimony. Subjects completed a gambling task alongside a confederate subject, and later we falsely told subjects that their partner had cheated on the task. Some subjects viewed a digitally-manipulated video of their partner cheating; some were told that video evidence of the cheating exists; and others were not told anything about video evidence. Subjects were asked to sign a statement confirming that they witnessed the incident and that their corroboration could be used in disciplinary action against the accused. See-video subjects were three times more likely to sign the statement than Told-video and Control subjects. Fabricated evidence may, indeed, produce false eyewitness testimony; we discuss probable cognitive mechanisms.
Can fabricated evidence induce false eyewitness testimony?

“The time was, we thought of photographs as recorders of reality. Now we know they largely invent reality”, writes Holland Cotter of *The New York Times* (Cotter, 2008). He is right of course: digital-trickery is part-and-parcel of everyday life. Yet fabricated images or video-footage can be extraordinarily compelling and difficult to detect, and forensic experts are increasingly called upon in criminal and civil cases to determine whether digital evidence has been altered (Peterson, 2006). In this article we examine whether doctored video-footage can induce people to testify about an event they never witnessed. Unlike previous eyewitness studies, we set out to examine whether people will falsely accuse another person of committing a misdemeanor when there are ostensibly real consequences to making that accusation. To this end we developed a new experimental procedure—a laboratory analog for exposing witnesses to fabricated evidence and obtaining false eyewitness testimony. Before describing this novel procedure, we review the scientific literature that shows verbal suggestions, and fabricated evidence in the form of digitally-manipulated images, can alter people’s autobiographies.

*Misinformation, false beliefs and false memories*

Researchers have amply demonstrated the ease with which false information can lead people to report witnessing events that never happened, or to give inaccurate reports about events that did happen (Loftus, 2005). Misinformation can come in different forms, such as in a leading question or in a co-witness’s testimony (Gabbert, Memon, & Allan, 2003; Loftus & Palmer, 1974; see Wright, Memon, Skagerberg, & Gabbert, 2009 for a recent review). But in recent years, memory researchers have used an especially compelling form of misinformation—digitally-manipulated photos and
videos—to elicit false beliefs and memories in people (Garry & Wade, 2005; Nash, Wade, & Lindsay, 2009; Strange, Sutherland, & Garry, 2006; Wade, Garry, Nash, & Harper, 2009; Wade, Garry, Read, & Lindsay, 2002; see Garry & Gerrie, 2005 for a review). Sacchi, Agnoli and Loftus (2007), for instance, exposed adults to photographs of significant public events; including a major peace demonstration that took place in Rome. Some subjects viewed doctored photos of the demonstration that suggested the event was more violent than it truly was. Subjects exposed to the doctored photo remembered the demonstration as being more violent and reported less inclination to participate in future demonstrations than did subjects exposed to the real photo. These findings, and the false memory literature more generally, show that false information or fabricated evidence can be powerfully suggestive; it can alter individuals’ beliefs and cultivate rich false memories about both public and personal events.

Based on false memory studies, scientists have argued that false suggestions could induce people to testify about events they never witnessed (Loftus, 2003). Yet there are important differences between false memory studies and real-life crimes. Specifically, in false memory studies subjects usually report erroneous information about something they have witnessed, or something they have done, in response to a misleading suggestion. Importantly, there are no consequences for reporting inaccurate information. However, in the real world, when eyewitnesses report erroneous information they are providing testimony about another person’s actions, and they know that their testimony will be used against the accused in a criminal trial. What we do not know, then, is whether false information might lead people to provide testimony when doing so has real ramifications.

In the current research, we ask whether doctored video-footage can induce people to testify about a misdemeanor they never witnessed. Doctored photos and videos are a
powerful form of suggestion which makes them ideal for studying eyewitness phenomena. If people believe they have witnessed an event, and are motivated to recall that event, doctored video-footage could constitute a source of compelling and perceptually-detailed images that may encourage them to develop false beliefs (Johnson, Hashtroudi, & Lindsay, 1993; Lindsay, 2008). In turn, if these distortions are sufficiently convincing and realistic, subjects may bear witness to a fictitious event. We tested this hypothesis using the false video procedure.

