Do Beliefs Alter the Relationship between the Confidence and Accuracy of Facial Recognition Judgements?

by

Olivia T. Webb

A thesis submitted to the faculty of Radford University in partial fulfillment of the requirements of the degree of Master of Arts in the Department of Psychology

Thesis Advisor: Dr. Kathleen Arnold

March 2025

Dr. Kathleen Arnold
Thesis Advisor

Date 4/25/2025

Dr. Daniel Berry
Committee Member

Date 4/23/2025

Dr. Thomas Pierce
Committee Member

Date 4/23/2025

Dr. Michalea Clark
Committee Member

Date 4/23/2025

Abstract

There is a growing argument that beliefs can alter our metacognitive judgments (Benjamin et al., 1998; Koriat & Bjork, 2006; Mueller & Dunlosky, 2017; Rhodes & Castel, 2008). This theory-based judgment approach has been primarily studied using judgments of learning (JOLs); however, confidence has not been examined. My thesis aimed to examine the effects of beliefs in general memory on the confidence-accuracy relationship. My primary hypotheses are that altering one's belief about memory will alter their confidence in their own memory, but that it will not affect the accuracy of their memory. This will lead to a difference in the confidence-accuracy relationship. My secondary hypothesis examines why this causal relationship is occurring, by predicting that this relationship is mediated by the participants beliefs about recognition memory. Participants were exposed to one of two types of messaging that were designed to alter their beliefs in the accuracy of facial recognition memory: that people are generally very good at facial recognition or that people are generally very poor at facial recognition. They participated in an old/new facial recognition test and rated their confidence in memory for each face. Results indicate that there is no difference in confidence between the two conditions and there was a strong relationship between confidence and accuracy in both conditions. This relationship was not affected by the beliefs condition. There is indirect-only mediation with belief altering statements influencing the criterion/confidence-accuracy relationship, only through the mediator of beliefs about memory. These findings can be applied to eyewitness identification and can help determine what factors must be considered when using confidence to measure recognition accuracy.

Keywords: Confidence, Accuracy, Recognition, Metacognition.

Olivia T. Webb
Department of Psychology, 2025
Radford University

Table of Contents

List of Figures

List of Tables

Do Beliefs Alter the Relationship Between the Confidence and Accuracy of Facial Recognition Judgements?

United States Supreme Court Justice Harry Blackmun once stated that “our criminal justice system is less than perfect” and “human error is inevitable” (*Callins v. Collins*, 1994). As of 2023, the National Registry of Exonerations recorded over 3,000 cases of wrongful convictions (“The Impact,” 2023). The Innocence Project is an organization that helps exonerate individuals who were falsely convicted of crimes. Nationally, 69% of false convictions from the Innocent’s Project were due to eyewitness misidentifications (Innocence Project, 2020). The Supreme Court first established the rules for evaluating eyewitness misidentifications in 1977 and later updated the rules in 2012; however, they have not revised the ruling to stay updated with the current science and statistics on exonerations (*Manson v. Brathwaite*, 1977; *Perry v. New Hampshire*, 2012).

Most mistaken identifications are not intentional. This can be seen in the case of Ronald Cotton. He was wrongfully convicted of raping Jennifer Tompson-Cannino in 1984. A large portion of the case rested on the eyewitness testimony of the victim, and he was sentenced to life in prison plus fifty-four years. He was eventually exonerated due to DNA evidence (Innocence Project, 2023). In later interviews Jennifer discusses how she pictures Ronald as her rapist even after learning that he was innocent and meeting him. The repeated recall and the suggestions by police led to her developing false memories.

Loftus and Ketcham (1991) reported a similar case with a kidnapping witness, John Picha. Picha witnessed the kidnapping of a 7-year-old boy in 1987. When viewing lineups, Picha originally pointed the finger at a hotel employee with an alibi. Two months later, he viewed a new lineup with a different man named, Howard Haupt. After repeated interviews, Picha identified Haupt with high confidence. Haupt was later found to be not guilty of the crime due to insufficient evidence. All witnesses were exposed to post event misinformation and analysis of initial interview transcripts showed little to no confidence in their answers. Witnesses used phrases like, “[Haupt’s] the closest,” and investigators used leading questions to get witnesses to pick Haupt. This is yet another example of how suggestable our memories can be.

Our memories are susceptible to internal and external influences and represent a constructive process, not an exact reproduction of the past. Thus, we can often make errors or distort past experiences (Schacter, 2012). Bartlett’s (1932) famous study states that these distortions are an adaptive process to assist in recall. In this study, he had college students listen to a Native American folk tale about a battle with ghosts. British participants were then asked to recall as much of the story as they could. Bartlett (1932) found that the students tended to add new or altered information to the story to better fit their cultural schema. We create schemas to help organize our memories and relate them to an already known concept or theme. A schema is a framework for our knowledge that we rapidly develop and use to help make quick judgments about the world (Tompary et al., 2020). Schemas have been shown to help with accurate encoding, storage, and retrieval of information that aligns with a specific schema (Alba & Hasher, 1983; Brewer & Treyens, 1981; Castel, 2005; Davenport & Potter, 2004; Mandler, 1984; Miller & Gazzaniga, 1998; Palmer, 1975).

Schematic memories are beneficial overall; however, they can lead to errors. An example of this type of error would be if a person is trying to remember a party. Overall, the schema will help them recall more aspects of the party because they have gone to many and know what takes place at the event. It can lead to errors when an event outside the schema occurs, like someone doing homework during the party. When describing the event, the individual will forget about that event or invent a narrative as to why that event is occurring to better fit the set schema. It is usually not an issue because accuracy is not always needed in daily life. However, there are rare times when accuracy is needed (such as in the case of eyewitness testimony). This is when our schematic memory hinders us. Studies found that there is a high rate of false alarms for schematic consistent lures (false information consistent with the schema; Lampinen et al., 2001; Miller & Gazzaniga, 1998; Neuschatz et al., 2002; Pezdek et al., 1989; Webb & Dennis, 2020). Webb and Dennis (2020) found that participants struggled in their recognition of non-schema-consistent information unless the encoding resources were directed towards non-schematic information processing. They believe it is because participants must actively shift their attention to non-schema-consistent details at encoding. If we are not prompted to shift our attention to non-schematic information, we will not properly encode it.

Researchers have found that we make gist-based memory errors, a type of semantic memory error. This is when people only take the general theme of the information and are unable to encode or recall the specific details (Brainerd et al., 2012; Brainerd & Reyna, 2002; Koutstaal & Schacter, 1997; Reyna & Brainerd, 1995). Reyna et al. (2016) argued that witness reports that occur right after the crime rely on verbatim memory; however, in the courtroom the witness must rely on gist memory. Gist memories are when a person remembers the general meaning or purpose of the situation, but they do not focus on the specifics of the memory (Thompson, 2014). An example of a gist memory error would be that a witness may have been attacked with a screwdriver, but in court they just remember the weapon was a tool and mistakenly change it to a hammer. Verbatim memories are precise representations of the events (Thompson, 2014). An example of this would be that a few minutes after a crime, a witness can tell officers the shirt color of the perpetrator. When recalling a crime, using verbatim memory is found to decrease suggestibility when compared to gist memory (Pansky & Tenenboim, 2011).

Memories can also be influenced by repeated recall. As previously mentioned, the cases of Ronald Cotton and Howard Haupt are examples of memory changes due to repeated recall (Loftus & Ketcham, 1991; Innocence Project, 2023). Both witnesses only identified the perpetrator after repeated lineups and questioning from police. Studies have found that recognition memory does not improve across repeated tests (Challis & Roediger, 1993; Payne & Roediger, 1987). Tversky and Marsh (2000) found that repeatedly recalling information with a specific perspective in mind can alter our memories. They had participants read a story and write a biased letter to one of the characters. They were then asked to recall the original story. It was found that the perspective taken in the letter influenced the amount and type of details recalled. Participants' perspective also influenced the valence of the memory errors made. An example of this was when participants on the prosecution's side of the story remembered the positive or neutral memory items as negative, incriminating events.

Memories have also been found to be influenced by post-event misinformation. Individuals who witnessed the 2003 Swedish Foreign Minister's murder only reported 58% of the perpetrator's features correctly, at least in part because they discussed the case while they were waiting to be interviewed (Granhag et al., 2013). By discussing the case, witnesses unintentionally planted false information into fellow witnesses' minds. They began to edit their memories to include these newly discussed details, and they were unable to separate discussed details from the actual event. A classic example of post-event misinformation was obtained by Loftus and Zanni (1975). Participants viewed a car crash, and they were asked questions about the scene. By simply changing the way the question was asked, participants recalled things that were not there, like a broken headlight. Overall, our memories are malleable and small factors can alter or create false memories (Clark & Godfrey, 2009; Deffenbacher, 1991; Wells & Loftus, 2003).

Preventing Memory Errors

To help decrease the number of incorrect identification, the Department of Justice developed a guide to interview witnesses. This guide provides specific training for any individual that interacts with a witness (e.g., 911 operator, preliminary investigating officer, detectives). It generally trains them to ask open questions and avoid leading ones. It also recommends that officers separate any potential witnesses ("Eyewitness Evidence A Guide for Law Enforcement-Research Report," 1999). Psychologists have also suggested another method to correct memory errors: cognitive interviewing (Beatty & Willis, 2007; Fischer et al., 2000; Fisher & Geiselman, 1992). This technique uses structured questions, but allows the individual to elaborate on their answers (Beatty & Willis, 2007). An example of this method would start with the basic interview question of "what did the perpetrator look like?" The witness would give an answer and then the interview would probe the witness for more information. Specifically, interviewers will try to examine how a witness developed their answer, any difficulties answering the question, and the witnesses' perceived meaning of the question. For example, "how did you know he was a tall man?" The goal of the interview is to lead the participants to "memory codes [cues]" that have robust and relevant information. The technique starts with open-ended narration and then probes of the memory codes to get a detailed and accurate story of the event (Fisher & Geiselman, 1992). People interviewed with this technique can remember more about their physical activity from 35 years ago compared to those asked with normal interviewing techniques. They were also able to have more detailed responses (Fischer et al., 2000). Cognitive interviewing and modified cognitive interviewing (a streamlined and less demanding version of cognitive interviewing) has been found to increase correct recall, without increasing the amount of memory errors (Wright & Holliday, 2007).

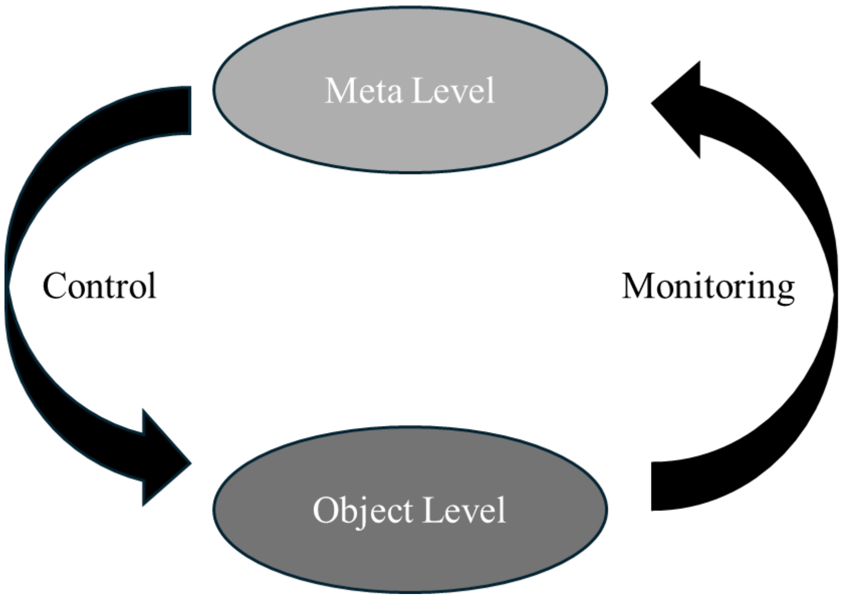
When specifically testing individuals' recognition memory, Wixted and Wells (2017) recommend pristine eyewitness procedures should be used. Procedures for pristine conditions include having one suspect per lineup, ensuring the suspect does not stand out, telling eyewitnesses that the suspect may not be in the lineup, double blind testing (i.e., testing in which both the person administering the line-up and the eyewitness does not know who the suspect is), and collecting confidence statements. The benefit of having one suspect per line up is to lessen the chance that the witness is just guessing and happening to choose one of the suspects. They argue that "a lineup that contains only suspects (no fillers) is like a multiple-choice test with no wrong answer" (Wixted & Wells, 2017, p. 11). It is also important to make sure the suspect does not stand out. If the suspect is the only one who fits the description, then the witness is going to pick them regardless of whether they are the offender. It is important that witnesses know that the perpetrator may not be in the lineup to avoid a forced choice. For example, a witness will choose the person who is closest to their memory of what the suspect looks like, even if they are not confident that person is the suspect. Double-blind testing also prevents suggestions from the individual administering the lineup. Wixted and Wells (2017) also specifically argue for the collection of the metacognitive judgment of confidence because it has been found to be an indicator of accuracy.

Examining Metacognition

With all these memory errors, our brains have created processes to evaluate our own memories. The idea of metacognition was first proposed by Tulving and Madigan (1970) when they called for methods to experimentally study memory. They called it, “knowledge of its own knowledge” (Tulving & Madigan, 1970, p. 477). The idea became popular and was labeled metacognition. Nelson and Narens (1990) created a simplified model of our metacognitive system (Figure 1). They broke it down into two levels (Meta and Object), and in between there is a flow of information (Control and Monitoring). The object level refers to our thoughts (e.g., answering questions on a practice exam). The meta level consists of the cognitions that we have about our thoughts (e.g., I am not confident these answers are correct) (Bares, 2011). Control modifies the object level. It can start, continue, or end basic thought processes (e.g., I need to continue studying for my exam because I need to know a lot of math formulas.). These control operations are influenced by the data collected from monitoring (Koriat & Helstrup, 2007) (e.g., I do not yet have the math formulas memorized.). Monitoring occurs when the meta-level is informed by the object level. This is the process that collects information and modifies the state of the meta-level based on the situation (e.g., “I believe I will pass my exam.”). Monitoring can best be measured using self-reports (Nelson & Narens, 1990). There are various types of monitoring processes, and they can occur at different points in acquisition such as before encoding “this is going to be difficult to learn,” during encoding “I am learning this quickly,” and after encoding “I will be able to remember this on my exam tomorrow.” Each of these processes have weak correlations with each other, indicating that they are based on different processes and kinds of input (Leonesio & Nelson, 1990).

Figure 1

Original Metacognitive Model From Nelson and Narens (1990)



Ease of Learning (EOL)

Ease of learning (EOL) is a judgment that occurs before acquisition, and it makes a prediction regarding our ability to learn materials in the future. We predict how easy or hard it would be to learn a specific item (Nelson & Narens, 1990). These judgements can inform our behaviors (Jemstedt et al., 2017; Nelson & Leonesio, 1988; Underwood, 1966). For example, if you think it would be easy to learn a new language, you will most likely allocate less time to studying it. An older study found that EOL judgments can accurately predict the rate of learning (Underwood, 1966). Jemstedt et al. (2017) found that EOLs can predict the ease of learning word pairs, but that this relationship is impacted by the variability of difficulty between the items in the "to be learned" list. Overall, they found that more variability led to higher accuracy and vice versa. Interestingly, Jemstedt et al. (2018) found that people's beliefs about what qualities make words easier to learn are a moderator for how processing fluency influences EOL judgments. Little research has been focused on EOLs (Dunlosky & Metcalfe, 2009). However, EOL judgments still have been found to be influenced by people's beliefs on their own knowledge.

