Imagination or Proof: The Use of Imagery and False Evidence in Eliciting Internalized False Confessions

Kendall McMillen
CUNY John Jay College, im.swimmer@yahoo.com

Follow this and additional works at: https://academicworks.cuny.edu/jj_etds
Part of the Criminology and Criminal Justice Commons, and the Social Psychology Commons

Recommended Citation

This Thesis is brought to you for free and open access by the John Jay College of Criminal Justice at CUNY Academic Works. It has been accepted for inclusion in Student Theses by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
Imagination or Proof: The Use of Imagery and False Evidence in Eliciting Internalized False Confessions

A Thesis Presented in Partial Fulfillment of the Requirements
for the Masters in Forensic Psychology
John Jay College of Criminal Justice
City University of New York

Kendall McMillen
May 2018
# Table of Contents

Abstract .................................................................................................................................................. 4

Introduction and Literature Review ........................................................................................................ 5

Internalized False Confessions .................................................................................................................. 6

False Evidence and Internalized False Confessions .................................................................................. 7

Eyewitness evidence .................................................................................................................................. 8

Doctored video ......................................................................................................................................... 9

Shown versus told evidence .................................................................................................................... 11

Imagery and False Memories ................................................................................................................... 11

Internalization and confabulation ............................................................................................................. 12

Imagery and internalized false confessions .............................................................................................. 13

False evidence vs. imagination ................................................................................................................ 14

Present Study .......................................................................................................................................... 15

Method .................................................................................................................................................... 16

Design ..................................................................................................................................................... 16

Participants .............................................................................................................................................. 16

Procedure and Materials ....................................................................................................................... 17

Gambling quiz ......................................................................................................................................... 17

Doctored video ........................................................................................................................................ 19

Confession form ..................................................................................................................................... 19
Abstract

Rates of internalized false confessions has become increasingly salient through the use of DNA evidence in exonerating those who have been wrongfully convicted. False confessions also carry heavy legal consequences, and internalized false confessions (Appleby & Kassin, 2016; Kassin & Neumann, 1997; Kassin & Wrightsman, 1980), where the individual actually believes they committed a crime, may be especially damning (Kassin S., 2007). The presentation of false evidence has been found to be extremely influential in eliciting a false confession (Horselenberg, Merckelbach, & Josephs, 2003; Kassin & Kiechel, 1996; Nash & Wade, 2009), as well as imagery (Garry, Manning, Loftus, & Sherman, 1996; Garry & Wade, 2005; Nash, Wade, & Lindsay, 2009; Shaw & Porter, 2015), but they have not been directly compared in previous literature. This was an exploratory study replicating the gambling paradigm created by Nash and Wade (2008), aimed at examining and directly comparing the effects of different forms of false evidence (i.e., shown-doctored video footage, told-doctored video footage, shown-eyewitness, told-eyewitness) and imagery, and their rates in eliciting false confessions. Participants were accused of stealing money from the study and cheating during a gambling task, and rates of confession, internalization, and confabulation were analyzed through Fisher’s exact p – values.

Keywords: false confessions, internalized false confessions, false evidence, confabulation, internalization, imagery
Imagination or Proof: The Use of Imagery and False Evidence in Eliciting Internalized False Confessions

The prevalence of wrongful convictions in the United States has become increasingly salient with the incorporation of DNA evidence. The Innocence Project has reported 347 exonerations through the examination of DNA (The Innocence Project, 2016) since their creation in 1992. Of those cases, they report that 25% of exonerees gave a false confession (The Innocence Project, 2016). False confessions carry heavy legal consequences for the defendant as juries do not adequately discount confessions even when they believe they were coerced (Kassin & Wrightsman, 1980; Kassin & McNall, 1991), and confessions have been found to influence jury decisions more than other forms of evidence, including DNA evidence (Kassin & Neumann, 1997; Appleby & Kassin, 2016).

Even though false confessions are not uncommon, many people may believe that they would never confess to a crime they did not commit. A number of factors have been identified that may increase the risk of a suspect providing a false confession, including the presentation of false evidence and imagery. While there is extant research examining the influence of the former on coaxing confessions (Horselenberg, Merckelbach, & Josephs, 2003; Nash & Wade, 2009; Kassin & Kiechel, 1996), researchers have only recently started to examine the latter (Shaw & Porter, 2015). Critically, no research has directly compared the two techniques in their ability to elicit false confessions and, in particular, internalized false confessions.

Thus, the aim of the present study is to address this gap in the literature and will directly compare the potency of previously studied forms of false evidence (i.e., doctored-video evidence, eyewitness testimony), as well as visualization on inducing an internalized false confession during interrogation. As false confessions carry lasting consequences such as
incarceration, understanding what causes a person to confess and subsequently internalize a crime they did not commit is necessary to avoid future wrongful convictions and provide justice to all involved. In what follows, I will examine a) prior literature in regards to the potency of false evidence in eliciting false confessions, in particular, internalized false confessions, b) imagery and its role in creating false memories, and c) then examine the source-monitoring framework of memory retrieval and creation that may contribute to the creation of internalized false confessions. Finally, the present study will be introduced which compares multiple forms of false evidence and imagery and their ability to have a participant provide an internalized false confession.

**Internalized False Confessions**

Generally, people know of two types of false confessions: voluntary and coerced-compliant false confessions (McCann, 1998). Those who voluntarily confess to a crime do so without any pressure and is internally driven, usually motivated by attempting to protect the true guilty party, because of mental illness or delusion, or a desire to gain notoriety. Coerced-compliant false confessions are given “when a suspect confesses in order to escape or avoid an aversive interrogation or to gain a promised reward” (Kassin, 1997, p. 225). This happens when a suspect confesses to a crime, not because they believe they are guilty, but to gain immediate benefits that seem to outweigh the consequences of confessing. This can include ending the interrogation, receiving sleep and food, or having been promised leniency for confessing.