The false-video procedure

Our procedure is modelled on that used by Nash and Wade (2009), whose subjects completed a gambling task and were later accused of having cheated on the task. In that study, subjects who saw a doctored video of themselves ostensibly cheating—withdraw ing money from the bank when they should have deposited money into the bank—were more likely to believe they were guilty than were subjects who were merely told that video-evidence existed. In the present study, subjects also completed the gambling task, but they did it alongside a confederate subject. Later, instead of accusing subjects of cheating, we falsely told subjects that their partner had cheated on the task. Some subjects viewed fake video evidence of their partner cheating alongside them (See-video group); others were merely told that such video evidence exists (Told-video group). A third group were not told anything about video evidence (Control group). We asked all subjects whether they could corroborate the (false) accusation by signing a statement confirming that they had witnessed their partner cheating.

Based on Nash and Wade’s (2009) findings, and on a growing number of studies that suggest images are an especially persuasive form of evidence (Kassin & Dunn, 1997; McCabe & Castel, 2008), we predicted that See-video subjects would be more vulnerable
to belief distortions—and thus more likely to sign the witness statement—than Told-video and Control subjects.

Method

Subjects

Sixty university students (53% female, range 18-43 years, $M = 21.6$, $SD = 4.1$) participated individually and received £6. They were randomly assigned to the See-video, Told-video, or Control condition such that there were 20 subjects in each group. Two confederates—blind to the hypotheses and group randomization—assisted in the study, and were extensively trained to follow an interview protocol and behave consistently across subjects.

Procedure

*Phase 1.* Subjects were seated alongside Confederate A, who was posing as another subject. Confederate A was a student at Warwick University, and she confidently confirmed that she did not know any of the subjects who participated in the study. Subjects were told that the experimenters were investigating differences in gambling behaviour when individuals gamble with physical versus virtual ‘money’. Their task was to earn as much fake money as possible on a computerized gambling task. Subjects were falsely informed that the person who made the largest profit in the experiment would win a cash prize.

The confederate and the subject each received a pile of fake money to gamble with and they shared a pile of money that represented the bank. A video-camera was assembled at the back of the room (Figure 1 depicts the camera’s field of view). Together they were filmed independently completing the gambling task. The task consisted of 15 multiple-choice general knowledge questions, and the subject and confederate answered the same set of questions. Each question had four possible responses and each response
was associated with a different odds ratio. For example, the question “Of what is ‘Rhytiphobia’ the fear?” had the following possible answers and odds ratios: [a] Getting wrinkles (Odds = 2:1), [b] Getting dirty (Odds = 3:1), [c] Getting undressed (Odds = 5:1) and [d] Getting leprosy (Odds = 10:1). Subjects selected an answer and typed the amount of money they wished to gamble on each question (See Nash & Wade, 2009 for a screen-shot of a question). Subjects received feedback after each question. In response to correct answers, a green tick appeared on the monitor with instructions to collect money from the bank (situated between the subject and confederate); in response to incorrect answers, a red cross appeared on the monitor with instructions to return money to the bank. In pilot tests, subjects performed at 33% accuracy on average. Confederate A always made an identical series of bets: she answered five questions correctly (consistent with pilot subjects’ 33% accuracy) and only ever took money and returned money to the bank as instructed. Phase 1 lasted approximately 15 minutes, and both the experimenter and Confederate A were blind to the subject’s condition in this phase.

To create the doctored video to show to See-video subjects, immediately after Phase 1 we manipulated the video-recording using Final Cut Pro 5®. We selected a 10- to 20-second segment of the video that showed Confederate A answering a question correctly, and taking money from the bank as appropriate. We then digitally replaced the green tick on her monitor with a red cross (Figure 1). The resulting clip ostensibly showed the subject and Confederate from behind, seated at the computers, and the confederate collecting money from the bank when she clearly should have returned money.