Judgments of Learning (JOL)

Judgments of learning (JOL) is a monitoring process that occurs during or after acquisition. These judgments are predictions of how well a participant will do on future tests based on how they can currently recall the answers (Nelson & Narens, 1990). Research first found that students could accurately predict their performance on various recall tasks (Arbuckle & Cuddy, 1969; Groninger, 1979; King et al., 1980; Lovelace, 1984). However, these studies did not examine the strength of the relationship between JOL and accuracy. Further studies examining this question have found a moderately-sized correlation between accuracy and JOLs (Begg et al., 1989; Leonesio & Nelson, 1990; Vesonder & Voss, 1985). Leonesio and Nelson

(1990) found a .29 correlation between JOLs and recall accuracy when using a gamma correlation. This was a much larger relationship as compared to the EOL judgments that they examined. It was hypothesized that it was due to the timing of the judgment and overlearning of the items, making it easier to recall. Once the connection between JOLs and accuracy was established, researchers began to develop theories on how JOLs are developed.

There are two main contrasting theories for how JOLs are formed. The direct-access theory states that people can access a memory trace created during studying and they can make a JOL based on the strength of that trace (Cohen et al., 1991). Because the JOL is monitoring the strength of the trace, it should be consistent with later memory recall because the better the information is encoded, the better it is recalled (Koriat & Hestrup, 2007). The cue-utilization model states that learners cannot measure the strength of the memory directly, so they must use cues and beliefs to assess future recall (Begg et al., 1989; Benjamin & Bjork, 1996; Benjamin et al., 1998; Koriat, 1997). This leads people to consider perceived difficulty of study items, ease of recall during study, number of times studied, encoding strategies used, type of memory test, and beliefs in their own memory efficiency (Koriat & Hestrup, 2007).

The cue-utilization theory suggests that JOLs can be based on either one's experience during the memory process (e.g., "I learned this quickly") or on one's general understanding (or theory) for how their memory functions (e.g., "I tend to be good at remembering faces") (Koriat, 1997). Experience-based JOLs use factors like how easy it is to encode information and retrieval fluency as a basis of the JOL. Retrieval fluency is how easily the information can come to mind or be retrieved from our memory. They are using experiences to help inform a JOL. Studies have shown that the perceived ease of encoding can influence a JOL (Begg et al., 1989; Koriat, 1997; Matvey et al., 2001). Other studies have found that retrieval fluency can influence JOLs

(Benjamin & Bjork, 1996). All these studies support the argument that the experience of fluency can influence JOLs.

Theory-based judgments are the deliberate use of beliefs and theories on memory skills to make judgments on knowledge of memory performance (Dunlosky & Nelson, 1994; Koriat, 1997; Koriat & Bjork, 2006; Mazzoni & Kirsh, 2002). When examining our own knowledge, we may have the belief/theory that “I have poor a memory” (Dunning et al., 2003). This will then influence us to make a JOL based on that belief. When examining this theory, it is important to note that the accuracy of these JOLs depends on the truth behind the underlying belief (Koriat & Bjork, 2006). Overall, our experiences and *a priori* beliefs can both influence how we make JOLs.

Feelings of Knowing (FOK)

Feeling of knowing (FOK) judgments are when participants fail to recall the information but make a judgment on if they would correctly recognize the correct answer. FOKs are judgments about if the items could be recognized even if one cannot recall them in the moment, while JOLs are examining if one can recall items they currently remember for a future test (Nelson & Narens, 1990). Hart (1965) proposed the recall-judge-recognition paradigm (RJR method) to explain the process behind FOK judgments. When asked a FOK question, the participant attempts to recall the answer. They then judge if they could recognize the correct answer compared to the incorrectly recalled ones, and finally they try to recognize the correct answer. Contemporary arguments on the topic state that FOK is experience-based. The FOK occurs because people infer that the topic they are trying to recall is in their memory based on relevant facts, like familiar cues or target accessibility (Dunlosky & Metcalfe, 2009). The cue familiarity hypothesis states that we make judgments based on how familiar the cue seems to us

(Metcalfe et al., 1993; Reder, 1988). The accessibility model argues that people make judgments based on how fast or how much they can recall target information. This judgment ignores the accuracy of the information and just looks at the quantity recalled (Benjamin et al., 1998; Koriat, 1993; 1994; 1995; Stone & Storm, 2021). The purpose of FOK is to determine if you should spend time searching your memory for the information or if you should search an outside source. This also helps determine how long to search your memory. If you begin your search with a strong FOK, as the judgment decreases you may end the search (Dunlosky & Metcalfe, 2009). For example, your teacher asks if anyone knows signal detection theory. You have a high FOK because you have heard the term before, so you raise your hand. You are called on and begin your search for the definition. But after some searching your FOK lowers, and you terminate your search for the definition, telling your teacher that you can't remember. FOK judgements are related to a separate metacognitive monitoring process: confidence.

Confidence

Another type of monitoring judgment is confidence. Dunlosky and Metcalfe (2009) stated that confidence is a judgment on the likelihood that the response on a test is correct. Importantly, this occurs after retrieval of the answer because they are stating their confidence in a given answer (Nelson & Narens, 1990). Similar to previously mentioned monitoring theories, there are three distinct perspectives on confidence (direct-access, information-based, and experience-based; Koriat, 1997; Koriat & Levy-Sadot, 1999). The direct-access perspective states that confidence is based on our access to and strength of the memory trace. The information approach argues that we use an analytical process where information retrieved from memory is examined to develop an educated determination on the likelihood that the decision is correct. Our confidence on the test is based on the amount of evidence gathered that supports each answer

(Griffin & Tversky, 1992; Koriat et al., 1980; Koriat & Adiv, 2012; McKenzie, 1997). Experience-based theory argues that subjective confidence is inferential, but the cues are immediate and based on task performance instead of taking information from long term memory. An example of this would be a student answering a test question and making their judgment based on how easily they came up with the answer (perceived fluency). This perspective argues that we have certain beliefs about these mnemonic cues that influences our judgment (Koriat & Adiv, 2012). People have the belief that the faster we come up with an answer, the more we know it. I intend to focus on the information-based perspective by altering beliefs about recognition memory.

Initial confidence (i.e., confidence judgements made during the very first recognition attempt) has been found to be a strong indicator of accuracy (Wixted et al., 2018). It is important to point out that, like memory, confidence is malleable, so only initial confidence should be used to predict accuracy. Factors such as repeatedly recalling the event, talking with others or watching news stories make the relationship between confidence and accuracy weaken over time (Wixted et al., 2018; Wixted & Wells, 2017). Original confidence-accuracy research argued that there was a weak correlation; however, a meta-analysis found that the correlation could actually be reliably found and was higher for individuals making correct identifications (Sporer et al., 1995). The early finding of a weak relationship was due to an incorrect method of calculating the confidence-accuracy relationship. Previous research used a method called calibration; however, there were some issues with how the calculation was being applied. When selecting a suspect from a line up, calibration analysis calculates the "relationship between the subjective probability that an ID [identification] is correct (measured using a 100-point confidence scale) and the objective probability that it is correct" (Mickes, 2015, p. 93). The formula for calibration accuracy is $C = \frac{\#\, correct\ IDs}{\#\, correct\ IDs + \#\, incorrect\ IDs}$. This calculation is used for different levels of confidence. For example, to get calibration accuracy for the confidence levels of 90-100, researchers would take the correct and incorrect IDs in that confidence range. To examine the confidence-accuracy relationship researchers would compare the calibration accuracy with the confidence level. For example, a participant with a confidence level of 90-100% had a calibration accuracy of 90, indicating perfect calibration. Correct IDs are the suspect identifications made from the target-present lineups, while the operational definition of incorrect IDs changes depending on the studies. One way to calculate the incorrect IDs is by counting all errors with filler (distractor) identifications included. Another approach is to ignore all the filler identifications and only count the innocent suspect identifications. Filler identifications are known to be innocent line up members. Innocent suspect identifications (wrongful identifications) are when someone who looks like the actual perpetrator is incorrectly identified as committing the crime. Returning to the case of Ronald Cotton, Ronald would be the innocent suspect identification. He looked like the actual perpetrator but was not the man who convicted the crime. He was put in a lineup of about four other individuals. Those other individuals are fillers and if they were chosen, then it would be a filler identification.

Mickes (2015) believed this variability in the incorrect ID's led to a weak confidence-accuracy relationship. She suggested the new method of the Confidence-Accuracy Characteristic (CAC) that lead to a stronger confidence-accuracy relationship when reanalyzing previous studies. CAC specifically is the method of plotting "the relationship between confidence and accuracy for correct and incorrect suspect IDs made with varying degrees of confidence" (Mickes, 2015, p. 96). This method corrects the inconsistencies with incorrect ID's by specifically defining the incorrect ID as incorrect suspect identifications. CAC analysis allows for more flexibility when creating the confidence scale. Calibration analysis only allows for a 100-point scale, while CAC allows any type of confidence scale (Mickes, 2015). Since then, research on confidence has grown, with repeated studies using CAC analysis to show that there is a relationship between confidence and accuracy (Mickes 2015; Pezdek et al., 2021; Sauer et al., 2010; Sauer et al., 2008; Wixted & Wells, 2017).

Factors Influencing the Confidence-Accuracy Relationship

Emotionality of Memories. When it comes to emotional memories, the confidence-accuracy relationship is extremely weak (Neisser & Harsch, 1992; Schmidt, 2004; Schmolck et al., 2000; Talarico & Rubin, 2003). Kensinger (2007) hypothesized that negative emotions enhance our memory for "intrinsic" item details (features inherent to the emotional item or event), while it decreases the "extrinsic" contextual details. An example of an intrinsic item would be the color of the blood knife and a extrinsic detail would be the TV show playing in the background. However, when looking at overall emotional memory, our emotions enhance the vividness of the memory, making us more confident (Kensinger, 2009). This inflated confidence leads us to the assumption that memory is more accurate than it really is. Although all memories are susceptible to this confidence-accuracy disconnect, studies have argued that emotional memories are more susceptible to the disconnect (Dougal & Rotello, 2007; Schmolck et al., 2000; Windmann & Kutas, 2001). Kensinger (2009) argued that our emotions influence what people think they will remember about the event (i.e., our metamemory). A witness will remember some aspects of the event (intrinsic or extrinsic details) and those pieces of information will inform their metacognitive judgment. If this is an emotional memory, less information is needed to help them in their judgment. She also argues that negatively valanced information is associated with greater processing fluency that leads to the confidence-accuracy disconnect.

Unconscious Transference. Another factor that can reduce the confidence-accuracy relationship is unconscious transference. Unconscious transference is when “an eyewitness to a crime mistakes a familiar but innocent person from a police lineup” (Ross et al.,1994, p. 918). Individuals incorrectly identify the feeling of familiarity with confidence in their identification. This leads to a miscalibration of the confidence-accuracy relationship; therefore, high confidence does not mean greater accuracy (C. A. Carlson et al., 2023). There is a weak relationship between confidence and accuracy. It is important to identify the source of familiarity to determine if the confidence-accuracy relationship is reliable.

Estimator Variables. There has been a growing argument over estimator variables. These are variables that police can only approximate (for example: lighting, distance from the perpetrator, and passage of time (Moore et al., 2024; Semmler et al., 2018). There are two lines of argument for estimator variables. Some argue that a robust confidence-accuracy relationship remains even when these variables are suboptimal (Carlson et al., 2017; Semmler et al., 2018). Semmler et al. (2018) mentioned that with suboptimal estimator variables the discriminability is decreased, but the confidence-accuracy relationship remains strong with participants adjusting for those variables. Others argue that this relationship weakens (M. A. Carlson et al., 2023; Giacona et al., 2021; Lockamyeir et al., 2020; Winsor et al., 2021). When examining witnesses’ distances from a crime, there was a strong relationship at 3 meters and 10 meters. At 20 meters, the relationship is no longer seen (Lockamyeir et al., 2020). M. A. Carlson et al. (2023) examined how the estimator variable of witness’s sleepiness influenced the confidence-accuracy relationship. Participants with lower levels of sleep had a weaker relationship compared to those with high levels of sleep. Age of the witness was also seen to influence the confidence-accuracy relationship with 10 – 17-year-olds displaying the relationship, but four- to six-year-old did not. A weak relationship begins to arise at age seven (Winsor et al., 2021). Giacona et al. (2021) claimed that when there are multiple suboptimal estimator variables, the accuracy is reduced even when a witness is highly confident, and the conditions were otherwise pristine. This indicates that even with investigators best efforts, these estimators can still greatly influence the confidence-accuracy relationship.

What are Beliefs and How Do We Change Them?

A belief is the mental acceptance or conviction that some idea is true (Schwitzgebel, 2010). There are at least two aspects of a belief: representation content (topic of the belief) and assumed veracity (Stephens & Graham, 2004). Many of our beliefs are unconscious thoughts or outside our awareness. Our beliefs can be described as "enduring, unquestioned ontological representations of the world" and contain "primary convictions about events, causes, agency, and objects that subjects use and accept as veridical" (Connors & Halligan, 2015, p. 2). Beliefs can be seen as mental scaffolding to help judge the environment, explain new observations, and create a shared meaning of the world (Halligan, 2007). They allow us to assess our current experiences and then apply them to our past or future.

Fishbein and Ajzen (1975) broke down beliefs into three categories: descriptive, inferential, and informational beliefs. Descriptive beliefs are based on our senses and direct experience with an object. Generally, we trust our senses, so beliefs made based on them are held with more certainty. Inferential beliefs are those about unobservable characteristics or dispositions (ex. Anna is honest). These can be made based on previously learned relationships (schemas) or using logic rules (Bruner, 1957). An example of using logic rules to form a belief would be Sally is taller than Tom and Tom is taller than Anna, therefore Sally is taller than Anna (Fishbein & Ajzen, 1975). An inference had to be drawn based on prior knowledge to form this belief. The difference between these two types can be represented on a continuum. This is because many objects could appear to be direct observations, but they cannot be directly perceived. Instead, they are concepts defined by past experiences or teaching (e.g., The rat looks happy; however, we are just making an inference on its emotion.). Informational beliefs are when we do not use direct experience or inferences. Instead, we use an outside source to inform our decision (social media, friends, news). Information from outside sources is not always accepted; this depends on the individual's general perceived trustworthiness in the source (Fishbein & Ajzen, 1975). Fox News may release the headline, "Antifa led January 6th insurrection." A viewer must first believe that Fox News is a trustworthy source. If so, the belief about January 6th is formed.

Connors and Halligan (2015) argued that beliefs have four functions that tend to overlap with each other. First, they provide a consistent and coherent representation of the world. This representation informs behaviors and allows us to pursue goals and avoid threats. They inform higher order cognitive functions, like planning and decision making. Beliefs also create an explanatory framework to help interpret the world and process information. We tend to have a web of beliefs, and this web allows us to add new observations quickly into our old memories. This web will change as we observe new things, and it helps us better understand and adjust to the environment. The framework also allows us to configure and calibrate lower-level modular cognitive systems, like memory, language, and perception. Through top-down processing, beliefs will inform how we interpret sensory information. From a social perspective, beliefs function to form social relationships and understand other motivations. They can also provide people with a sense of community and safety when one is with individuals of the same beliefs.

Various researchers have examined the dimensions of a belief. Connors and Halligan (2015) have created 10 different characteristics of beliefs based on this research (Table 1). These characteristics acknowledge that there is variability in these characteristics depending on the person and type of beliefs. There is also a strength of how much we defend this belief and reject contrary evidence. By keeping these factors in mind, researchers can change people’s beliefs.