There is a third type of false confession, identified as coerced-internalized false confessions, where the individual begins to believe that they actually committed the crime they are accused of by means of duress and coercive interrogation techniques (McCann, 1998). False confessions in general have been found to be an extremely potent factor in convicting a
defendant, and internalized false confessions may make them even more so. Because a defendant has confused their memories in an internalized false confession and may have been given specific and true facts during interrogation that were incorporated into their confession, they are more likely to give a more detailed confession than those who simply admitted to the crime to escape an interrogation (Kassin, 2007). Internalizing the crime then may cause the defendant to continually admit their guilt to multiple individuals and not just during interrogation, allowing multiple “confessions” to be presented as evidence, whereas only coerced-compliant confessors may proclaim their innocence after their initial false confession. Internalization could also lead the confessor to debate whether or not to appeal their case or retract their confession because they still are not truly certain of their innocence. Having more detailed and multiple confessions, may influence a judge or jury to believe that the defendant is guilty and, in turn, discount any retraction or evidence of coercion (Kassin S., 2007). As Kassin (2007) also pointed out, juries may not understand the difference between an internalized false confession and a coerced-compliant confession because “they result from a deeper, more profound, less intuitive form of social influence” (p. 182). The presentation of false evidence during an interrogation has been found to be a powerful technique to elicit these types of false confessions (Horselenberg et al., 1996; Kassin & Kiechel, 1996; Nash & Wade, 2008).

**False Evidence and Internalized False Confessions**

Presenting a suspect with false evidence confuses an individual’s memories between what they remember happening, and what is being physically presented as contradictory fact. Most research that has been conducted on false confessions and the presentation of false evidence has focused on eyewitness evidence and doctored videos. Each type will be discussed
in turn, as well the difference between when an individual is actually shown the evidence, and only told that evidence exists.

**Eyewitness evidence.** Kassin and Kiechel (1996) created one of the first paradigms to examine the effect of false evidence on false confessions and internalization. For this study, they manipulated false eyewitness evidence and stress level to elicit an internalized false confession. Participants were asked to complete a typing task on a computer under different stress levels. They were specifically told not to press the ALT-key on the keyboard or else the computer would crash and their data would be lost. Sixty seconds into the experiment, the computer would automatically crash, and the researcher would accuse them of pressing the ALT-key. For half of the participants, the researcher would then present eyewitness testimony from a confederate who claimed they saw the participant press the key. The other half of participants were not given any false eyewitness account or other evidence. Students were then asked to sign a confession and were sent to a waiting room where a second confederate asked them what happened and recorded their conversation. Their responses were coded to identify internalization – belief that they committed the act – and any subsequent confabulation – ‘remembering’ specific details of the act that did not happen.

They found that 69% of participants willingly signed the confession where 28% internalized the incident and 9% provided confabulated details. They also found that those in the high stress level condition and false evidence group were more likely to sign the confession (100%) and internalize (65%) the incident. The presence of a witness (false evidence) alone was found to significantly increase compliance and internalization.
Horselenberg et al. (1996) also used the ALT-key paradigm and eyewitness evidence to examine the effect that individual differences such as susceptibility to leading questions, as well as the severity of consequences for confessing, have on false confessions. They found that even when participants are punished for their confession and would lose most of their participation fee, 82% still signed the confession where 42% internalized the confession and 58% then confabulated details. When examining the influence of individual differences in their study, none of the individual differences tested were significant. Subsequent studies have only found “susceptibility to leading questions” to be a significant individual difference (Forrest et al., 2005), showing that though only certain individual differences may lead to an internalized false confession, the presentation of false evidence is a much greater influence on confession. The presentation of false eyewitness evidence potentially confused the memories of the participants enough to internalize pressing the ALT-key, so that they confessed even when they faced punishment.

**Doctored video.** The previous studies only used eyewitness information as false evidence, but police can use almost any form of false evidence, including creating false evidence to show them (Najdowski & Bonventre, 2014). To expand the research on false confessions and the presentation of other types of false evidence, Nash and Wade (2008) created a new paradigm to research false confessions and increase the stakes of confessing. They created a gambling paradigm and told participants they were recruited for a study on differences in gambling behaviors between when people can actually handle their money and those who only theoretically know how much they have. All participants were told they were in the “physical money” condition, and they were given a stack of fake money and another stack was placed on a desk as a “bank.” They then completed a gambling task where they answered 15 questions on a
computer and placed a bet on their answer being correct. When they got a question wrong they had to pay the bank, and each time they got a question correct they took money from the bank. Each participant was filmed under the guise of observing their behavior while gambling.

Afterwards, participants were told for one of the questions they answered incorrectly, they took money from the bank when they should have returned it. Half of the participants were shown a doctored-video of them stealing money, while the other half were only told that the video caught them cheating. They were asked to sign a confession and told that they would not receive their compensation for participating, and then sent to a waiting room where a confederate asked them what happened. Their conversations with the confederate were coded for full and partial internalization and confabulation. All of the participants signed the confession, and 20% partially internalized the event, and 63% fully internalized the event. Those who actually viewed the edited video were more likely to have some form of internalization (i.e., partial or full internalization), but both groups were equally likely to fully internalize the crime.

One limitation of the study was that participants may have confessed to stealing the money because they thought they took the money by accident and did not consciously mean to take more than they should. To address this issue and to mimic the usually implausible allegations of a real crime on an innocent individual, the researchers conducted a second experiment where they accused participants of stealing money on three questions answered incorrectly. Ninety-three percent of participants signed the confession and took responsibility for stealing money three times, with 30% partially internalizing the event and 43% fully internalizing the event. Ten percent then partially confabulated and hypothesized what might have happened, and 7% fully confabulated details.
Shown versus told evidence. In Nash and Wade (2008), both those who saw the video and only told of the video were equally likely to confess, but those who saw the video were more likely to fully internalize the event than those only told about the video. Subsequently, those who were only told of the video were more likely to offer some form of confabulation (i.e., either partial or full) more than those who saw the video. This difference could be due to the differences of each medium leading to “acceptance” as internalization or “false memory.” As will be discussed, being given evidence of your wrongdoing may lead a person to simply accept this evidence though it conflicts with their actual memory—the evidence then becomes their “memory.” Only being told information and subsequently imagining the evidence and event together could lead to source attribution errors, where the information from an imagined event is assimilated into their actual memory, and they begin to create a false memory. Being shown the evidence may lead to more full internalization because the individual simply accepts the evidence as fact, and there is no need to confabulate or come up with details. On the other hand, only being told of an event may lead to only partial internalization because the person is then forced to imagine the event and what could have happened to lead to this information, creating some type of confabulation.