Phase 2. Subjects returned to the laboratory 5-7 hours later, expecting to complete further gambling-related tasks. However, they were told their absent partner (Confederate A) was suspected of cheating by taking money from the bank when she should have returned money. Specifically, the experimenter informed subjects that the
profit their partner recorded far exceeded what the computer records stated, and that she had been known to cheat in previous experiments. The experimenter explained that the Department was keen to take disciplinary action against the cheat to prevent her from ruining other experiments, yet the video-recording was insufficient to prove the accusation because the cheat repeatedly blocked the camera’s view. It was not possible for the experimenter to be blind to the subject’s condition in Phase 2 because the experimenter delivered the misleading suggestion. To ensure minimal variation across subjects, the experimenter was carefully trained to follow a script and behave identically in each session.

The experimental manipulation occurred as follows: See-video and Told-video subjects were informed that the video-recording from Phase 1 had captured the confederate improperly taking money from the bank on one occasion, but that this occasion did not account for the large money discrepancy. See-video subjects also viewed an approximately 15-s doctored video-clip that ostensibly showed the confederate taking money after answering a question incorrectly. Control subjects were not told about or shown any incriminating footage.

After receiving the video or verbal ‘evidence’ (or no evidence in the case of the control group), subjects were asked whether they had witnessed any cheating, and asked whether they would sign a statement so that their corroboration of the accusation could be used as evidence. The experimenter handwrote the following statement on a pre-constructed ‘Disciplinary Incident Report Form’:

Student suspected to have knowingly cheated in an experiment with an incentive prize fund. On at least one occasion the participant was seen taking ‘money’ from the bank in this experiment inappropriately. We have reason to believe that this was a deliberate act.
Beneath this statement was a printed section that asked for the subject’s signature to confirm that [a] they witnessed the other subject improperly taking money from the bank, and [b] they understood that their corroboration would result in disciplinary action against the accused student. As well as asking subjects to read through the form carefully, the experimenter read this section aloud to ensure that subjects understood what they were being asked to confirm. The experimenter emphasized to subjects that they should not sign the witness statement unless they actually saw their partner cheat. Specifically, the experimenter said to subjects, “Please do not sign this form just because I am asking you to sign it. Only sign the form if you actually remember seeing them cheat.” See-video subjects were also told that they should not sign merely to confirm that they witnessed the incriminating video-footage: “By signing you are not saying that you have seen them cheating on the videoclip I just showed you. You are saying that you actually saw the cheating yourself.” If subjects did not sign the form after the first request, the experimenter made one additional request for subjects to sign. A final section on the form asked subjects to write down any details of the cheating incident that they could remember.

We used two techniques to assess whether subjects believed the cover story. First, subjects waited in a waiting room while the experimenter consulted with her supervisor. Here, Confederate B—who was posing as a subject waiting for another experiment—initiated a covertly audio-recorded conversation with the subject, encouraging them to describe the incident. As well as probing for suspicion of the cover story, this procedure served to assess subjects’ belief that they witnessed the cheating, as per Kassin and Kiechel’s (1996) false confession (“computer-crash”) paradigm. Confederate B was usually blind to the subject’s condition, but when blind was broken (e.g., the subject mentioned that she watched a video) we checked transcripts of the
conversation to ensure that Confederate B never deviated from the script. Second, subjects returned from the waiting room, supposedly to continue the experiment, and were asked to write down what they thought the study was genuinely investigating.

Finally, subjects were debriefed. The experimenter explained the true nature of the study and the necessity for deception. Subjects were given the opportunity to ask questions and were paid for participating. All subjects reacted positively to debriefing, with many—especially those who had been willing to sign the witness statement—expressing great interest in the outcomes and applications of the research.