Table 1

Ten Characteristics of Beliefs

Characteristics		Literature
Different Origins	We can form a belief based on various sources. Ex. Direct experience or a trusted source.	Hughes and Sims (1997), Langdon (2013)
Vary in levels of evidence	Some beliefs need high levels of evidence to support them, while others do not need much evidence.	Lamont (2007)
"Held" at different awareness levels	Some are implicit, unconscious, or only inferred from behaviors, while others are conscious and even ruminated on.	Young et al. (2003)
Vary in generality and scope	They could refer to a specific object or person, or to a whole class of objects or people.	Freeman (2007)
Vary in degree of personal reference	A person can hold a belief about themselves, extend it to their in-group, apply it to out-groups, or all people equally.	Freeman (2007)
Held with different levels of conviction	Some beliefs can be stanchly held and defended, while others can be uncertain. This can change over time and across different contexts.	Bisiach et al. (1991), Connors and Coltheart (2011), Peters et al. (2004)
Vary in resistance to change	This looks at the response to contrary evidence and social pressure. People and a specific belief can vary in how open they are to disconfirming evidence.	
Vary in impact of cognition and behavior	There are some beliefs that people act on, while others people only verbally support the belief and do not act on it.	Bortolotti (2013)
Produce different emotional consequences	Some will be harmless or sometimes self-serving, while others can create distress.	Beck (1976)
Vary in degree they are shared	Some are common, while others are unusual.	David (1999)

Note. Characteristics are taken from Connors and Halligan (2014).

Belief Change

Belief change is the process of transforming a past belief to reflect the revised information of the topic (Trevors, 2023). Sharot et al. (2023) took a similar view stating that it is the "conscious or unconscious process of weighing the value of an old belief against the expected value of a potential new belief" (p. 3). Both definitions emphasize the comparison of the new information with a past belief.

Various key factors have been used to alter beliefs (Trevors, 2023). One theory of belief change is the Knowledge Revision Component framework (KreC framework). It examines the cognitive processes behind belief change (Kendeou et al., 2013, 2014, 2019). When looking at belief change, the passive activation principle should be taken into consideration. It states that memories can be passively reactivated by incoming information that cues the old belief to activate (Butterfuss & Kendeou, 2021). This theory states that old beliefs can be reactivated and hurt future learning/behaviors. To successfully change the memory, we must have old/incorrect information reactivated at the same time as the correct/new information. The co-activation is essential for integration of the new information by encoding the connections between the old and new information in long term memory. This will create a competition between the ideas, and as more information supporting the correct explanation is collected, the correct belief becomes the dominant one (Van Den Broek & Kendeou, 2008).

Another way people accept new beliefs is by making plausibility judgments. This is when people evaluate the potential truthfulness of the new information against their preexisting beliefs (Trevors, 2023). If a student does not think it is plausible that humans can stop climate change, they are less likely to update past incorrect beliefs or engage with new information about the topic. Therefore, if a new belief is viewed as more plausible, it is more likely to change a prior belief. Explanations for what makes a belief seem plausible include: the new belief contains similar qualities or ideas to their old belief, have more supporting evidence, and are delivered by credible sources (Lombardi et al., 2016). Cialdini (2001) further suggested that people are more susceptible to belief change when their credible source is an authority figure. Plausibility judgments closely relate to the validation process where information is being evaluated for plausibility specifically based on prior beliefs and knowledge (Cook & O'Brien, 2014; Richter & Maier, 2017). Validation is an "epistemic gatekeeper" that filters out inconsistent information, preventing it from being retained in memory (Richter, 2015). This process helps protect our beliefs because we are filtering out any opposing information that seems implausible when referencing our prior information (Trevors, 2023). This process must be taken into consideration when attempting to change a belief. We must reframe the new belief as more plausible compared to the previous belief.

Dual processing theory has been discussed as a possible explanation for belief formation and change. To break down the theory, processing is split into two systems. System 1 is an intuitive, unconscious, and automatic process that relies on heuristics. System 2 is the analytical, conscious, and controlled process that is slower and relies on systematic processing (Evans, 2008). According to this theory, incorrect beliefs are due to an overreliance on System 1 thinking, however we can alter the beliefs if we switch the thinking to System 2 (Chi, 2013; Trevors, 2023). This position is supported by misinformation research. When examining how misinformation is accepted, Pennycook and Rand (2019) found that people do not engage in critical thinking when they read false headlines. To prevent these false beliefs, researchers suggest people take critical thinking courses or engage system 2 thinking (Muis et al., 2018; Wilson, 2018). One study found that students reduced their beliefs in the paranormal and pseudoscientific beliefs when given a science and critical thinking course (Wilson, 2018). Therefore, influencing which system an individual uses could increase the chances of belief change.

Emotional valance has been shown to influence our belief acceptance. Emotional valance is the (positive/negative) psychological state that occurs in response to a relevant stimulus (Scherer, 2005; Trevors, 2023). Negative emotions have been linked to the duration of corrective messages and can lower learning gains (Nauroth et al., 2014; Trevors & Duffy, 2020). Sharot and Garrett (2016) argued that beliefs are more likely to be adjusted when the new information is positive compared to negative. They call this idea asymmetric updating. People have been shown to update their beliefs more when presented with good news, compared to bad (Sharot et al., 2011). Participants accepted the positively valanced information more and had a more positively biased belief (Chowdhury et al., 2014; Eil & Rao, 2011; Korn et al., 2012, 2014; Kuzmanovic et al., 2015; Möbius et al., 2022; Moutsiana et al., 2013; Sharot et al., 2011; Wiswall & Zafar, 2015). Sharot and Sunstein (2020) found that people are information avoidant when it induces negative affect. This would make people less open to that information when attempting to induce belief change.

Another process that influences belief change is motivation. In college students, researchers found that tasks with utility value (the task is useful to participants) led to increased belief change and engagement compared to tasks that had attainment value (the task supports the participants identity or schema) (Johnson & Sinatra, 2013). Another type of motivational variable is personal relevance; however, it can both help and hurt belief change. Interventions that refer to participants' identities can increase the personal relevance of the message. This leads to an increase in utility value of the task, increasing attention and effort on the task leading to better belief change (Dole & Sinatra, 1998; Eccles, 2009; Gregoire Gill et al., 2022). On the other side of the coin, studies have shown that higher levels of personal relevance can lead to the rejection of these attempted corrections to a belief (Nauroth et al., 2014; Trevors, 2022). Other cognitive and affective factors may be at play when using personal relevance. This theory does argue that personal relevance is a necessary part of changing a belief because it provides motivation to process correct content, but it is insufficient alone. The intervention success depends on the message and on how someone can cope with a challenge to their beliefs (Gregoire Gill, 2003; Gregoire Gill et al., 2022).

One type of belief changing intervention is that of refutation texts. This is a test that identifies an incorrect belief, refutes it, and presents evidence to reinforce the understanding of the new explanation (Hynd, 2001). These texts help promote knowledge revision by activating both the correct and incorrect knowledge. This leads to the cognitive conflict to help overwrite the incorrect belief (Kendeou et al., 2013, 2014, 2019). It is important to note that this text should not simply state that the belief is untrue, but it should provide compelling supporting explanations (Trevors, 2023). Lewandowsky et al. (2020) also recommended identifying the fallacies underlying the incorrect belief and only talk about the incorrect belief once to best refute the belief. Overall, meta-analyses have shown that refutation texts are a low cost and simple intervention that consistently and robustly leads to belief change (Schroeder & Kucera, 2022; Tippett, 2010). With all these belief changing methods in mind, I intend to induce positive and negative beliefs about general recognition memory.

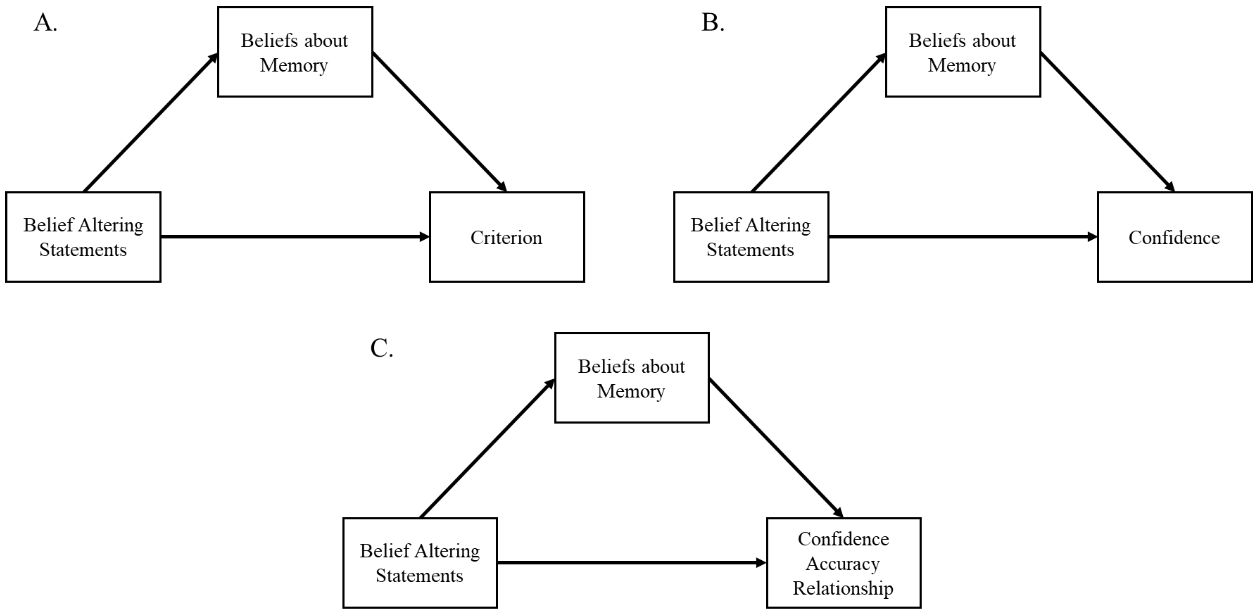
Present Study

There is a growing support for the position that beliefs can alter our metacognitive judgments (Benjamin et al., 1998; Koriat & Bjork, 2006; Mueller & Dunlosky, 2017; Rhodes & Castel, 2008). Although the link between JOLs and beliefs has been well documented, there has been considerably less research connecting confidence ratings and beliefs. The prior research on JOLs and beliefs cannot simply be generalized to confidence ratings without further research. A key difference between these two forms of judgment is the point in time at which they are made. JOLs are collected before the test is taken (e.g., Walking into the testing room a student says, "I am going to fail this test."), while confidence is assessed after the test is taken but before the answers are revealed (e.g., When talking to a friend after the test, a student says, "I think I passed that test.") (Nelson & Naren, 1990). Because confidence is taken after the test, there is additional information (their experience on the test) being factored in when making the judgment.

In this thesis, I examine if beliefs also influence confidence judgments, and thus, in turn, the confidence-accuracy relationship. Koriat and Adiv (2012) have begun to investigate this connection; however, their study examines social beliefs and confidence in a testing environment. They found that confidence can be influenced by how consistently the belief is supported. They specifically looked at the strength of the belief and the participant's levels of confidence. My study examined beliefs about memory instead of social beliefs. Also, I investigated the confidence-accuracy relationship seen in recognition tests instead of just confidence about social beliefs. It is important to investigate this question because there are recommendations to use confidence judgments to judge the accuracy of eyewitness memory. If beliefs influence this relationship, the judicial system should consider it when judging the accuracy of the memories. Participants were randomly assigned to two conditions that were designed to attempt to alter their beliefs about recognition memory: participants in one condition encountered information suggesting recognition memory is bad (Memory Horrible), and the participants in other condition encountered information suggesting recognition memory is good (Memory Great). My primary hypotheses are that altering one's belief about memory will alter their confidence in their own memory, but that it will not affect the accuracy of their memory. This will lead to a difference in the confidence-accuracy relationship. My hypothesis suggests that confidence should be higher in the Memory Great condition, while it should be lower in the Memory Horrible condition. My secondary hypothesis examines why this causal relationship is occurring by predicting that this relationship is mediated by the participants' beliefs about recognition memory (Figure 2). Specifically, I examined if the manipulation did indeed change beliefs about memory and if beliefs in memory mediated the relationship between the assigned condition and confidence, accuracy, and the confidence-accuracy relationship.

Figure 2

Predicted Mediation Paths



Method

Participants

Participants were recruited from Radford University through the SONA system in exchange for course extra credit (*n* = 451). Data were collected for one fall semester (about 4 months). Participants were all 18 years of age or older; anyone under that age was excluded from analysis. Participants were also excluded if they did not complete the whole test, took the test too fast/slow (under 13 minutes and over one hour), or they put the same answer for every question. Because deception was used, participants were excluded if they removed their permission to use their data. A total of 176 participants were excluded from the total. This led to a final sample of 275 participants. The participants' demographics were representative of the Radford Psychology Participation Pool (Table 2).

Table 2

Demographic Characteristics (n = 275)

Characteristics		*n*	%	*M*	*SD*
Gender	Female	211	76.7%		
	Male	58	21.1%		
	Non-Binary	5	1.8%		
	Prefer not to Answer	1	0.4%		
Age	18-41 Years Old			19.15	3.19
Year	Freshman	221	80.4%		
	Sophomore	23	8.4%		
	Junior	14	5.1%		
	Senior	17	6.2%		
Major	Nursing	107	38.9%		
	Psychology	39	14.2%		
	Criminal Justice	21	7.6%		
	Allied Health	15	5.5%		
	Exercise Science	7	2.5%		
	Undecided	7	2.5%		
	Biology	6	2.2%		
	Other	54	19.6%		

Note. Majors in the “Other” group contain four or fewer participants to a major.

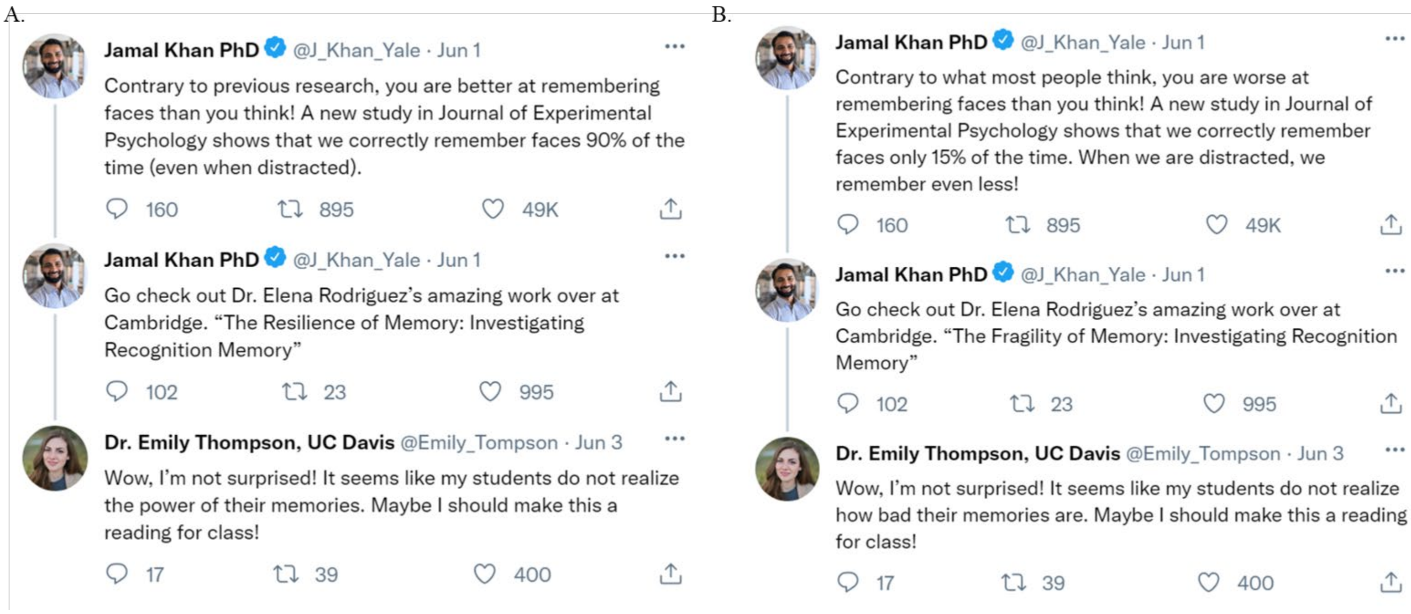
Materials

Belief Altering Statements

Six tweets were developed to change the beliefs of participants. Participants read a series of three tweet threads to attempt to change their beliefs about the reliability of the general public's memory. Three threads were developed for the "Memory is Horrible" condition, and three were created for the "Memory is Great" condition. Each of these conditions have the threads formatted the same way (usernames, likes, retweets, profile picture); however, the content is slightly different (Figure 3). Tweet threads were written by ChatGPT and then edited to include belief change methods. Tweet threads were used because they were shorter statements, like headlines, possibly preventing participants from engaging in analytical thinking and resulting in them accepting the proposed belief (Pennycook & Rand, 2019).