Imagery and False Memories

As indicated by the aforementioned research, showing or telling a suspect about false evidence can lead to internalization, or believing that the participant actually committed an offense. However, since false evidence may not always be readily available and is harder to produce, police are more often encouraged to use imagery and repetitive visualization accompanied with key facts of the crime during interrogation (Walters, 1996). That is, police officers may ask the suspect to imagine and tell them what might have happened during the
crime. Generally, individuals may be skeptical of the potency of imagination to induce a false confession, but research has shown that simply imagining an event that did not happen can create false memories and produce internalization (Garry & Wade, 2005; Garry, Manning, Loftus, & Sherman, 1996; Johnson, Foley, Suengas, & Raye, 1988).

**Internalization and confabulation.** Many studies that examine internalization also try to identify confabulations (Horselenberg, Merckelbach, & Josephs, 2003; Kassin & Kiechel, 1996; Nash & Wade, 2009; Shaw & Porter, 2015). Confabulations are a potential byproduct of internalization, where an individual provides false details in accordance with their perceived guilt. Shaw and Porter (2015) pointed out that confabulation may contribute to one of two different forms of internalization. When presented with false evidence of a crime, some participants may simply “accept” the evidence and believe in their guilt simply due to the new evidence, though they have no memory of it. Other participants may take that evidence and confabulate details, actually creating a false memory of the crime (Shaw & Porter, 2015).

When an individual is only told of evidence, or is asked to imagine an event, they have to engage in reality monitoring, or “processes involved in discriminating between memories of real and imagined events” (Johnson, 1988, p. 390). Memories can have two different sources – internal, or imagined and thought based, and external, or perceptual experiences – and source attribution errors, where a person is unable to distinguish between real and imagined memories, could occur. These errors may confuse a defendant when trying to differentiate between what they remember happened and what police told or guided them to imagine. Giving suspects plausible “facts” that can easily be incorporated into a memory can make reality monitoring harder for an individual to engage in (Henkel & Coffman, 2004). Highly emotional and stressful situations have also been found to increase source attribution errors, and these false memories
are very convincing especially when memories are retrieved through heuristic judgments that rely on qualitative attributes such as emotion, perception, and context (Henkel & Coffman, 2004).

Raye and Johnson (1980) examined reality monitoring and the individual’s ability to identify the source of a memory when there are two external sources (i.e., remembering which of two people said a word), and having one external (i.e., remembering words said by one person) and one internal source (i.e., remembering words they said themselves). They found that those who had both an internal and external source were significantly better at distinguishing who said what word than those who had two external sources. In an interrogation where false evidence is presented, an individual may accept the evidence (i.e., external source) as fact, even though they know it conflicts with their real memory (i.e., internal source). But when having to differentiate between an imagined event and their memory (i.e., two internal sources), source attribution errors may occur more often. Garry et al. (1996) found this to be true when they asked participants to imagine a childhood event that they know the participant did not experience, such as being saved by a lifeguard or punching through a window. After repeated visualization, participants were more likely to become more confident that the event did actually happen.

**Imagery and internalized false confessions.** The aforementioned studies only asked participants to imagine an emotional childhood event. Shaw and Porter (2015) conducted a study that actually examined the effectiveness of imagination on internalizing and creating a false memory of a crime. They conducted three interviews with participants where they presented them with two events with key information (e.g., where they lived, the season it took place during) that happened in their childhood- one true, emotional event, and one fake event that lead to police contact (i.e., committing a theft, committing an assault, committing an assault
with a weapon). The participants were then asked to describe each of the events, and when they could not recall the criminal act (as none of the participants recalled the event during the initial interview), the researcher encouraged them to try and remember and visualize what had happened. At the conclusion of the third interview, 70% of participants were deemed to have created a false memory, providing on average 71.76 false details of the criminal act. This suggests that imagery can be a powerful tool to elicit false confessions, and subsequent studies have found that imagery can sometimes be more influential than false evidence (Garry & Wade, 2005; Nash, Wade, & Lindsay, 2009).

**False evidence vs. imagination.** Though this study did not use a false criminal act, Garry and Wade (2005) also conducted a study on source attribution errors that pitted false, doctored-photographs (false evidence) and false narratives (imagination). They had 44 undergraduates recruit a family member who had never experienced a hot-air balloon ride and randomly assigned them to either false photo or false narrative. The researchers created a booklet containing photos and narratives of three real events that the participant experienced, and either a Photoshopped false picture of them on a hot-air balloon ride, or a 45 word long false narrative of the hot-air balloon ride. They were interviewed multiple times to report what they could recall about each event, their confidence that the event occurred, and which medium was better at jogging their memory. In support of their hypothesis, false narratives were more likely to generate false information and confabulate details for 82% of participants as opposed to only 50% of those who were provided doctored (fake) photos. However, as addressed by the authors, these findings could be due to the constraints of each medium, where pictures provide specific details and leave little room for imagination, and narratives allow the individual to form more false information because of how little is provided. Nash et al. (2009) also found that
imagination and false evidence were not significantly different in creating a false memory of having performed a specific action one week previously. This again highlights the distinction between acceptance of evidence, and the production of confabulation between actually seeing evidence and imagining an event.

As seen by the previous literature, imagination and visualization can be a powerful tool in creating false memories and eliciting an internalized false confession. It has been found, in some cases, to be equally as effective as false evidence in creating a false memory (Nash, Wade, & Lindsay, 2009), and in some studies has been found to be better than false evidence in eliciting a more detailed narrative and false memory (Garry & Wade, 2005; Shaw & Porter, 2015).