Results and Discussion

Overall, 20% of subjects signed the witness statement; 17% on the first request (7 See-video subjects; 2 Told-video; 1 Control), and 3% on the second request (2 See-video subjects only). Most importantly though, our experimental manipulation influenced the likelihood of subjects corroborating the false accusation (after the first request, Fisher’s exact $p = .06$, Cramer’s $V = .35$; after the second request, Fisher’s exact $p = .006$, Cramer’s $V = .45$; see Figure 2). Comparing individual conditions, after the first request, See-video subjects were more likely to sign the witness statement than were Control subjects (Fisher’s exact $p = .04$, $\phi = .38$), but the difference between See-video and Told-video subjects did not reach conventional significance (Fisher’s exact $p = .13$, $\phi = .30$). Due to the small numbers of Control and Told-video subjects signing the witness statement, it was not possible to conduct meaningful statistical tests on the difference between these two conditions. After the second request to sign, See-video subjects were more likely to sign the witness statement than both Told-video subjects (Fisher’s exact $p = .03$, $\phi = .46$) and Control subjects (Fisher’s exact $p = .008$, $\phi = .39$).

Furthermore, 4 of the 12 subjects who signed the witness statement (3 See-video; 1 Told-video) wrote down additional incriminating details or described incriminating
details to Confederate B [e.g., “I saw the ‘x’ sign crossed out on her screen and she reached out for a note from the bank” “I sensed that she took money from the bank very often…but the symbol ‘x’ was always (seemed to me) on the screen”].

Additional checks suggest that subjects were unaware of the true nature of the experiment and genuinely believed that it was investigating gambling behavior. Only one subject indicated suspicion about the experiment in her conversation with Confederate B (“this isn’t part of the experiment, is it?”); this subject was removed from analyses and replaced. In their notes about the true purpose of the study, all subjects wrote a hypothesis involving gambling behavior. Additionally, during debriefing See-video subjects often expressed surprise that the video was doctored and that the study was investigating false accusations (e.g., “Really? You faked the video?” “Wow, I had no idea”).

From a theoretical perspective, although our data do not speak directly to the psychological mechanisms responsible for the false-video effect, we suspect that seeing the fake footage might have provided See-video subjects with several different bases on which to judge that they witnessed the cheating. First, our doctored footage may have helped subjects to imagine their partner cheating, and research shows that seeing relevant images encourages people to make errors about what happened in the recent past (Henkel & Carbuto, 2008). Second, the footage may have provided a ‘visual fluency’ that made the cheating event feel familiar. People often misattribute feelings of fluency for familiarity when thinking about an event. Familiarity, in turn, is mistaken for a signal that the event must have happened in the past (Bernstein, Whittlesea, & Loftus, 2002; Jacoby, Kelley, & Dywan, 1989). Finally, the video evidence might have offered a high level of credibility to the claim that subjects may have witnessed cheating. Models of false belief such as that of Mazzoni and Kirsch (2002) propose that when
events appear extremely likely to have occurred, people might lower their criteria for believing that the event occurred and for accepting particular mental images as real memories. In other words, the persuasiveness and apparent conclusiveness of the video evidence might have made subjects less cautious about accepting and reporting that they witnessed our confederate cheat (for more on the mechanisms underlying the power of fabricated evidence see Nash, Wade, & Brewer, in press). This finding squares with recent research showing that people evaluate the credibility of a suggestion before accepting it as fact (Echterhoff, Hirst, & Hussy, 2005; Schweitzer & Saks, 2009).