Figure 3

Belief Altering Tweets



Note. A. One example of Memory is Great Tweet threads. B. One example of Memory is Horrible Tweet threads.

To follow the suggestions of Cialdini (2001), these tweets were written by an authority figure in memory. Each of these tweets were written by fictitious experts in the field. This was done by adding the words “PhD” and “Dr.” to the username or handle. Morris et al. (2012) found that usernames have a large influence on how people judge the credibility of the content and author of the tweet. They found that usernames with topically relevant names are rated as more credible. They also thought that usernames that used non-standard grammar and abbreviations would be less credible. It was also found that profile pictures of people or a content relevant image also influenced credibility judgments compared to non-relevant default images. These tweets contained headshots or memory recognition-related images (image of a brain scan) to increase credibility. Tweets also contained a verified blue checkmark and mention the name of a prestigious college.

To increase the chances of belief change, the tweets followed the co-activation principle. The new information was activated concurrently with the old belief that it is refuting (Van Den Broek & Kendeou, 2008). This was done by having the tweet acknowledge the opposing view, while presenting the new idea.

Gregoire Gill (2003; 2022) argued that personal relevance is necessary when changing a belief, so one of the threads contained a reply applying the content of the tweet to something relevant to our participants. All participants were students, so there was a comment discussing how this knowledge could be applied to learning.

The participants viewed one thread that is flagged for misinformation and contains comments from other experts disproving the original tweet (Figure 4). This was a type of refutation text. Meta-analyses show that refutation texts can consistently and robustly promote belief change (Schroeder & Kucera, 2022; Tippett, 2010). Using all these methods, I predicted that the participants would change their beliefs in general recognition memory.

Figure 4

Example of Refutation Tweets (Memory Great)

Belief Change Questionnaire

Participants completed five questions measuring their beliefs about general recognition memory (Table 3). Portions of the questionnaire have been adapted from Magnussen et al. (2006) and Simons and Chabris (2011). The questions measuring general beliefs on memory were taken from this survey. In my study, the questionnaire had an unacceptable Cronbach's alpha ($\alpha = .45$). Questions were presented on a 1 (strongly disagree)-5 (strongly agree) Likert scale. Some questions from the original scales are formatted as yes/no questions, so they were adapted for a Likert scale. Higher scores on the questionnaire indicate that participants believe human memory is generally accurate.

Table 3

Belief Change Questionnaire

Rate the following statements on a scale from 1(strongly disagree)- 5(strongly agree).
1. There is a limit to the amount of information we can store in our brains. [R]
2. Once someone has seen a person and formed a memory about them, that memory does not change.
3. Human memory works like a video camera, accurately recording the people we see and hear so that we can review and remember them later.
4. In my opinion, the testimony of one certain eyewitness should be enough evidence to convict a defendant of a crime.
5. I cannot trust my memory when trying to recognize faces. [R]

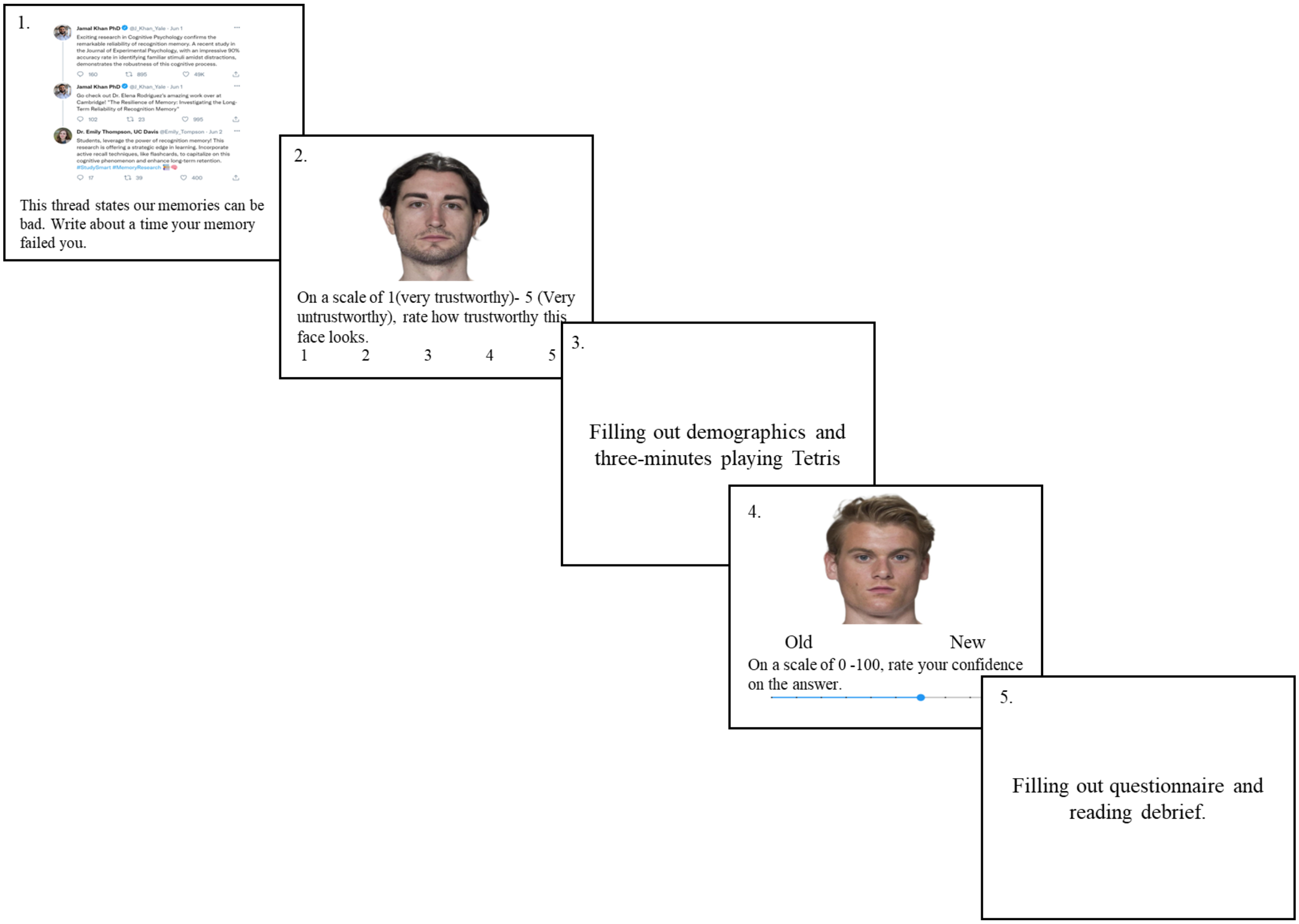
Note. R = reverse coded items

Faces for the Old/New Recognition Task

A series of 48 faces from the Chicago Face Database and the Face Research Labs London Set were shown to participants (DeBruine & Jones, 2021; Ma et al., 2015). These are free databases that are commonly used for facial recognition research (Heinbockel et al., 2024; Jones et al., 2021; Marini et al., 2021; Wixted & Wells, 2017). Photoshop was used to crop out the t-shirt color and to standardize the background color. Each face is of a white male with a neutral facial expression. The image was a frontal shot of the face (Figure 5). A collection of various races would make the faces easier to identify as old/new, so images of white males were used to prevent ceiling effects with recognition.

Figure 5

Steps of the Procedure



Note. 1. Reading belief altering tweet threads. 2. Encoding Phase. 3. Gap to complete demographics and play three minutes of Tetris. 4. Recognition phase and confidence question. 5. Complete the belief change questionnaire and debrief the participants.

Procedure

Participants completed a Qualtrics survey that started with an informed consent document. Participants were randomly assigned negative or positive belief statements. Participants read the tweet threads and answered three questions for each thread. The three questions were attention checks quizzing the participants on the content of the tweets. These questions were not relevant for data collection and instead made sure participants were thinking about the threads that they read. Each thread was presented one at a time to prevent cognitive overload. They read three tweet threads, which were designed to alter their beliefs in recognition memory. Next, they completed a free response question, "These threads state that our memories can be bad/good. Write about a time your memory failed/helped you." These questions were not analyzed. This made sure participants are applying the newly altered belief to their personal memories.

After the tweets, participants read the instructions for the encoding phase of the old/new recognition task. During the instruction phase, participants were told that they will view 24 faces and rate the perceived trustworthiness of each face. The trustworthy rating was not relevant to data collection and instead was used to make sure the participants were examining each face. The encoding phase began, and each face was presented one at a time with the Likert scale question on trustworthiness below it. Participants looked at each face and took as long as they wanted to answer the questions. This test was taken on Qualtrics in their preferred environment.

After this encoding phase, there was a break to report their demographics. Participants competed questions about their age, major, year in college, and gender. To prevent ceiling effects, the participants played Tetris in the survey for three minutes before they were able to move on.

The next part of the old/new recognition test was the recognition phase. Participants were instructed to view the faces and determine if they had seen the face before (old) or had never seen it (new). The participants were also asked to rank their confidence on a slider of 0% (completely guessing) to 100% (absolutely sure I'm correct) for each face. The phase began, and the participants viewed 48 faces (24 old and 24 new) and answered the old/new and confidence questions for each face. These questions were presented below the face. Faces were presented in random order for each participant, and all were equally presented as an old or new face. Like the encoding phase, the task was self-paced. Due to the use of deception during the belief altering statements, participants were informed that all the statements were fictitious and were provided with possible resources on the real research behind memory. They also received the option to withdraw their responses from the study.

Results

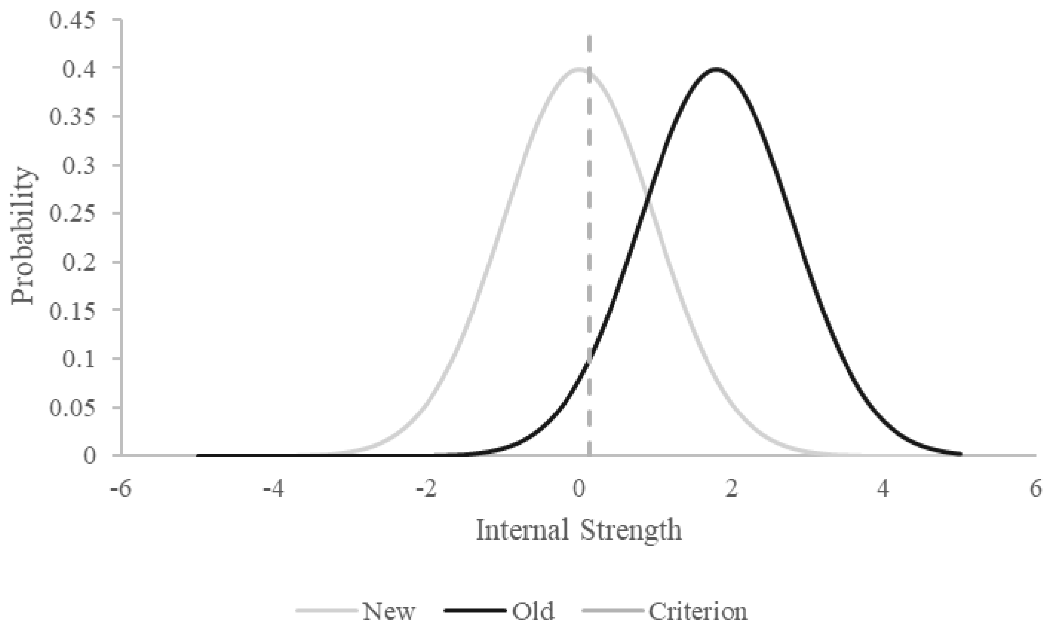
Recognition Accuracy

To determine recognition accuracy, R 4.4.2 and Excel were used to create a signal detection-based model. A simple independent observation model was run containing two Gaussian distributions (noise and signal + noise). Signal detection assumes that the curves are both normally distributed. A second assumption is that noise is always present. There are two possible assumptions of variance for this statistic, equal variance, and unequal variance (or not assuming the variances are equal, but they might be). Equal variance states that the variance of the signal distribution is equal to the variance of the noise distribution. Unequal variance states that the distributions can have different or equal distributions. For this research, unequal variance was used because it includes the possibility of both types of distributions.

After running the simple independent observation model, it was determined that participants can significantly discriminate between the new (noise) and old (signal) stimuli in both the Memory Great (χ^2 (1) = 2114.87, $p < .001$) and Memory Horrible (χ^2 (1) = 2180.42, $p < .001$) conditions (Figure 6). Criterion values were calculated by first finding a criterion value for each participant, then an independent samples t-test was run between the two conditions. Participants do not shift their criterion depending on the memory conditions as there was no significant difference between the Memory Great (M = .13, SD = .34) and the Memory Horrible (M = .13, SD = .37) criterion, $t(273) = -0.10$, $p = .924$, $d = -0.00$. The same process of aggregating the participants' trials was used to calculate the d' for each group. As hypothesized, both conditions can differentiate between the old and new responses the same amount because there was no significant difference between the Memory Great (M = 1.80, SD = 0.86) and the Memory Horrible (M = 1.80, SD = 0.80) d's, $t(273) = -0.01$, $p = .990$, $d = -0.00$.