Present Study

Prior research has shown that false evidence may lead to more acceptance in terms of an internalized false confession (Kassin & Kiechel, 1996; Garry & Wade, 2005; Horselenberg, Merckelbach, & Josephs, 2003; Nash & Wade, 2009) and imagery elicits more confabulation and creation of a false memory (Garry & Wade, 2005; Garry, Manning, Loftus, & Sherman, 1996; Shaw & Porter, 2015). Unfortunately, the extant literature lacks a direct comparison of false evidence and imagery in eliciting an internalized false confession of an immoral or criminal act, and the previous studies that examine the use of false evidence in eliciting a false confession only used one form of false evidence each (e.g., eyewitness evidence or doctored video). The present study aims to fill this gap and add to the existing literature on the ability of each technique to elicit an internalized false confession. This study will replicate Nash and Wade’s (2008) gambling paradigm to further explore and directly compare the effects of previously studied forms of false evidence as well as imagery. Due to the lack of research comparing each
technique in eliciting internalized false confessions, we remain agnostic as to which will elicit more internalized false confessions.

Method

Design

This study randomly assigned participants to one of five manipulations: being shown a doctored-video of the participant stealing money (shown-video), merely being told that there is video evidence of them stealing money (told-video), having a research assistant as an eyewitness directly tell the participant they saw them steal money (shown-witness), being told that the research assistant had seen them stealing money (told-witness), and being instructed to imagine the event and what might have happened when they stole the money (imagery). Three dependent variables were measured (i.e., confession, internalization, confabulation) with three levels each (i.e., no confession, confession on initial prompt, confession on second prompt; full internalization, partial, or none; full confabulation, partial, or none; respectively). Frequencies of each dependent variable were counted, and qualitative analysis was used to examine participant responses and obtain frequencies of internalization and confabulation. Manipulation checks were conducted by running the first six participants as a test-run of each condition and determining the believability of each manipulation.

Participants

Fifty-nine undergraduate psychology students from an urban Northeast college participated in the present study. All participants received course credit for their participation and were randomly assigned to one of the five manipulations. The mean age of participants was 19.73 (SD = 2.97) with a range of 18-32. The sample was majority female (78%), with 24%
being Caucasian, 22% Hispanic, 19% African-American, 8% Asian, and 7% identifying as “other.” The majority of participants were freshman (59%) taking Psychology 101.

**Procedure and Materials**

This study replicated the gambling paradigm used by Nash and Wade (2008). Each session was conducted within the lab facilities at an urban Northeast college and each participant was run individually. One researcher, one confederate, and one research assistant conducted each session and followed a script to maintain uniformity across participants.

**Gambling quiz.** Upon arriving to the lab, participants entered a small room with a computer and video camera. Participants read and signed an informed consent form to participate in and be videotaped during the study. Immediately after signing the consent form, the researcher told the participants that this was a study examining the difference in gambling habits and behaviors between individuals who can actually see and handle physical money, and those who only theoretically know how much they have by way of a computerized score. The researcher told each participant that they were placed in the “physical money” condition and were given a pile of fake money to gamble with, and another pile of fake money on the desk to represent a “bank.” They were then informed that as an incentive to win as much money as possible, the money they were initially given was what the prior participant had won, and the amount the participant wins is what the next participant will start off with. This was included to increase the severity of the cheating allegation.

Next, the researcher explained that the participants would complete a 20-question, multiple-choice quiz with trivia questions ranging in topic, such as popular culture, historical facts, and animal knowledge (see Appendix A). This was a computerized questionnaire adapted
from Nash and Wade (2008) with four possible answer choices for each question. Each question was accompanied by an odds ratio that dictated how much the participant would win or lose based on their chosen bet. For example, participants were asked, “What is a group of rhinoceros called?” with the choices, a “herd” (odds = 2:1), a “crash” (odds = 3:1), a “coalition” (odds = 5:1), and a “clattering” (odds = 10:1).

For each question, the participant was to place a monetary bet of their choosing on their answer being correct. If they answered a question correctly, the computer screen would show a large green checkmark and they would be instructed to take their winnings from the bank. If they answered incorrectly, a red ‘X’ would appear on the screen and they would be told to pay the bank from their own pile. Participants were told that each potential answer choice had an odds ratio accompanied with it and the researcher explained that the odds dictated how much they would win or lose based on the participant’s monetary bet. They were warned not to bet any more than what they possessed or they would not be able to pay the bank. The participant was then shown a sample question before beginning the study to see the format of the quiz.

The participant was then given the chance to ask the researcher any questions. When the participant was ready, the researcher then started the video camera if the participant was in one of the video conditions and told the participant to begin the quiz. The researcher remained in the room at a separate desk. If the participant was in one of the witness conditions, the primary researcher would leave the room under the guise of having other work and instructed a research assistant to observe the participant’s behavior.

After the participant finished the quiz, they were instructed to count their money and tell the researcher how much they had won. If the participant was in the eyewitness or imagery
conditions, the participant was then instructed to sit in a waiting room while the researcher was ostensibly transferring their quiz data to their personal computer. If they were in one of the video conditions, they were also told that their video had to be saved, and that the researcher would call them back in to be debriefed.

**Doctored video.** Participants waited for 15 minutes before they were brought back in and accused of cheating. These 15 minutes was deemed to be a reasonable amount of time for the researcher to ostensibly converse with their advisor about possible solutions to the problem. If the participant was placed in the shown-video condition, this time was also taken to edit video footage of the participant. To edit the video footage, the researcher used Final Cut Pro on an iMac laptop and took a short 10- to 20- second clip of a question that the participant had answered correctly so the green checkmark was visible on the computer screen and the participant took money from the bank. The researcher then digitally over-layed a short clip of the red ‘X’ from a question that the participant had answered incorrectly over the green checkmark so it appeared as if the participant was taking money from the bank when they had gotten a question wrong. If the participant was in the told-witness condition, the research assistant would leave the room and wait in a separate lobby until the conclusion of the trial.