One competing explanation for our results is that subjects who signed the witness statement merely did so to comply with the experimenter’s request, and not because they genuinely believed that they witnessed their partner cheating. Indeed, we might expect compliance to play a role in our study because in false confession research—where people are induced to confess to acts they never committed—subjects will sign a confession statement despite maintaining belief in their innocence (Kassin, 2008; Kassin & Kiechel, 1996). We are confident, though, that compliance alone cannot account for our findings. Recall that subjects were told not to sign the statement unless they believed they saw their partner cheat, and some of the subjects who signed the witness statement also described—either in writing or in conversation with Confederate B—increditing accounts of what they allegedly witnessed. In addition, if compliance alone were responsible for our results, then we would expect to see similar rates of false accusations in the control and experimental conditions. However, only 1 Control subject corroborated the accusation, and See-video subjects were significantly more likely to sign the statement than Control subjects. Together these findings lead us to conclude that compliance alone cannot adequately explain our data.
To assess whether subjects genuinely believed they witnessed their partner cheating, we adopted the confederate-in-the-waiting-room procedure (e.g., Kassin & Kiechel, 1996; Nash & Wade, 2009). Interestingly, though, subjects were somewhat reluctant to talk to our confederate about the incident, and an independent rater who examined transcripts of the conversations between subjects and Confederate B judged that 52\% of subjects did not discuss the accusation even after repeated encouragement from Confederate B. Why should subjects resist talking about the cheating? An examination of subjects’ comments revealed at least two possible reasons. First, subjects may have been concerned about providing the confederate with confidential information (“I don’t know if I should say” “I’m not sure if really you’re meant to be hearing about it”). Second, it is possible that subjects did not want to be perceived as “snitches.” Indeed, our independent rater judged that subjects who signed the witness statement were less likely to discuss the accusation than those who did not sign, $\chi^2(1, N = 60) = 4.54, p = .03, \phi = .32$. In fact, both of the two subjects who signed the statement and discussed the accusation actually denied to the confederate that they signed (“I did remember seeing the cross or tick, but I didn’t tell [the experimenter]”). In short, although the confederate-in-the-waiting-room procedure has proved helpful in assessing subjects’ beliefs and minimizing experimenter demand in some false confession studies, this technique needs to be modified for future false accusation research. One solution might be to replace Confederate B with a friend of the subject or another person who the subject is likely to trust and confide in.

To conclude, decades of research on suggestion reveals that people can hold distorted beliefs about their past experiences (Loftus, 2005; Wright et al., 2009). The primary goal of this research was to examine whether suggestion—in the form of a doctored video or a claim that video evidence exists—could induce people to testify
about a misdemeanor they never witnessed, when doing so would have obvious consequences for the accused. To this end, we have presented a new technique for examining the effects of fabricated evidence upon eyewitnesses. Our results show that people will sign a witness statement and assent to providing (false) eyewitness testimony in response to a compelling yet false 10-20-second video-clip. Moreover, when subjects signed the witness statement they were well aware that their actions would be used to punish a peer. At the time of writing this paper, several studies have shown that people will provide erroneous, incriminating information in response to misleading suggestions (Gabbert et al., 2003; Loftus & Palmer, 1974). Yet to the best of our knowledge, our study is the first to show that people will make false accusations against another real person (as opposed to, for example, an actor seen in a simulated-crime video) when they believe the accused will be punished as a result. These findings have implications for law enforcement officials, policy makers, and virtually any person involved in criminal or civil hearings in which physical evidence may have been presented. In fact, in a recent criminal case, video-footage was presented in a New York court of a suspected rioter appearing to resist arrest. However, the charges against the suspect were dropped after it was discovered that two clips that seemingly proved his innocence had been deleted from the footage (Democracy Now, 2005).

Our research suggests that fabricated evidence need not enter the courtroom to interfere with justice. Rather, showing potential witnesses fabricated evidence—or perhaps even genuine evidence that is somehow misleading—might induce them to testify about entire experiences they have never actually had.
References


Author Note

We thank Hannah Bailey, Giles Poulter, Jenna McKeown and Olwen Bryer for their assistance in data collection and coding, and Dan Wright and an anonymous reviewer for their helpful comments on an earlier draft.

Address correspondence to Kimberley Wade, Psychology Department, University of Warwick, Coventry, CV4 7AL, United Kingdom, Direct: +44 (0)24 7657 5680; Fax: +44(0)24 4652 4225; Email: K.A.Wade@warwick.ac.uk.
Figure captions

Figure 1. The video-doctoring process. The real subject is on the right; Confederate A (the ‘cheat’) is on the left.

Figure 2. Percentage of subjects who signed a witness statement as a function of condition.
Figure 2

The bar chart illustrates the percentage of subjects signing witness statements under different conditions: Control, Told-video, and See-video. The chart compares signing on first request (solid black) versus signing on second request (gray). The See-video condition shows a significantly higher percentage of subjects signing on the first request compared to the other conditions.