Figure 6

Signal Detection Model for Both Conditions

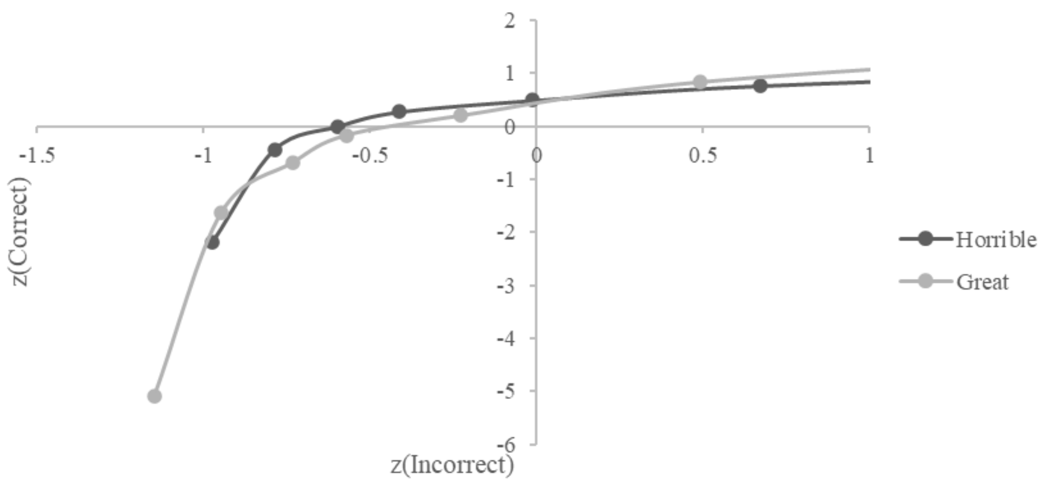


Note. The d’ and criterions for both conditions were the same, so these Gaussian distributions are a visual representation of both conditions (Memory Great and Memory Horrible).

To account for the possibility of unequal variance, linearized ROC curves (Figure 7) were used to check assumptions (Barch, n.d.). The curve was calculated by taking the cumulative frequencies of the false alarms and hits for each condition. A total of six points calculated based on three levels of confidence (low: 50%, moderate: 51-75%, high: 76-100%). These points were converted into z scores and graphed using Excel. The process of converting the points into z scores turned the normal ROC curve into a linearized one. The slope of each condition is roughly one, so there is an equal variance between the signal and the noise distributions. The ends of each condition do curve to extremes specifically for the false alarms rate indicating a possibility of a ceiling effect. Based on these findings, there is equal variance between the noise and signal distributions, so a normal ROC curve, area under the curve (AUC) and partial area under the curve (pAUC) could be obtained.

Figure 7

Linearized ROC Curve for Both Conditions

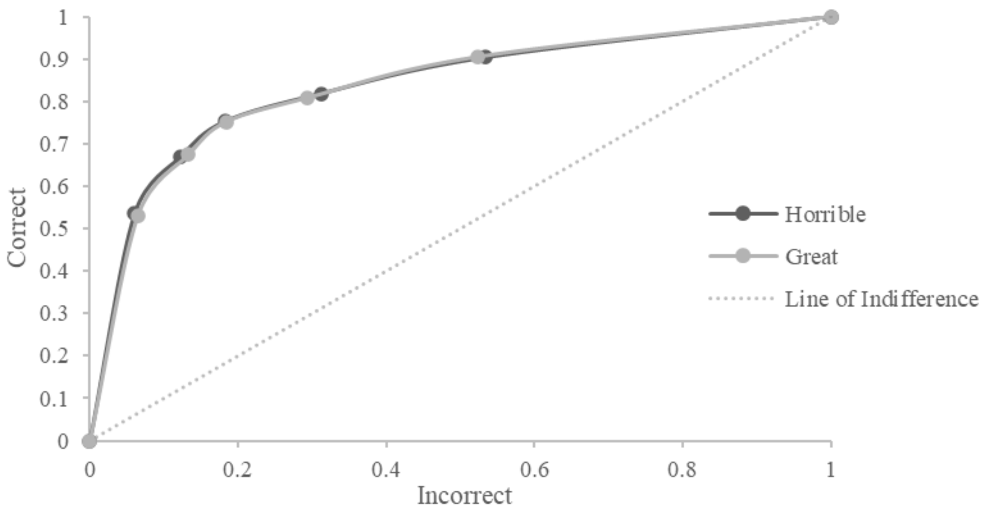


Note. The x axis is the z score of the proportion of correct answer on the old/new recognition test. The y axis is the z score of the proportion of incorrect answers.

A normal ROC curve was run obtained by taking the cumulative frequencies of correct and incorrect rates for each condition (Figure 8). Like the linearized ROC curve, they were broken down by confidence level. There was no difference in discrimination between the two conditions as seen by the overlapping ROC curves. Both conditions did have good discrimination between hits and false alarms with the ROC curve bending towards the top left corner. The argument for good discrimination can also be supported by the AUC being 0.83 (95% CI: .82, .84) for the Memory Horrible condition and .83 (95% CI: .82, .84) for the Great condition. Perfect discrimination would have an AUC of 100. PAUC was also calculated between each point to see if there is a difference in discrimination between each confidence level (Table 4). For each confidence level, participants were able to significantly discriminate between hits and false alarms.

Figure 8

ROC Curves for Both Conditions



Note. The graph contains an indifference line (dotted line) that represents the false alarm and hit ratio when a participant is just guessing. Each point on the curve is a decision threshold and the correct and incorrect rate is plotted for each threshold. The curve bows towards the top left corner, indicating good discriminability. For each decision criteria, the proportion of correct is more than the incorrect proportion.

Table 4

Partial Areas Under the Curve

Horrible			
Range	pAUC	Lower CI	Upper CI
0.00 - 0.06	0.02	0.01	0.02
0.06 - 0.12	0.04	0.02	0.05
0.12 - 0.18	0.04	0.02	0.06
0.18 - 0.31	0.10	0.07	0.13
0.31 - 0.53	0.19	0.16	0.22
0.53 – 1.00	0.44	0.41	0.48
Great			
Range	pAUC	Lower CI	Upper CI
0.00 - 0.06	0.02	0.01	0.02
0.06 - 0.13	0.04	0.03	0.05
0.13 - 0.18	0.04	0.02	0.05
0.18 - 0.29	0.09	0.06	0.11
0.29 - 0.52	0.20	0.17	0.23
0.52 – 1.00	0.45	0.42	0.49

Note. pAUC = partial area under the curve. 95% Confidence intervals (CI) for pAUC's were calculated. Range = distance between the two points on the ROC curve.

Confidence

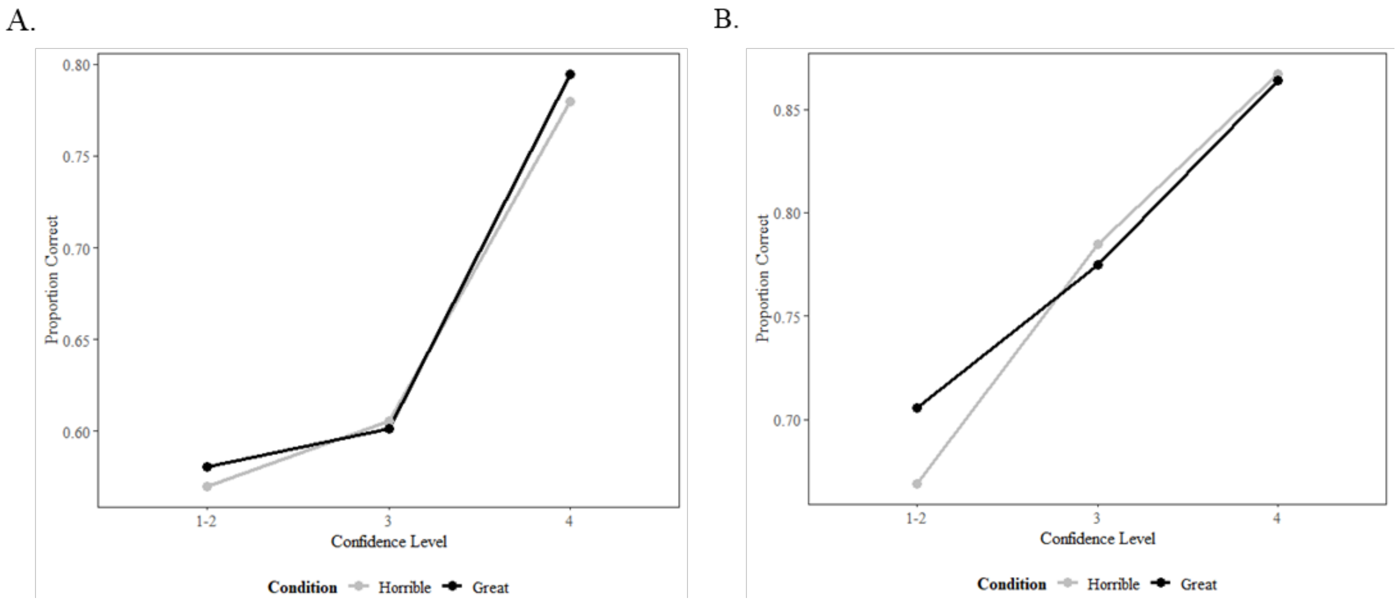
To compare confidence levels for the two belief altering statement groups, an average confidence score was created for each participant. The assumption of univariate normality of confidence was examined using the Kolmogorov-Smirnov goodness-of-fit test and graphing confidence on a histogram. A Kolmogorov-Smirnov goodness-of-fit test was used due to the large sample size. Both tests rejected this assumption, so confidence was not normally distributed, $D(275) = .06$, $p = .019$. With this violated assumption in mind, a nonparametric independent samples Mann-Whitney test was run to examine the difference between the confidence levels of the two conditions. All the assumptions for this test were met. There was no significant difference between confidence levels for the Memory Great (*Mdn* = 77.61, $n = 136$) and Memory Horrible (*Mdn* = 76.27, $n = 139$) conditions, $U = 10077.00$, $p = .343$. The results did not confirm my hypothesis that there would be a significant difference between the confidence levels with the Memory Great group having higher levels of confidence.

Confidence Accuracy Characteristic (CAC) Plots

Using R 4.4.2 and the ggplot2 package, the confidence-accuracy relationship was plotted. Confidence results were originally categorized into four separate bins (0-25%, 26-50%, 51-75%, 76-100%). However, like previous studies, the lowest two confidence bins were combined due to the small n's (Tekin et al., 2021) (0-50%, $n = 244$; 51-75%, $n = 266$; 76-100%, $n = 273$). Two graphs were created using the data from the old-new recognition test (old response and new response). Accuracy for the old responses were calculated using $\frac{hits}{hits + false\ alarms}$ for each confidence bin. Accuracy for new responses were calculated using $\frac{correct\ rejections}{correct\ rejections + misses}$ for each confidence bin. For each graph, two lines were constructed representing each belief condition (Figure 9).

Figure 9

CAC Plots



Note. Panel "A." This graph shows the confidence-accuracy relationship for the old responses. Proportion correct is lower overall for old responses compared to the new responses. There is also little difference in the proportion correct between confidence bin low and moderate. Panel "B." This graph shows the confidence-accuracy relationship for the new responses. For both conditions, as confidence increases, the proportion correct increases as well.

To determine if the manipulation (belief conditions) significantly changed the confidence-accuracy relationship between the groups, two 3 (confidence bins) x 2 (belief conditions) Mixed Model ANOVAs were run with the old and new responses. For the ANOVAs, the participants' trials were sorted by confidence bin and old/new responses with a mean accuracy calculated from each participants' confidence bin for old/new responses. So, each participant had six mean accuracy data points (three confidence bins x two old/new responses).

For the old responses, participants were more accurate when they were more confident, as shown by a significant main effect for confidence levels, $F(1.89, 383.44) = 45.28$, $p < .001$, $\eta^2_{partial} = .18$. A Tukey post hoc test indicated that participants were significantly more accurate when they have high (76-100%) confidence ($M = .79$, $SD = .21$, $n = 273$) compared to when they have low confidence (0-50%, $M = .57$, $SD = .32$, $n = 244$, $p < .001$) and moderate confidence (51-75%) confidence ($M = .63$, $SD = .29$, $n = 266$, $p < .001$). Accuracy was also significantly higher when the participants had moderate confidence compared to low confidence ($p = .02$). The accuracy between the Horrible ($M = 74.67$, $SD = 13.40$, $n = 139$) and Great ($M = 75.75$, $SD = 15.68$, $n = 136$) conditions was not different, as seen by the lack of significant main effect for belief conditions, $F(1, 203) = 2.13$, $p = .146$, $\eta^2_{partial} = .01$. For old responses, the belief conditions did not have any effect on the confidence-accuracy relationship as there was no significant interaction between belief condition and confidence level for the old responses, $F(1.89, 383.44) = .84$, $p = .429$, $\eta^2_{partial} = .00$.

The results are replicated with the new responses with significant main effect for confidence levels, $F(1.77, 373.44) = 29.92$, $p < .001$, $\eta^2_{partial} = .12$. A Tukey post hoc test indicated that participants were significantly more accurate when they have high (76-100%) confidence ($M = .86$, $SD = .18$, $n = 273$) compared to when they have low (0-50%, $M = .68$, SD

= .30, *n* = 244, $p < .001$) and the moderate (51-75%) confidence (*M* = .77, *SD* = .25, *n* = 266, $p <$ .001). Accuracy was also significantly higher when the participants had moderate confidence compared to low confidence ($p = .002$). Accuracy was not different between the Horrible (*M* = 74.67, *SD* = 13.40, *n* = 139) and Great (*M* = 75.75, *SD* = 15.68, *n* = 136) conditions with no significant main effect for belief conditions, $F(1, 211) = 0.90$, p = .343, $\eta^2_{partial} = 0.00$. For new responses, the confidence-accuracy relationship was the same across the two belief conditions as shown by no significant interaction between belief condition and confidence level for the old responses, $F(1.77, 373.44) = 0.71$, $p = .475$, $\eta^2_{partial} = 0.00$.

Beliefs About Memory

To compare beliefs about memory for the two belief altering statement conditions, a composite mean was calculated from the questionnaire. Questions one and five were reverse scored. The assumption of univariate normality was examined using the Kolmogorov-Smirnov goodness-of-fit test and graphing beliefs on a histogram. Beliefs about memory were not evenly distributed, $D(275) = .11$, $p < .001$. Due to the violated assumption, a nonparametric independent samples Mann-Whitney test was used to compare beliefs about memory for the two conditions. The assumptions for this test were met and there was a significant difference between the beliefs about memory for the Memory Great (*Mdn* = 2.8, *n* = 136) and the Memory Horrible (*Mdn* = 2.4, *n* = 139) conditions, $U = 11471.00$, $p = .002$. This indicates that my manipulation was successful in changing participants' beliefs.