**Confession form.** After the 15-minute waiting period, the participant was taken back to the room and was told that something went wrong with their trial. If they were in one of the video conditions, the researcher informed the participant that upon review of the video, it showed the participant taking money from the bank on three questions that they had answered incorrectly. Shown-video participants were then shown the doctored-video of them taking money on one of their incorrectly answered questions. They were allowed to view the video twice if asked. Told-video participants were told that the video clearly showed the red ‘X’ and
them taking money from the bank. If they asked to see the video, they were told that was impossible because they were still in the middle of their study session.

If they were in one of the eyewitness conditions, the researcher told them that the research assistant in the room had told them that they had seen the participant take money from the bank on three questions they got wrong. Shown-witness participants were then told directly by the research assistant that they had looked up at their quiz a few times and had seen the red ‘X’ on the screen and the participant taking money from the bank on three questions. In the told-witness condition, the research assistant was absent from the room, and if the participant asked to speak with them, they were told that the assistant had to leave due to a prior engagement. If the participant was in the imagery condition, they were told by the researcher that there was a very large discrepancy between what the quiz’s data said the participant should have won, and what the participant actually reported as having won. In reality, the quiz does not keep track of the participant’s winnings, nor does it collect any data on the participant’s answers or bets. They were then told that though the quiz did not keep track of all the participant’s data, it was determined that for three questions the participant answered incorrectly, they must have taken money from the bank instead of paying the bank. The participant was then asked to try and imagine or visualize what happened, with suggestions as to try and remember which questions they got wrong, how much money they could have taken, and even what the researcher was doing during that time.

The participant was told that, even though the study used fake money, this was a problem because the participant cheated, so the data collected from their quiz was invalid. They were also told that because the next participant would receive an incorrect amount of money, the next participant’s data was also invalid. Because of this problem, the researcher stated that they had
to call their supervisor and ask how to handle the situation and were “told” that the participant had to fill out a “conditional payment form” (Appendix B).

The participants were then given the “conditional payment form” that they were asked to sign. Participants were told that this was a general form used by the college in cases when a mistake was made on the part of a participant, (Nash & Wade, 2008) and acted as the confession for the study. The form was filled in by hand for each participant by the researcher with the paragraph, “During the trial, participant took credit from the ‘bank’ on three separate questions when they were instructed to return it. Due to the nature of the study, this trial data cannot be used. Data from the subsequent trial is also invalid.” Participants were told that as long as they signed the paper and took full responsibility for taking the money, they would receive their class credit for participating in the study. If they refused to sign on the first prompt, they were told that they would have to have a meeting with their psychology professor who would review the alleged evidence and would then decide if the participant would receive the credit. If the participant refused to initially sign the form, they were then prompted a second time to sign the form.

**Internalization.** After the accusation and confession, the participant was told to sit in the waiting room again while the participant contacted their advisor again to make sure they had done everything that was needed. A confederate posing as the next participant was also sitting in the waiting room and would ask the participant what happened. Each conversation was surreptitiously audio recorded. The confederate, blind to the participant’s assigned condition, followed a vague script and was allowed to ad lib in order to respond to various things the participant would say and ask.
Demographics questionnaire. After 5 minutes while the participant spoke with the confederate, the participant was then escorted back into the room and was thoroughly debriefed on the true nature of study by the primary researcher. Participants were then asked to complete an ex post facto consent form for the use of the audio recorded conversation with the confederate. Participants were then asked to complete a short demographics questionnaire (Appendix C), indicating their age, sex, race, and educational year, as well as indicate on a Likert Scale how much they believed that they actually took more money than they were supposed to. Finally, they were given a debrief and experimental design statement to keep and review.

Internalization coding. To determine whether internalization occurred, the audio recorded conversations with the confederate were transcribed into written transcripts. These transcripts were coded into full internalization, partial internalization, or no internalization using the coding method used by Nash and Wade (2008). The researcher and confederates read through transcripts of the audio recorded conversations, and highlighted sentences within the conversation. Sentences were deemed as full internalization if the participant said they definitively believed they committed the act (e.g., “I stole money from the bank when I shouldn’t have”), and partial internalization if they believed that they may have committed the act (e.g., “I think I might have messed up”). Sentences were also deemed to show no internalization if the participant stated they did not cheat or only reported what the researcher told them (e.g., “She told me I cheated, but I didn’t”). Statements were also coded for confabulation, or creating false, specific details of the crime. They were coded as full confabulation if participants stated that they committed the crime and they provided details (e.g., “I was concentrating too hard and didn’t notice my mistake”) and partial confabulation if they speculated or rationalized as to how they may have stolen the money (e.g., “I must have zoned-out and forgot to pay the bank and
took instead”). If the participant gave no confabulating details or rationalizations, they were determined to have no confabulation. Based on the majority of internalization or confabulation statements made by the participant, coders deemed the participant to have fully, partially, or not internalized or confabulated the accusation.

Results

Across all manipulations, 91.5% of participants signed the confession on either the first or second request, with 54.2% confessing on the first request and 37.3% on the second. For internalization, 59.3% of the sample showed some form of internalization: 28.8% fully internalized the crime and 30.5% partially internalized the crime. Finally, 28.8% of participants confabulated details where 3.4% fully confabulated the details of the act and 25.4% hypothesized what might have happened.

To analyze which manipulation produced more false confessions, rates of internalization, and rates of confabulation, three 5 (False evidence: shown-video vs. told-video vs. shown-witness, told-witness vs. imagery) x 3 (Confession: None vs. first prompt vs. second prompt; Internalization: full vs. partial vs. none; Confabulation: full vs. partial vs. none) Fisher’s exact tests were conducted. Table 1 shows the percentages and frequencies of compliance, internalization, and confabulation for each of the five manipulations, along with the Fisher’s exact test $p$-value.

Confession

In terms of rate of confessions, there was no main effect across all conditions ($p = .673$). Thus, regardless of condition, all participants were equally likely to sign the confession (Table 1). To examine the effect of any false evidence against imagery, all of the false evidence
conditions were then combined (i.e., shown-video, told-video, shown-witness, told-witness) to form one false evidence category to compare to imagery. A 2 (i.e., false evidence versus imagery) x 2 (i.e., any confession versus no confession) Fisher’s exact test found \( p = .274 \) though, again showing that both the presentation of false evidence and use of imagery were not significantly different in eliciting a false confession. Interestingly, though not statistically significant, 100% of the imagery participants signed the confession form, whereas only 89% of those presented with false evidence did.