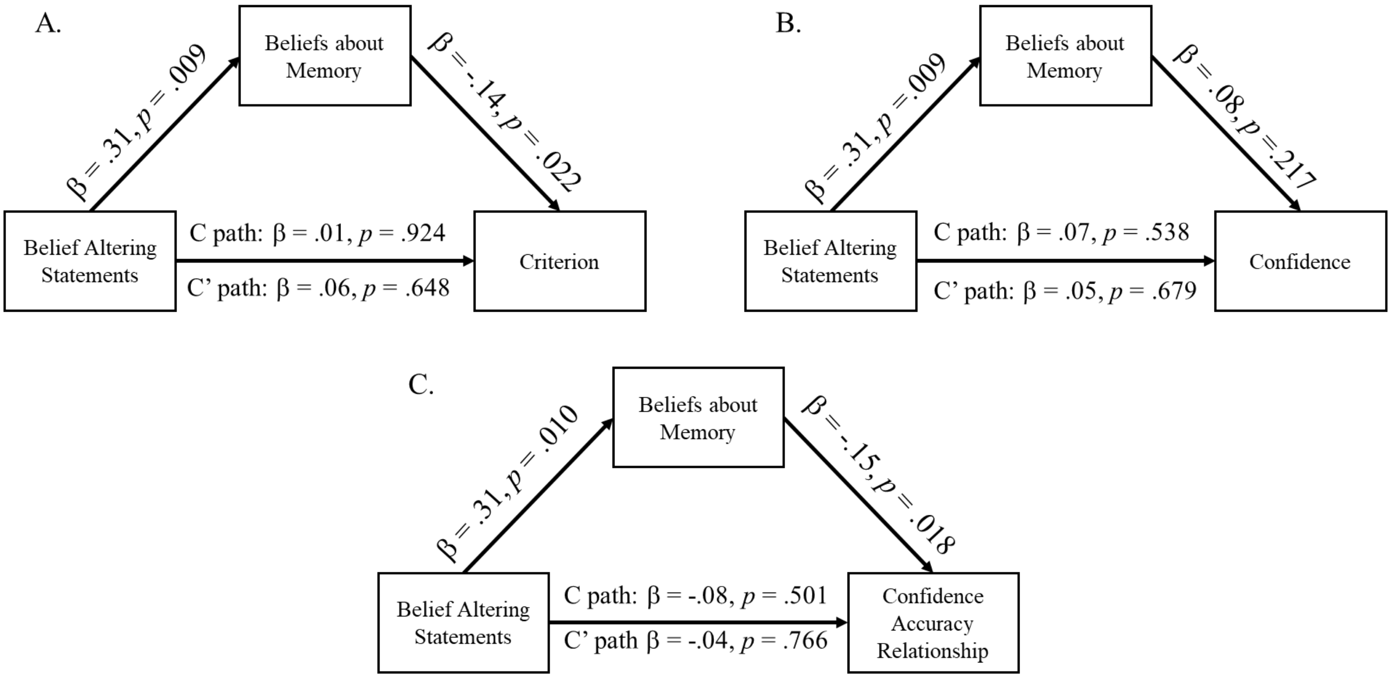
Mediation Model

The series of Hayes' PROCESS macro mediation analyses (Hayes, 2022) using model 4 with 5,000 bootstrapped resamples were conducted to see if beliefs about memories mediate the relationship between belief altering statements and the three dependent variables (criterion, confidence, and confidence-accuracy relationship: Figure 9). To determine mediation, three simple linear regressions were conducted.

To determine mediation, four conditions must be met. First, the belief conditions must be significantly related to the three dependent variables (c-path). Second, belief conditions must be significantly related to beliefs about memory (a-path). Third, beliefs about memory should be a significant predictor of the three dependent variables (b-path). Lastly, belief conditions should no longer be a significant predictor of the three dependent variables (c') when the mediator beliefs about memory is included as a second predictor in the same regression model. If only the first three conditions are met and the relationship's strength is reduced, then partial mediation can be determined and assessed.

Figure 10

Mediation Models for Each Dependent Variable



Note. C. $N = 272$. Correlations for three participants could not be determined due to a lack of variance between accuracy and confidence.

Model A examined the relationship between belief altering statements and the criterion with beliefs about memory mediating the relationship. In Figure 10a, there was no significant relationship between the belief altering statement condition and criterion, as shown by a non-significant c-path, $\beta = .01$, $p = .924$. There was a significant relationship between the belief altering statements and beliefs about memory (a-path; $\beta = .31$, $p = .009$), and a significant relationship between beliefs about memory and criterion (b-path; $\beta = -.14$, $p = .022$), indicating an indirect relation. The indirect (a*b) path was significant ($\beta = -.04$, 95% CI: −.11, − .00). The relationship between belief altering statements and criterion was indirectly predicted through beliefs about memory.

In Model B, the relationship between belief altering statements and confidence was examined with beliefs about memory as a mediator. Figure 10b demonstrates that there was a non-significant relationship between belief altering statements and confidence (c- path), $\beta = .07$, $p = .538$. As previously seen, there was a significant relationship between belief altering statements and beliefs about memory (a-path; $\beta = .31$, $p = .009$); however, there was no significant relationship between beliefs about memory and confidence (b-path; $\beta = .08$, $p = .217$). There was also a non-significant indirect (a*b) path ($\beta = .02$, 95% CI: −.02, .08).

In Model C, the relationship between the belief altering statements and the confidence-accuracy relationship was examined with beliefs about memory being the mediator. The confidence-accuracy relationship was calculated by getting a correlation between confidence and accuracy for each participant. Figure 10c shows that there is no significant relationship between belief altering statements and the confidence-accuracy relationship, as seen by the nonsignificant c-path, $\beta = -.08$, $p = .501$. There was a significant relationship between belief altering statements and beliefs about memory, as seen by a significant a path, $\beta = .31$, $p = .010$. There was also a significant relationship between beliefs about memory and the confidence accuracy relationship (b-path; $\beta = -.15$, $p = .018$). There is a weak indirect-only mediation with belief altering statements influencing the confidence-accuracy relationship, only through the mediator of beliefs about memory as seen by the significant indirect effect ($\beta = -.05$, 95% CI: $-.10$, $-.00$).

Discussion

This study aimed to examine if altering an individual's beliefs about memory could influence a person's confidence and create a confidence accuracy disconnect. It also wanted to further examine why this relationship would occur through a mediation analysis. Overall, the manipulation was not successful in directly changing participants' confidence-accuracy relationship through belief statements. Confidence levels did not differ between the two conditions. Participants in the Memory Great condition did not have higher levels of confidence than those in the Memory Horrible condition. As predicted, accuracy was the same between the two groups, but because confidence was also equal, there was no confidence-accuracy disconnect between the two conditions. Therefore, my primary hypothesis must be rejected. Instead of finding a weak relationship between confidence and accuracy, there was a similar confidence-accuracy relationship across both conditions. Confidence was able to predict the participants' accuracy.

When further investigating this relationship through mediation analysis, there was also no direct relationship between the belief altering statements and dependent variables of the criterion, confidence, and confidence-accuracy relationship. Instead, there is a very weak indirect-only relationship with belief altering statements influencing the confidence-accuracy relationship and criterion, only through the mediator of beliefs about memory. When looking at individual pathways, the belief altering statements were related to the beliefs about memory, however; beliefs about memory were not significantly related to confidence. The original hypothesis acted under the assumption that a direct relationship would arise and wanted to investigate why this relationship was occurring. Because there are weak indirect-only relationships for the two of the models and not a direct one, my secondary hypothesis is not supported.

Implications

Stretch and Wixted (1998b) claim that individuals adjust their confidence criteria using a likelihood-ratio decision model explaining why the disconnect in the confidence-accuracy relationship was not found. The likelihood ratio decision model is the theory that we take a ratio of the probability of old faces over new faces to help us make an old/new decision. If the ratio is within a certain threshold, the participant will say that the face is old and vice versa for new items. When the likelihood ratio is low, participants will adjust the boundaries of these thresholds and require more evidence to make a high confidence decision that an item is old (Stretch & Wixted, 1998a). Participants could be adjusting their threshold boundaries for each confidence level based on their accuracy leading to the presence of a relationship instead of a weak confidence-accuracy relationship.

Pezdek et al. (2021) showed that participants can be sensitive to changes in the accuracy of their memory and therefore confidence can be a good predictor of accuracy. Similarly, in this study, participants also appeared to be metacognitively aware of when they were more or less accurate, as shown by a robust confidence-accuracy relationship in both conditions. Like Pezdek et al. (2021), participants were aware of the instances that they would be less accurate and took that into account when making a confident judgment. Overall, having a confidence-accuracy relationship across the conditions aligns with previous findings (Lockamyeir et al., 2020; Mickes, 2015; Pezdek et al., 2021; Sauer et al., 2010; Sauer et al., 2008; Sporer et al., 1995; Wixted & Wells, 2017).

The belief statements did not directly influence the dependent variables (criterion, confidence, confidence-accuracy). In all mediation models, there was a relationship between belief changing statements and beliefs about memory indicating manipulation successfully changing beliefs. The belief change methods used in the tweet threads were effective in influencing beliefs about memories (Cialdini, 2001; Gregoire Gill, 2003; Gregoire Gill et al., 2022; Morris et al., 2012; Pennycook & Rand, 2019; Schroeder & Kucera, 2022; Tippett, 2010; Van Den Broek & Kendeou, 2008). There was no significant relationship between belief changing statements and confidence. Also, the indirect (a*b) path was weak and not significant. A possible explanation for the weak relationship is that there could be an additional variable influencing confidence. Koriat (2012) mentioned that confidence judgments primarily rely on mnemonic cues from task performance instead of applying declarative knowledge and beliefs. Participants may have gathered enough evidence from the initial performance-based cues, that they did feel the need relay on their beliefs to the judgments.

The weak indirect-only relationships for the dependent variables of the criterion and confidence-accuracy could be explained by the direction of the relationships in the a and b paths. The relationship between the belief changing statements and beliefs about memory is positive, while the relationship between the beliefs about memory and the two dependent variables is negative. As beliefs about memories increase, the criterion decreases leading to a more liberal criterion. For the confidence-accuracy relationship, as beliefs about memory increase, confidence-accuracy decreases. These a path and b path relationships create a significant negative ab indirect relation and it leads to a weak direct relationship between the belief changing statements and the dependent variables. The belief-changing statement influences on the criterion and confidence-accuracy relationship operates only through beliefs about memories.

An individual's criterion can be influenced by simply instructing participants to be more conservative or liberal when making decisions. Also, if you offer an incentive, a participant will change their criterion (Estes & Maddox, 1995; Healy & Kubovy, 1978). If participants are changing their decision-making criterion based on instructions or payoffs that they may not believe in, then it makes sense that the criterion will change when they internalize a belief. This could explain the relationship between beliefs about memory and the criterion. The indirect relationship for the confidence-accuracy variable indicates that as participants believe humans memory is reliable, the weaker the confidence-accuracy relationship. This indicates that the participants are possibly becoming overconfident when making accuracy judgments. This supports previously proposed theories (Benjamin et al., 1998; Koriat & Bjork, 2006; Mueller & Dunlosky, 2017; Rhodes & Castel, 2008); however, further investigation is needed to see why the relationship between beliefs and confidence-accuracy was significant while confidence alone was not a significant relationship.

In both conditions, participants were able to successfully distinguish between what faces were old and what were new as seen by the significant chi squared tests. These results help confirm that the old/new recognition task was a successful method of testing facial recognition. These results were expected as the task has been used many times for facial recognition studies (Chance & Goldstein, 1979; Pezdek et al., 2020, 2021; Searcy et al., 1999). This experiment's old/new recognition test was specifically designed using experiments by Pezdek et al. (2021, 2020). Their participants were also able to discriminate between old and new faces. Although between their two studies, they overall had less discrimination with the maximum d' at .95, compared to this study's d' of 1.80. The high d' and the curved tail on the linearized ROC curves indicate that a ceiling effect is occurring. More time between the exposure and recognition portions of the memory tasks is needed because the participants were too easily discriminating between the old and new responses.

Limitations/Future Directions

A possible limitation of this study would be the type of population being used. Because SONA was used, the participants were all students taking an Introduction to Psychology course. Part of the curriculum is to teach students about memory and possible errors. Thinking back to methods used to alter beliefs because participants previously encountered information about memory; the participants in this study may have needed higher levels of evidence to support the change compared to other beliefs (Connors & Halligan, 2015; Lamont, 2007). It seems this manipulation does reach the proper level of evidence to change the belief, however using a psychology student population could lead to an inaccurate measurement of the manipulation's effectiveness. It brings into question how much did the manipulation influence beliefs and how much was really the class instructor. Specifically, participants in the Memory Great group could have higher levels of belief change if they did not encounter the class content about memory errors. In the Memory Horrible condition, the strength of the manipulation could be over inflated because they learned about memory errors. Also, the class instructor and the fabricated experts are both credible authority figures, so the manipulation could have an uphill battle to convince participants that the tweets are more credible (Cialdini, 2001). Ideally, future studies should not use Introduction to Psychology students to avoid all these issues. If that is not possible, the data collection should stop before the content on memory is taught in the class. These safeguards will lead to a more accurate measurement of the belief altering statements influence on beliefs about memory.

The a path indicates that the manipulation is changing beliefs; however, an additional study should more directly test the manipulation. Further studies should give a pre and post belief survey with the belief altering statements in between. Along the same lines, participants may alter their beliefs about others' memory with the tweets but are not applying that belief to themselves. Writing about a time that their memory was great or failed may not be enough for them to apply the belief to themselves. Koriat and Adiv (2012) argued that a belief will not be endorsed unless it passes a certain confidence threshold, depending on the belief being accepted. It is possible that the tweets were not able to surpass the participants' threshold for personal endorsement. Future studies could increase the strength of belief manipulation. Instead of just reading tweets and writing about their memories, participants could be given a task that would mislead them to believe that they failed or aced a facial recognition test. This would help them apply the belief to themselves and could better influence their confidence judgments on future old/new recognition tests.

It is also important to acknowledge that there is a chance of expectancy bias. Because the tweets are trying to change the participants' beliefs and memories, they could create an expectancy effect just by reading them. Also having a belief change survey could inform the participants of the main goal of the study and it could have altered the old/new recognition test. To help prevent this issue, the belief change questionnaire was placed at the end of the study. There is a possibility that placing the questionnaire at the end is not enough to remove this bias, and the characteristics the tweets should be further examined for bias.

With many court convictions still heavily relying on eyewitness memory, it is important to examine this confidence-accuracy relationship. There is a growing consensus that confidence judgments could be used to help measure the accuracy of witness statements (Wixted et al., 2018; Wixted & Wells, 2017), and the legal system has begun to accept it (People of the State of New York vs. Boone, 2019). If the legal system decides to use confidence to judge accuracy, then they must be aware of the factors that can influence the strength of this relationship. Previous research has begun to study these factors by testing emotional states (Kensinger, 2009;

Pezdek et al., 2021), distance from crime (Lockamyeir et al., 2020), sleep levels (M. A. Carlson et al., 2023), witnesses age (Winsor et al., 2021), witnesses drug usage (Pezdek et al., 2020), and other estimator variables (Carlson et al., 2017; Semmler et al., 2018). When making determinations on witness accuracy, the justice system must consider all factors that influence the confidence-accuracy relationship to decrease the number of false convictions in the United States.

References

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*(2), 203–231. https://doi.org/10.1037/0033-2909.93.2.203

Alberini, C. M. (2005). Mechanisms of memory stabilization: Are consolidation and reconsolidation similar or distinct processes? *Trends Neurosci*, *28*. 51–56.

Arbuckle, T. Y., & Cuddy, L. L. (1969). Discrimination of item strength at time of presentation. *Journal of Experimental Psychology*, *81*(1), 126–131. https://doi.org/10.1037/h0027455

Barch, D. (n.d.). *Chapter 8: Signal Detection Theory*. In *Advanced Statistics I & II*. Bookdown. https://bookdown.org/danbarch/psy_207_advanced_stats_I/signal-detection-theory.html

Bares, C. B. (2011). Emerging metacognitive processes during childhood: Implications for intervention development with children. *Child and Adolescent Social Work Journal*, *28*(4), 291–299. https://doi.org/10.1007/s10560-011-0233-1

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, *71*(2), 287–311. https://doi.org/10.1093/poq/nfm006

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders*. International Universities Press.

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*(5), 610–632. https://doi.org/10.1016/0749-596X(89)90016-8

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*(1), 55–68. https://doi.org/10.1037/0096-3445.127.1.55

Benjamin, A. S., & Bjork, R. (1996). Retrieval fluency as a metacognitive index. *Implicit Memory and Metacognition*, 309–338.

Bisiach, E., Rusconi, M. L., & Vallar, G. (1991). Remission of somatoparaphrenic delusion through vestibular stimulation. *Neuropsychologia*, *29*(10), 1029–1031. https://doi.org/10.1016/0028-3932(91)90066-H

Bortolotti, L. (2013). Delusion. In E. N. Zalta, *Stanford encyclopedia of philosophy*. Stanford University Press.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, *11*(5), 164–169. https://doi.org/10.1111/1467-8721.00192

Brainerd, C. J., Reyna, V. F., & Holliday, R. E. (2012). Development of recollection: A fuzzy-trace theory perspective. *Origins and Development of Recollection: Perspectives from Psychology and Neuroscience*, 101–143. https://doi.org/10.1093/acprof:oso/9780195340792.003.0005

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, *13*(2), 207–230. https://doi.org/10.1016/0010-0285(81)90008-6

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5*, 73–99.

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, *64*(2), 123–152. https://doi.org/10.1037/h0043805

Butterfuss, R., & Kendeou, P. (2021). KReC-MD: Knowledge revision with multiple documents. *Educational Psychology Review*, *33*(4), 1475–1497. https://doi.org/10.1007/s10648-021-09603-y

Callins v. Collins, 510 U.S. 1141 (1994).

Carlson, C. A., Dias, J. L., Weatherford, D. R., & Carlson, M. A. (2017). An investigation of the weapon focus effect and the confidence–accuracy relationship for eyewitness identification. *Journal of Applied Research in Memory and Cognition*, *6*(1), 82–92. https://doi.org/10.1037/h0101806

Carlson, C. A., Wooten, A. R., & Carlson, M. A. (2023). Was he the perpetrator or a bystander? Testing theories of unconscious transference for eyewitness identification. *Journal of Applied Research in Memory and Cognition, 13*(4), 543–555. https://doi.org/10.1037/mac0000155

Carlson, M. A., Carlson, C. A., & Fitzsimmons, C. L. (2023). The sleepy eyewitness: Self-reported sleep predicts eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *12*(4), 513–530. https://doi.org/10.1037/mac0000076

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging*, *20*(4), 718–721. https://doi.org/10.1037/0882-7974.20.4.718

Challis, B. H., & Roediger, H. L. (1993). The effect of proportion overlap and repeated testing on primed word fragment completion. *Canadian Journal of Experimental Psychology /*

Revue Canadienne de Psychologie Expérimentale, *47*(1), 113–123. https://doi.org/10.1037/h0078774

Chance, J. E., & Goldstein, A. G. (1979). Reliability of face recognition performance. *Bulletin of the Psychonomic Society*, *14*(2), 115–117.

Chi, M. (2013). Two kinds and four sub-types of misconceived knowledge, ways to change it, and the learning outcomes. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (2nd ed.). Routledge.

Chowdhury, R., Sharot, T., Wolfe, T., Düzel, E., & Dolan, R. J. (2014). Optimistic update bias increases in older age. *Psychological Medicine*, *44*(9), 2003–2012. https://doi.org/10.1017/S0033291713002602

Cialdini, R. (2001). The science of persuasion. *Scientific American*, *284*(2).

Clark, S. E., & Godfrey, R. D. (2009). Eyewitness identification evidence and innocence risk. *Psychonomic Bulletin & Review*, *16*(1), 22–42. https://doi.org/10.3758/PBR.16.1.22

Cohen, R. L., Sandler, S. P., & Keglevich, L. (1991). The failure of memory monitoring in a free recall task. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *45*(4), 523–538. https://doi.org/10.1037/h0084303

Cold Spring Harb Perspect Biol-2015-Squire-a021766.pdf. (n.d.).

Connors, M. H., & Coltheart, M. (2011). On the behaviour of senile dementia patients vis-à-vis the mirror: Ajuriaguerra, Strejilevitch and Tissot (1963). *Neuropsychologia*, *49*(7), 1679–1692. https://doi.org/10.1016/j.neuropsychologia.2011.02.041

Connors, M. H., & Halligan, P. W. (2015). A cognitive account of belief: A tentative road map. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01588

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes*, *51*(1–2), 26–49. https://doi.org/10.1080/0163853X.2013.855107

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*, *15*(8), 559–564. https://doi.org/10.1111/j.0956-7976.2004.00719.x

David, A. S. (1999). On the impossibility of defining delusions. *Philosophy, Psychiatry, & Psychology*, *6*(1), 17–20.

DeBruine, L., & Jones, B. (2021). *Face Research Lab London Set* (p. 281699214 Bytes) [Dataset]. https://doi.org/10.6084/M9.FIGSHARE.5047666.V5

Deffenbacher, K. A. (1991). A maturing of research on the behaviour of eyewitnesses. *Applied Cognitive Psychology*, *5*(5), 377–402. https://doi.org/10.1002/acp.2350050502

Dole, J., & Sinatra, G. (1998). Reconceptalizing change in the cognitive construction of knowledge. *Educational Psychologist*, *33*(2), 109–128. https://doi.org/10.1207/s15326985ep3302&3_5

Dougal, S., & Rotello, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review*, *14*(3), 423–429. https://doi.org/10.3758/bf03194083

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Sage Publications.

Dunlosky, J., & Nelson, T. O. (1994). Does the sensitivity of judgments of learning (JOLs) to the effects of various study activities depend on when the JOLs occur? *Journal of Memory and Language*, *33*(4), 545–565. https://doi.org/10.1006/jmla.1994.1026

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83–87. https://doi.org/10.1111/1467-8721.01235

Eccles, J. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, *44*(2), 78–89. https://doi.org/10.1080/00461520902832368

Eil, D., & Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, *3*(2), 114–138. https://doi.org/10.1257/mic.3.2.114

Estes, W. K., & Maddox, W. T. (1995). Interactions of stimulus attributes, base rates, and feedback in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1075–1095. https://doi.org/10.1037/0278-7393.21.5.1075

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Eyewitness Evidence A Guide for Law Enforcement-Research Report. (1999). *US Department of Justice, Office of Justice Programs*.

Fischer, R. P., Falkner, K. L., Trevisan, M., & McCauley, M. R. (2000). Adapting the cognitive interview to enhance long-term (35 years) recall of physical activities. *Journal of Applied Psychology*, *85*(2), 180–189. https://doi.org/10.1037/0021-9010.85.2.180

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behaviour: An introduction to theory and research* (Vol. 27).

Fisher, R. P., & Geiselman, R. E. (1992). *Memory-enhanching techniques for investigative interviewing: The cognitive interview*. Charles C Thomas.

Freeman, D. (2007). Suspicious minds: The psychology of persecutory delusions. *Clinical Psychology Review*, *27*(4), 425–457. https://doi.org/10.1016/j.cpr.2006.10.004

Giacona, A. M., Lampinen, J. M., & Anastasi, J. S. (2021). Estimator variables can matter even for high-confidence lineup identifications made under pristine conditions. *Law and Human Behavior*, *45*(3), 256–270. https://doi.org/10.1037/lhb0000381

Granhag, P. A., Ask, K., Rebelius, A., Öhman, L., & Mac Giolla, E. (2013). 'I saw the man who killed Anna Lindh!' An archival study of witnesses' offender descriptions. *Psychology, Crime & Law*, *19*(10), 921–931. https://doi.org/10.1080/1068316X.2012.719620

Gregoire Gill, M. (2003). Is it a challenge or a threat? A dual-process model of teachers' cognition and appraisal processes during conceptual change. *Educational Psychology Review*, *15*(2), 147–179. https://doi.org/10.1023/A:1023477131081

Gregoire Gill, M. G., Trevors, G., Greene, J. A., & Algina, J. (2022). Don't take it personally? The role of personal relevance in conceptual change. *The Journal of Experimental Education*, *90*(1), 1–22. https://doi.org/10.1080/00220973.2020.1754152

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*(3), 411–435. https://doi.org/10.1016/0010-0285(92)90013-R

Groninger, L. D. (1979). Predicting recall: The "feeling-that-I-will-know" phenomenon. *The American Journal of Psychology*, *92*(1), 45. https://doi.org/10.2307/1421478

Halligan, P. (2007). Belief and illness. *The Psychologist*, *20*(6), 358–361.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology*, *56*(4), 208–216. https://doi.org/10.1037/h0022263

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (3rd ed.). The Guilford Press.

Healy, A. F., & Kubovy, M. (1978). The effects of payoffs and prior probabilities on indices of performance and cutoff location in recognition memory. *Memory & Cognition*, *6*(5), 544–553. https://doi.org/10.3758/BF03198243

Heinbockel, H., Wagner, A. D., & Schwabe, L. (2024). Post-retrieval stress impairs subsequent memory depending on hippocampal memory trace reinstatement during reactivation. *Science Advances*, *10*(18), eadm7504. https://doi.org/10.1126/sciadv.adm7504

Hughes, T. A., & Sims, A. C. P. (1997). Folie a deux. In D. Bhugra & A. Munro, *Troublesome disguises: Undiagnosed psychiatric syndromes*.

Hynd, C. R. (2001). Refutational texts and the change process. *International Journal of Educational Research*, *35*(7–8), 699–714. https://doi.org/10.1016/S0883-0355(02)00010-1

Innocence Project. (2023). *Ronald Cotton*. Innocence Project. https://innocenceproject.org/cases/ronald-cotton/

Innocence Project. (2020). *How eyewitness misidentification can send innocent people to prison*. Innocence Project. https://innocenceproject.org/how-eyewitness-misidentification-can-send-innocent-people-to-prison/

Jemstedt, A., Kubik, V., & Jönsson, F. U. (2017). What moderates the accuracy of ease of learning judgments? *Metacognition and Learning*, *12*(3), 337–355. https://doi.org/10.1007/s11409-017-9172-3

Jemstedt, A., Schwartz, B. L., & Jönsson, F. U. (2018). Ease-of-learning judgments are based on both processing fluency and beliefs. *Memory*, *26*(6), 807–815. https://doi.org/10.1080/09658211.2017.1410849

Johnson, M. L., & Sinatra, G. M. (2013). Use of task-value instructional inductions for facilitating engagement and conceptual change. *Contemporary Educational Psychology*, *38*(1), 51–63. https://doi.org/10.1016/j.cedpsych.2012.09.003

Jones, A. L., Schild, C., & Jones, B. C. (2021). Facial metrics generated from manually and automatically placed image landmarks are highly correlated. *Evolution and Human Behavior*, *42*(3), 186–193. https://doi.org/10.1016/j.evolhumbehav.2020.09.002

Kendeou, P., Butterfuss, R., Kim, J., & Van Boekel, M. (2019). Knowledge revision through the lenses of the three-pronged approach. *Memory & Cognition*, *47*(1), 33–46. https://doi.org/10.3758/s13421-018-0848-y

Kendeou, P., Smith, E. R., & O’Brien, E. J. (2013). Updating during reading comprehension: Why causality matters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 854–865. https://doi.org/10.1037/a0029468

Kendeou, P., Walsh, E. K., Smith, E. R., & O’Brien, E. J. (2014). Knowledge revision processes in refutation texts. *Discourse Processes*, *51*(5–6), 374–397. https://doi.org/10.1080/0163853X.2014.913961

Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science*, *16*(4), 213–218. https://doi.org/10.1111/j.1467-8721.2007.00506.x

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, *1*(2), 99–113. https://doi.org/10.1177/1754073908100432

King, J. F., Zechmeister, E. B., & Shaughnessy, J. J. (1980). Judgments of knowing: The influence of retrieval practice. *The American Journal of Psychology*, *93*(2), 329. https://doi.org/10.2307/1422236

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review*, *100*(4), 609–639. https://doi.org/10.1037/0033-295X.100.4.609

Koriat, A. (1994). Memory's knowledge of its own knowledge: The accessibility account of the feeling of knowing. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition* (pp. 115–136). The MIT Press. https://doi.org/10.7551/mitpress/4561.003.0008

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General*, *124*(3), 311–333. https://doi.org/10.1037/0096-3445.124.3.311

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. https://doi.org/10.1037/0096-3445.126.4.349

Koriat, A., & Adiv, S. (2012). Confidence in one's social beliefs: Implications for belief justification. *Consciousness and Cognition*, *21*(4), 1599–1616. https://doi.org/10.1016/j.concog.2012.08.008

Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *32*(5), 1133–1145. https://doi.org/10.1037/0278-7393.32.5.1133

Koriat, A., & Helstrup, T. (2007). Metacognitive aspects of memory. In S. Magnussen & T. Helstrup (Eds.), *Everyday memory* (pp. 251–274). Psychology Press.

Koriat, A., & Levy-Sadot, R. (1999). *Dual-process theories in social psychology*. Guilford Press.

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 107–118. https://doi.org/10.1037/0278-7393.6.2.107

Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *The Journal of Neuroscience*, *32*(47), 16832–16844. https://doi.org/10.1523/JNEUROSCI.3016-12.2012

Korn, C. W., Sharot, T., Walter, H., Heekeren, H. R., & Dolan, R. J. (2014). Depression is related to an absence of optimistically biased belief updating about future life events. *Psychological Medicine*, *44*(3), 579–592. https://doi.org/10.1017/S0033291713001074

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language*, *37*(4), 555–583. https://doi.org/10.1006/jmla.1997.2529

Kuzmanovic, B., Jefferson, A., & Vogeley, K. (2015). Self-specific optimism bias in belief updating is associated with high trait optimism. *Journal of Behavioral Decision Making*, *28*(3), 281–293. https://doi.org/10.1002/bdm.1849

Lamont, P. (2007). Critically thinking about paranormal belief. In S. Della Sala, *Tall tales about the mind and brain separating fact from fiction*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198568773.001.0001

Lampinen, J. M., Copeland, S. M., & Neuschatz, J. S. (2001). Recollections of things schematic: Room schemas revisited. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *27*(5), 1211–1222. https://doi.org/10.1037//0278-7393.27.5.1211

Langdon, R. (2013). Folie à deux and its lessons for two-factor theorists. *Mind & Language*, *28*(1), 72–82. https://doi.org/10.1111/mila.12009

Leonesio, R. J., & Nelson, T. O. (1990). Do different metamemory judgments tap the same underlying aspects of memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(3), 464–470. https://doi.org/10.1037/0278-7393.16.3.464

Lewandowsky, S., Cook, J., & Lombardi, D. (2020). *Debunking handbook 2020*. https://doi.org/10.17910/B7.1182

Lockamyeir, R. F., Carlson, C. A., Jones, A. R., Carlson, M. A., & Weatherford, D. R. (2020). The effect of viewing distance on empirical discriminability and the confidence–accuracy relationship for eyewitness identification. *Applied Cognitive Psychology*, *34*(5), 1047–1060. https://doi.org/10.1002/acp.3683

Loftus, E., & Ketcham, K. (1991). *Witness for the defense: The accused, the eyewitness, and the expert who puts memory on trial*. St Martin's Press.

Loftus, E. F., & Zanni, G. (1975). Eyewitness testimony: The influence of the wording of a question. *Bulletin of the Psychonomic Society, 5*(1), 86-88.

Lombardi, D., Danielson, R. W., & Young, N. (2016). A plausible connection: Models examining the relations between evaluation, plausibility, and the refutation text effect. *Learning and Instruction*, *44*, 74–86. https://doi.org/10.1016/j.learninstruc.2016.03.003

Lovelace, E. A. (1984). Metamemory: Monitoring future recallability during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 756–766. https://doi.org/10.1037/0278-7393.10.4.756

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135. https://doi.org/10.3758/s13428-014-0532-5

Magnussen, S., Andersson, J., Cornoldi, C., De Beni, R., Endestad, T., Goodman, G. S., … Zimmer, H. (2006). What people believe about memory. Memory, 14(5), 595–613. https://doi-org.radford.idm.oclc.org/10.1080/09658210600646716

Magnussen, S., & Helstrup, T. (2007). *Everyday memory*. Psychology Press.

Mandler, J. M. (1984). *Stories, scripts, and scenes: Aspects of schema theory*. L. Erlbaum Associates.