**Internalization**

In regards to internalization, there was no main effect across all condition \( (p = .175) \), showing that each manipulation was as equally likely to produce internalization (Table 1). Comparing *any* form of false evidence against imagery in producing *any* form of internalization using a 2 (i.e., false evidence versus imagery) x 2 (i.e., any internalization versus no internalization) Fisher’s exact test also showed no main effect \( (p = .442) \), again finding that imagery is equally as likely to elicit internalization as false evidence.

**Confabulation**

In terms of confabulation, the results of the analysis was approaching statistical significance \( (p = .069) \). Specifically, shown-video and imagery participants were significantly different \( (p = .037) \). Those who imagined the event were more likely to fully confabulate the event, whereas those who were shown the edited footage of them cheating were more likely to have only partial confabulation. Comparing all forms of evidence combined against imagery in examining any form of confabulation using a 2 (i.e., false evidence versus imagery) x 2 (i.e., any confabulation versus no confabulation) Fisher’s exact test did not produce a main effect \( (p = \)
and there were no significant differences between false evidence and imagery participants in terms of producing any confabulation.

**Discussion**

The present study examined whether the presentation of certain types of false evidence (witness vs. video and shown vs. told) and imagery elicit different rates of false confession, internalized false confessions, and confabulations. We remained agnostic in our hypothesis, and instead chose to explore and directly compare the differences of these techniques in rates of confession, internalization, and confabulation.

**Confession.** Overall, our results on compliance and willingness to sign a false confession replicate the findings of previous studies, showing that an overwhelming majority of participants will readily sign a false confession regardless of being presented with false evidence or imagery (Forrest, Wadkins, & Larson, 2005; Horselenberg et al., 1996; Kassin & Kiechel, 1996; Nash & Wade, 2008). Interestingly, 100% of imagery participants signed the confession form whereas only 89% who were presented with false evidence complied. When examining each form of false evidence independently, 100% of told-witness participants also signed the confession form, and told-video had the highest rate of non-confession (20%). We believe this is due to the differences in acceptance versus confabulation and false memory internalization, and replicates the findings of Nash and Wade (2008) where shown-video participants were more likely to confess on the first request. Told-witness participants and imagery participants were given false key facts (e.g., you took money on three questions you got wrong), and though imagery participants were directed to imagine what happened, told-witness participants were unable to confront the witness and may have instead automatically engaged in visualizing what
might have happened. This visualization may have begun the process of creating a false memory leading to at least partial internalization, and subsequent confession.

Those who actually viewed the doctored video did not have to imagine the event and may have simply accepted the evidence as truth. Told-video participants, alternatively, tended to become combative and question the researcher, sometimes stipulating that the researcher was mistaken and they were actually putting money back, or that the researcher did not watch the video correctly. Many of them asked to view the video, but the researcher claimed this was not allowed because they were still in the middle of the trial session. This study was also conducted in a university geared towards criminal justice and psychology, so many participants may have also been aware to question authority and argue for their rights. Because there is an alleged video that they are told about, the participant does not have to engage in imagination like the told-witness or imagery participants, thus not confabulating details and creating a false memory. Instead, they will be more likely to understand that there is an alleged video, but because it conflicts with their memory, they may be more likely to not accept it and instead trust their own memories. But further research is needed to examine these possibilities.

**Internalization and confabulation.** In terms of internalization, there were no significant differences in rates of internalized false confessions across all manipulations. This replicated the findings of Garry and Wade (2005) and Nash, Wade, and Lindsay (2009), indicating that each type of interrogation tactic is equally likely to elicit an internalized false confession. Though the differences were not significant, more participants in the shown-video condition fully internalized the crime. This again seemed to be due to acceptance of the evidence, rather than creation of a false memory or rationalization of what happened. Many participants stated when
asked by the confederate if they remembered taking the money, “Yeah, they have a video of it you know, but I have no memory of it.”

This leads to the significant difference between shown-video participants and imagery participants in terms of confabulation. Though across all manipulations there was no main effect in confabulation, imagery participants were more likely to fully confabulate details than those in the shown-video condition. Many interrogation techniques encourage police to give a suspect some key facts, and then ask them to imagine what might have happened given this information (Henkel & Coffman, 2004). This was emulated in the imagery condition during the accusation in this study by the researcher telling the participant that 1) there was a large discrepancy between what the quiz said they should have won and what they actually won, and 2) we determined that for three questions they got wrong, they actually took money from the bank instead of paying the bank. The participant was then asked to visualize and imagine how this could have happened. Source attribution errors have been found to be more common when an individual has to identify facts from two internal sources (i.e., the false, imagined event versus the perceived memory), and “facts” from an imagined event may erroneously be incorporated into the real memory (Johnson, 1988). This happens more often when the “fact” is plausible, or in highly emotional situations (Henkel & Coffman, 2004). The participant may have made imagined potential reasons for the discrepancy and what happened for three questions, and these rationalizations could have created source attribution errors leading to confabulation. As Shaw and Porter (2015) stated of imagery, being asked to imagine what “could” have happened leads to what “would” have happened, and finally what “did” happen.
Limitations and Future Directions

This study was able to build upon the body of previous research on false confessions by comparing internalization and confabulation using shown-video, told-video, shown-witness, told-witness false evidence, and an imagery technique, whereas previous studies only focused on one or two types of false evidence (Kassin & Kiechel, 1996; Horselenberg et al., 1996; Nash & Wade, 2008). Of course, there are other types of false evidence that interrogators could use, such as fake documents, fake polygraph results, or fake fingerprints, which may have different effects on compliance, internalization, and confabulation, due to the medium and limited details they provide. This study intended to examine the types of evidence previously used in literature, as they have not previously been directly compared. These other types of false evidence should be investigated for future studies on internalized false confessions, as they may be more readily available to police during interrogation.

Also, although this study accused participants of an unlikely series of events (i.e., stealing money on three separate occasions), and gave them a specific negative outcome for not confessing (i.e., possible loss of course credit), we were unable to allege punishment for if they did confess, like previous studies (Horselenberg, Merckelbach, & Josephs, 2003). Some subjects even indicated this lack of punishment or stress involved with confessing in their responses to the confederate in the waiting room through phrases such as, “They said I stole some money, I don’t really care, I just signed the confession to get out of meeting with my professor. I just wanted the credits.” Future research should then give explicit punishment to the participant if they confess in order to mitigate a false confession due to perceived leniency.

Finally, due to time constraints, a small overall sample size and unequal cell size lead to the present analyses to be under powered. Due to scheduling needs of two confederates for those
in the shown- and told-witness conditions, as well as no-show participants, there were only nine participants in the shown-witness condition and seven in the told-witness condition, almost half of the other three groups. It is difficult to truly compare groups of such varying size and thus, the present results should be interpreted with caution. With small group sizes, percentages of the told- and shown-witness conditions may also seem skewed as though the percentage of confession or confabulation may seem high, there may have really only been one or two participants that make up that percentage. This study does have the potential to be continued, so more time and availability of confederates would allow for a larger sample size to be collected and ensure power to test for significant differences.

Conclusion

Our study replicated the findings of previous studies, showing that the presentation of false evidence or use of imagery will elicit high rates of false confessions. This study was also able to build upon and analyze how different forms of false evidence affect internalization and confabulation, as well as how the use of imagery can be a significant factor in producing false memories and false confessions. The presentation of false evidence, in particular false videos, influenced a majority of participants to accept evidence and then confess to a crime they did not commit. Other types of false evidence or imagery can prompt confusion and confabulation of specific details, leading to the creation of a false memory and subsequent internalization. This study then questions whether police should be able to lie and present false evidence to suspects to get a confession, as well as use simple imagery, as it could possibly distort an individual’s memory and, in turn, lead to false confessions.
References

The Innocence Project (2016). Retrieved from The Innocence Project:
http://www.innocenceproject.org/


Table 1

Table 1: Percentages and frequencies of confession, internalization, and confabulation with Fisher’s exact p value

<table>
<thead>
<tr>
<th>Categorization</th>
<th>Condition</th>
<th>Fisher’s exact p value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shown-video</td>
<td>Told-video</td>
</tr>
<tr>
<td>No confession</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confession</td>
<td>First request</td>
<td>60% (9)</td>
</tr>
<tr>
<td></td>
<td>Second request</td>
<td>33.3% (5)</td>
</tr>
<tr>
<td>No internalization</td>
<td></td>
<td>20% (3)</td>
</tr>
<tr>
<td>Internalization</td>
<td>Partial internalization</td>
<td>20% (3)</td>
</tr>
<tr>
<td></td>
<td>Full internalization</td>
<td>60% (9)</td>
</tr>
<tr>
<td>No confabulation</td>
<td></td>
<td>53.3% (8)</td>
</tr>
<tr>
<td>Confabulation</td>
<td>Partial confabulation</td>
<td>46.7% (7)</td>
</tr>
<tr>
<td></td>
<td>Full internalization</td>
<td>0% (0)</td>
</tr>
</tbody>
</table>

*Fisher’s exact p value*
Q1: To the nearest thousand, what is the population of the Republic of Montenegro, formerly part of Yugoslavia?

Think carefully, then place a bet on ONE answer ONLY!

(A) 620 thousand (Odds = 2:1)

(B) 310 thousand (Odds = 3:1)

(C) 62 thousand (Odds = 5:1)

(D) 31 thousand (Odds = 10:1)

Q2: Which country was runner-up in the 1982 Football World Cup?

Think carefully, then place a bet on ONE answer ONLY!

(A) Brazil (Odds = 2:1)

(B) Italy (Odds = 3:1)

(C) West Germany (Odds = 5:1)

(D) Poland (Odds = 10:1)
Q3: What is the hottest planet in our solar system?

Think carefully, then place a bet on ONE answer ONLY!

(A) Mercury (Odds = 2:1)
(B) Jupiter (Odds = 3:1)
(C) Saturn (Odds = 5:1)
(D) Venus (Odds = 10:1)

Q4: The average human brain weighs roughly what percentage of a person's total weight?

Think carefully, then place a bet on ONE answer ONLY!

(A) 2% (Odds = 2:1)
(B) 14% (Odds = 3:1)
(C) 26% (Odds = 5:1)
(D) 38% (Odds = 10:1)
Q5: What is a group of rhinoceros called?

Think carefully, then place a bet on ONE answer ONLY!

(A) A "herd" (Odds = 2:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\

(B) A "crash" (Odds = 3:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\

(C) A "coalition" (Odds = 5:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\

(D) A "clattering" (Odds = 10:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\

Q6: Of what is "Rhytiphobia" the fear?

Think carefully, then place a bet on ONE answer ONLY!

(A) Getting wrinkles (Odds = 2:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\

(B) Getting dirty (Odds = 3:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\

(C) Getting undressed (Odds = 5:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\

(D) Getting leprosy (Odds = 10:1) $\_\_\_\_\_\_\_\_\_\_\_\_\_\_
Q7: Who was the 38th president of the United States?

Think carefully, then place a bet on ONE answer ONLY!

(A) Lyndon B. Johnson (Odds = 2:1)

(B) Gerald Ford (Odds = 3:1)

(C) Richard Nixon (Odds = 5:1)

(D) Jimmy Carter (Odds = 10:1)

Q8: The first human-made object to land on the moon was launched by which country?

Think carefully, then place a bet on ONE answer ONLY!

(A) The United States (Odds = 2:1)

(B) China (Odds = 3:1)

(C) The Soviet Union (Odds = 5:1)

(D) Germany (Odds = 10:1)
Q9: In the movie "The Lion King," what was Simba's mother's name?

Think carefully, then place a bet on ONE answer ONLY!

(A) Sarabi (Odds = 2:1) $     Enter a bet

(B) Sarafina (Odds = 3:1) $     Enter a bet

(C) Shenzi (Odds = 5:1) $     Enter a bet

(D) Zazu (Odds = 10:1) $     Enter a bet

Q10: What song by Michael Jackson contains the lyrics, "Annie are you OK?"