Manson v. Brathwaite, 432 U.S. 98 (1977).

Marini, M., Ansani, A., Paglieri, F., Caruana, F., & Viola, M. (2021). The impact of facemasks on emotion recognition, trust attribution and re-identification. *Scientific Reports*, *11*(1), 5577. https://doi.org/10.1038/s41598-021-84806-5

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, *29*(2), 222–233. https://doi.org/10.3758/BF03194916

Mazzoni, G., & Kirsh, I. (2002). Autobiographical memories and beliefs: A preliminary metacognitive model. In T. J. Perfect & B. L. Schwartz (Eds.), *Applied metacognition* (1st ed., pp. 121–145). Cambridge University Press. https://doi.org/10.1017/CBO9780511489976.007

McKenzie, C. R. M. (1997). Underweighting alternatives and overconfidence. *Organizational Behavior and Human Decision Processes*, *71*(2), 141–160. https://doi.org/10.1006/obhd.1997.2716

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 851–861. https://doi.org/10.1037/0278-7393.19.4.851

Mickes, L. (2015). Receiver operating characteristic analysis and confidence–accuracy characteristic analysis in investigations of system variables and estimator variables that affect eyewitness memory. *Journal of Applied Research in Memory and Cognition*, *4*(2), 93–102. https://doi.org/10.1016/j.jarmac.2015.01.003

Miller, M. B., & Gazzaniga, M. S. (1998). Creating false memories for visual scenes. *Neuropsychologia*, *36*(6), 513–520. https://doi.org/10.1016/S0028-3932(97)00148-6

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat, T. S. (2022). Managing self-confidence: Theory and experimental evidence. *Management Science*, *68*(11), 7793–7817. https://doi.org/10.1287/mnsc.2021.4294

Moore, K. N., Pennekamp, P., Yu, C., & Zwemer, D. U. (2024). Eyewitness confidence does not necessarily indicate identification accuracy. *Policy Insights from the Behavioral and Brain Sciences*, *11*(2), 149–155. https://doi.org/10.1177/23727322241268384

Morris, M. R., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012). Tweeting is believing?: Understanding microblog credibility perceptions. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 441–450. https://doi.org/10.1145/2145204.2145274

Moutsiana, C., Garrett, N., Clarke, R. C., Lotto, R. B., Blakemore, S.-J., & Sharot, T. (2013). Human development of the ability to learn from bad news. *Proceedings of the National Academy of Sciences*, *110*(41), 16396–16401. https://doi.org/10.1073/pnas.1305631110

Mueller, M. L., & Dunlosky, J. (2017). How beliefs can impact judgments of learning: Evaluating analytic processing theory with beliefs about fluency. *Journal of Memory and Language*, *93*, 245–258. https://doi.org/10.1016/j.jml.2016.10.008

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational Psychologist*, *53*(3), 165–184. https://doi.org/10.1080/00461520.2017.1421465

Nauroth, P., Gollwitzer, M., Bender, J., & Rothmund, T. (2014). Gamers against science: The case of the violent video games debate. *European Journal of Social Psychology*, *44*(2), 104–116. https://doi.org/10.1002/ejsp.1998

Neisser, U. & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall: Studies of 'flashbulb' memories* (pp. 9–31).: Cambridge University Press.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation* (Vol. 26, pp. 125–173). Elsevier. https://doi.org/10.1016/S0079-7421(08)60053-5

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(4), 676–686. https://doi.org/10.1037/0278-7393.14.4.676

Neuschatz, J. S., Lampinen, J. M., Preston, E. L., Hawkins, E. R., & Toglia, M. P. (2002). The effect of memory schemata on memory and the phenomenological experience of naturalistic situations. *Applied Cognitive Psychology*, *16*(6), 687–708. https://doi.org/10.1002/acp.824

Palmer, T. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519–526. https://doi.org/10.3758/BF03197524

Pansky, A., & Tenenboim, E. (2011). Inoculating against eyewitness suggestibility via interpolated verbatim vs. Gist testing. *Memory & Cognition*, *39*(1), 155–170. https://doi.org/10.3758/s13421-010-0005-8

Payne, D. G., & Roediger, H. L. (1987). Hypermnesia occurs in recall but not in recognition. *The American Journal of Psychology*, *100*(2), 145–165.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

People of the State of New York v. Boone, 2190-11 (People of the State of New York 2019).

Perry v. New Hampshire (United States Supreme Court 2012).

Peters, E., Joseph, S., Day, S., & Garety, P. (2004). Measuring delusional ideation: The 21-item Peters et al. delusions inventory (PDI). *Schizophrenia Bulletin*, *30*(4), 1005–1022. https://doi.org/10.1093/oxfordjournals.schbul.a007116

Pezdek, K., Abed, E., & Cormia, A. (2021). Elevated stress impairs the accuracy of eyewitness memory but not the confidence–accuracy relationship. *Journal of Experimental Psychology: Applied*, *27*(1), 158–169. https://doi.org/10.1037/xap0000316

Pezdek, K., Abed, E., & Reisberg, D. (2020). Marijuana impairs the accuracy of eyewitness memory and the confidence–accuracy relationship too. *Journal of Applied Research in Memory and Cognition*, *9*(1), 60–67. https://doi.org/10.1016/j.jarmac.2019.11.005

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 587–595. https://doi.org/10.1037/0278-7393.15.4.587

Reder, L. M. (1988). Strategic control of retrieval strategies. In *Psychology of Learning and Motivation* (Vol. 22, pp. 227–259). Elsevier. https://doi.org/10.1016/S0079-7421(08)60042-0

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*(1), 1–75. https://doi.org/10.1016/1041-6080(95)90031-4

Reyna, V. F., Corbin, J. C., Weldon, R. B., & Brainerd, C. J. (2016). How fuzzy-trace theory predicts true and false memories for words, sentences, and narratives. *Journal of Applied Research in Memory and Cognition*, *5*(1), 1–9. https://doi.org/10.1016/j.jarmac.2015.12.003

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615–625. https://doi.org/10.1037/a0013684

Richter, T. (2015). Validation and comprehension of text information: Two sides of the same coin. *Discourse Processes*, *52*(5–6), 337–355. https://doi.org/10.1080/0163853X.2015.1025665

Richter, T., & Maier, J. (2017). Comprehension of multiple documents with conflicting information: A two-step model of validation. *Educational Psychologist*, *52*(3), 148–166. https://doi.org/10.1080/00461520.2017.1322968

Ross, D. R., Ceci, S. J., Dunning, D., & Toglia, M. P. (1994). Unconscious transference and mistaken identity: When a witness misidentifies a familiar but innocent person. *Journal of Applied Psychology*, *79*(6), 918–930. https://doi.org/10.1037/0021-9010.79.6.918

Sauer, J., Brewer, N., Zweck, T., & Weber, N. (2010). The effect of retention interval on the confidence–accuracy relationship for eyewitness identification. *Law and Human Behavior*, *34*(4), 337–347. https://doi.org/10.1007/s10979-009-9192-x

Sauer, J. D., Brewer, N., & Weber, N. (2008). Multiple confidence estimates as indices of eyewitness memory. *Journal of Experimental Psychology: General*, *137*(3), 528–547. https://doi.org/10.1037/a0012712

Schacter, D. L. (2012). Constructive memory: Past and future. *Dialogues in Clinical Neuroscience*, *14*(1), 7–18. https://doi.org/10.31887/DCNS.2012.14.1/dschacter

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*(4), 695–729. https://doi.org/10.1177/0539018405058216

Schmidt, S. R. (2004). Autobiographical memories for the September 11th attacks: Reconstructive errors and emotional impairment of memory. *Memory & Cognition,* *32*(3), 443–454. https://doi-org.radford.idm.oclc.org/10.3758/BF03195837

Schmolck, H., Buffalo, E. A., & Squire, L. R. (2000). Memory distortions develop over time: Recollections of the O.J. Simpson trial verdict after 15 and 32 months. *Psychological Science*, *11*(1), 39–45. https://doi.org/10.1111/1467-9280.00212

Schroeder, N. L., & Kucera, A. C. (2022). Refutation text facilitates learning: A meta-analysis of between-subjects experiments. *Educational Psychology Review*, *34*(2), 957–987. https://doi.org/10.1007/s10648-021-09656-z

Schwitzgebel, E. (2010). Acting contrary to our professed beliefs or the gulf between occurrent judgment and dispositional belief. *Pacific Philosophical Quarterly*, *91*(4), 531–553. https://doi.org/10.1111/j.1468-0114.2010.01381.x

Searcy, J. H., Bartlett, J. C., & Memon, A. (1999). Age differences in accuracy and choosing in eyewitness identification and face recognition. *Memory & Cognition*, *27*(3), 538–552. https://doi.org/10.3758/BF03211547

Semmler, C., Dunn, J., Mickes, L., & Wixted, J. T. (2018). The role of estimator variables in eyewitness identification. *Journal of Experimental Psychology: Applied*, *24*(3), 400–415. https://doi.org/10.1037/xap0000157

Sharot, T., & Garrett, N. (2016). Forming beliefs: Why valence matters. *Trends in Cognitive Sciences*, *20*(1), 25–33. https://doi.org/10.1016/j.tics.2015.11.002

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience*, *14*(11), 1475–1479. https://doi.org/10.1038/nn.2949

Sharot, T., Rollwage, M., Sunstein, C. R., & Fleming, S. M. (2023). Why and when beliefs change. *Perspectives on Psychological Science*, *18*(1), 142–151. https://doi.org/10.1177/17456916221082967

Sporer, S. L., Penrod, S., Read, D., & Cutler, B. (1995). Choosing, confidence, and accuracy: A meta-analysis of the confidence-accuracy relation in eyewitness identification studies. *Psychological Bulletin*, *118*(3), 315–327. https://doi.org/10.1037/0033-2909.118.3.315

Squire, L. R., Genzel, L., Wixted, J. T., & Morris, R. G. (2015). Memory consolidation. *Cold Spring Harbor Perspectives in Biology*, *7*(8), a021766. https://doi.org/10.1101/cshperspect.a021766

Stephens, G. L., & Graham, G. (2004). Reconceiving delusion. *International Review of Psychiatry*, *16*(3), 236–241. https://doi.org/10.1080/09540260400003982

Stone, S. M., & Storm, B. C. (2021). Search fluency as a misleading measure of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(1), 53–64. https://doi.org/10.1037/xlm0000806

Stretch, V., & Wixted, J. T. (1998a). Decision rules for recognition memory confidence judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1397–1410. https://doi.org/10.1037/0278-7393.24.6.1397

Stretch, V., & Wixted, J. T. (1998b). On the difference between strength-based and frequency-based mirror effects in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1379–1396. https://doi.org/10.1037/0278-7393.24.6.1379

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, *14*(5), 455–461. https://doi.org/10.1111/1467-9280.02453

Tekin, E., DeSoto, K. A., Wixted, J. H., & Roediger Iii, H. L. (2021). Applying confidence accuracy characteristic plots to old/new recognition memory experiments. *Memory*, *29*(4), 427–443. https://doi.org/10.1080/09658211.2021.1901937

The impact of false or misleading forensic evidence on wrongful convictions. (2023). National Institute of Justice. https://nij.ojp.gov/topics/articles/impact-false-or-misleading-forensic-evidence-wrongful-convictions#:~:text=As%20of%202023%2C%20The%20National

The science of persuasion. (2001). *Scientific American*.

Thompson, V. A. (2014). What intuitions are… and are not. In *Psychology of Learning and Motivation* (Vol. 60, pp. 35–75). Elsevier. https://doi.org/10.1016/B978-0-12-800090-8.00002-0

Tippett, C. D. (2010). Refutation text in science education: A review of two decades of research. *International Journal of Science and Mathematics Education*, *8*(6), 951–970. https://doi.org/10.1007/s10763-010-9203-x

Tompary, A., Zhou, W., & Davachi, L. (2020). Schematic memories develop quickly, but are not expressed unless necessary. *Scientific Reports*, *10*(1), 16968. https://doi.org/10.1038/s41598-020-73952-x

Trevors, G., & Duffy, M. C. (2020). Correcting COVID-19 misconceptions requires caution. *Educational Researcher*, *49*(7), 538–542. https://doi.org/10.3102/0013189X20953825

Trevors, G. J. (2022). The roles of identity conflict, emotion, and threat in learning from refutation texts on vaccination and immigration. *Discourse Processes*, *59*(1–2), 36–51. https://doi.org/10.1080/0163853X.2021.1917950

Trevors, G. J. (2023). Belief change. In P. A. Schutz & K. R. Muis, *Handbook of educational psychology* (4th ed., pp. 291–313). Routledge. https://doi.org/10.4324/9780429433726-16

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, *21*(1), 437–484. https://doi.org/10.1146/annurev.ps.21.020170.002253

Tversky, B., & Marsh, E. J. (2000). Biased retellings of events yield biased memories. *Cognitive psychology*, 40(1), 1-38.

Underwood, B. J. (1966). Individual and group predictions of item difficulty for free learning. *Journal of Experimental Psychology*, *71*(5), 673–679. https://doi.org/10.1037/h0023107

Van Den Broek, P., & Kendeou, P. (2008). Cognitive processes in comprehension of science texts: The role of co-activation in confronting misconceptions. *Applied Cognitive Psychology*, *22*(3), 335–351. https://doi.org/10.1002/acp.1418

Vesonder, G. T., & Voss, J. F. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language*, *24*(3), 363–376. https://doi.org/10.1016/0749-596X(85)90034-8

Webb, C. E., & Dennis, N. A. (2020). Memory for the usual: The influence of schemas on memory for non-schematic information in younger and older adults. *Cognitive Neuropsychology*, *37*(1–2), 58–74. https://doi.org/10.1080/02643294.2019.1674798

Wells, G. L., & Loftus, E. F. (2003). Eyewitness memory for people and events. In I. B. Weiner (Ed.), *Handbook of psychology* (1st ed., pp. 149–160). Wiley. https://doi.org/10.1002/0471264385.wei1109

Wilson, J. A. (2018). Reducing pseudoscientific and paranormal beliefs in university students through a course in science and critical thinking. *Science & Education*, *27*(1–2), 183–210. https://doi.org/10.1007/s11191-018-9956-0

Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, *13*(5), 577–592. https://doi.org/10.1162/089892901750363172

Winsor, A. A., Flowe, H. D., Seale-Carlisle, T. M., Killeen, I. M., Hett, D., Jores, T., Ingham, M., Lee, B. P., Stevens, L. M., & Colloff, M. F. (2021). Child witness expressions of certainty are informative. *Journal of Experimental Psychology: General*, *150*(11), 2387–2407. https://doi.org/10.1037/xge0001049

Wiswall, M., & Zafar, B. (2015). How do college students respond to public information about earnings? *Journal of Human Capital*, *9*(2), 117–169. https://doi.org/10.1086/681542

Wixted, J. T., Mickes, L., & Fisher, R. P. (2018). Rethinking the reliability of eyewitness memory. *Perspectives on Psychological Science*, *13*(3), 324–335. https://doi.org/10.1177/1745691617734878

Wixted, J. T., & Wells, G. L. (2017). The relationship between eyewitness confidence and identification accuracy: A new synthesis. *Psychological Science in the Public Interest*, *18*(1), 10–65. https://doi.org/10.1177/1529100616686966

Wright, A. M., & Holliday, R. E. (2007). Enhancing the recall of young, young–old and old–old adults with cognitive interviews. *Applied Cognitive Psychology*, *21*(1), 19–43. https://doi.org/10.1002/acp.1260

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2007). *Schema therapy: A practitioner's guide*. Guilford Press.