Think carefully, then place a bet on ONE answer ONLY!

(A) Beat It (Odds = 2:1) $     Enter a bet

(B) Smooth Criminal (Odds = 3:1) $     Enter a bet

(C) Man in the Mirror (Odds = 5:1) $     Enter a bet

(D) Billie Jean (Odds = 10:1) $     Enter a bet
Q11: What is the most popular pizza topping at 'Domino’s Pizza’ in Japan?

Think carefully, then place a bet on ONE answer ONLY!

(A) Anchovies (Odds = 2:1) $ Enter a bet

(B) Tuna (Odds = 3:1) $ Enter a bet

(C) Salmon (Odds = 5:1) $ Enter a bet

(D) Squid (Odds = 10:1) $ Enter a bet

Q12: "Do you expect me to talk?" "No, I expect you to die." is a quote from which Bond movie?

Think carefully, then place a bet on ONE answer ONLY!

(A) Goldfinger (Odds = 2:1) $ Enter a bet

(B) Thunderball (Odds = 3:1) $ Enter a bet

(C) Diamonds are Forever (Odds = 5:1) $ Enter a bet

(D) The Spy Who Loved Me (Odds = 10:1) $ Enter a bet
Q13: What is the approximate gestation period of a goat?

Think carefully, then place a bet on ONE answer ONLY!

(A) 2 months (Odds = 2:1)  Enter a bet
(B) 5 months (Odds = 3:1) Enter a bet
(C) 8 months (Odds = 5:1) Enter a bet
(D) 11 months (Odds = 10:1) Enter a bet

Q14: The tugrik is the currency of which country?

Think carefully, then place a bet on ONE answer ONLY!

(A) Myanmar (Odds = 2:1) Enter a bet
(B) Malawi (Odds = 3:1) Enter a bet
(C) Mongolia (Odds = 5:1) Enter a bet
(D) Malta (Odds = 10:1) Enter a bet
Q15: Who was the only English Pope?

Think carefully, then place a bet on ONE answer ONLY!

(A) Adrian IV (Odds = 2:1)

(B) Clement VI (Odds = 3:1)

(C) Innocent IX (Odds = 5:1)

(D) John Paul I (Odds = 10:1)

Q16: Who was the author of the satirical paper, "A Modest Proposal"?

Think carefully, then place a bet on ONE answer ONLY!

(A) Jonathan Swift (Odds = 2:1)

(B) Voltaire (Odds = 3:1)

(C) William Godwin (Odds = 5:1)

(D) Rebecca Dingley (Odds = 10:1)
Q17: What is the sleepiest animal in the world, sleeping around 22 hours each day?

Think carefully, then place a bet on ONE answer ONLY!

(A) **Squirrel** (Odds = 5:1) $ Enter a bet

(B) **Koala** (Odds = 3:1) $ Enter a bet

(C) **Sloth** (Odds = 2:1) $ Enter a bet

(D) **Tiger** (Odds = 10:1) $ Enter a bet

Q18: Which team won the National Football League’s first Super Bowl?

Think carefully, then place a bet on ONE answer ONLY!

(A) **Pittsburgh Steelers** (Odds = 2:1) $ Enter a bet

(B) **Denver Broncos** (Odds = 3:1) $ Enter a bet

(C) **Greenbay Packers** (Odds = 5:1) $ Enter a bet

(D) **Cleveland Browns** (Odds = 10:1) $ Enter a bet
Q19: 1,024 Gigabytes is equal to one what?

Think carefully, then place a bet on ONE answer ONLY!

(A) **Terabyte** (Odds = 2:1) $\underline{\phantom{9999}}$ Enter a bet

(B) **Megabyte** (Odds = 3:1) $\underline{\phantom{9999}}$ Enter a bet

(C) **Kilobyte** (Odds = 5:1) $\underline{\phantom{9999}}$ Enter a bet

(D) **Petabyte** (Odds = 10:1) $\underline{\phantom{9999}}$ Enter a bet

Q20: The Black Forest is located in what European country?

Think carefully, then place a bet on ONE answer ONLY!

(A) **Germany** (Odds = 5:1) $\underline{\phantom{9999}}$ Enter a bet

(B) **Sweden** (Odds = 3:1) $\underline{\phantom{9999}}$ Enter a bet

(C) **Poland** (Odds = 10:1) $\underline{\phantom{9999}}$ Enter a bet

(D) **France** (Odds = 2:1) $\underline{\phantom{9999}}$ Enter a bet
Correct! Take $16 from the bank and then continue.
Wrong! Return $8 to the bank and then continue

Continue to Question 15
Conditional payment of credit for research participants

SECTION A: Details

Researcher: ____________________________________

Experiment: _________________________________________________________

Date: ___/___/_____

Participant’s name: _____________________________

SECTION B: Reason for conditional receipt of credit/payment:
(please write clearly)

During the trial, participant took credit from the ‘bank’ on three separate questions when they were instructed to return it. Due to the nature of the study, this trial data cannot be used. Data from subsequent trial is also invalid. (handwritten by researcher)

________________________________
________________________________
________________________________

SECTION C: Declaration by participant – PLEASE READ CAREFULLY

I confirm that I consent to only receive any credit or payment due to me for participation in the research project stated if I sign below, for the reason(s) stated in Section B above.

Name: ____________________________ Signature: ________________________ Date: ___/___/_____
SECTION D: Authorisation

Refusal authorised by: _________________________________
(Print name and title)

Signature: ___________________________________________

Date: ___/___/____
Appendix C

Demographics Questionnaire

Age: _______

Gender: _______

Education Year:

1) Freshman
2) Sophomore
3) Junior
4) Senior

Ethnicity:

1) Caucasian
2) Hispanic/Latino
3) African American
4) Native American or American Indian
5) Asian/Pacific Islander
6) Other: _______

To what extent did you believe you actually took money from the bank?

1 2 3 4 5 6 7
Not at all Completely