6-29-2020

Assessing the Strategic Use of Evidence Using a Psychologically Realistic Paradigm: Improving Diagnosticity of Elicited Information in the Interrogation Room

Amelia R. Mindthoff
amogh005@fiu.edu

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

ASSESSING THE STRATEGIC USE OF EVIDENCE USING A PSYCHOLOGICALLY REALISTIC PARADIGM: IMPROVING DIAGNOSTICITY OF ELICITED INFORMATION IN THE INTERROGATION ROOM

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

PSYCHOLOGY

by

Amelia Susan Mindthoff

2020

To: Dean Michael R. Heithaus
College of Arts, Sciences and Education

This dissertation, written by Amelia Susan Mindthoff, and entitled Assessing the Strategic Use of Evidence Using a Psychologically Realistic Paradigm: Improving Diagnosticity of Elicited Information in the Interrogation Room, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

_____
Ronald Fisher

_____
Timothy Hayes

_____
Ryan Meldrum

_____
Jacqueline Evans, Major Professor

Date of Defense: June 29, 2020

The dissertation of Amelia Susan Mindthoff is approved.

_____
Dean Michael R. Heithaus
College of Arts, Sciences and Education

_____
Andrés G. Gil
Vice President for Research and Economic Development
and Dean of the University Graduate School

Florida International University, 2020

DEDICATION

I dedicate my dissertation to my husband and to my parents. Their unconditional love and constant support have seen me through this incredible journey as I grew as a researcher and achieved the work that I had always dreamed of.

ACKNOWLEDGMENTS

First, I would like to thank my committee members for their support. Dr. Ryan Meldrum, I appreciate your insight on this project. Dr. Timothy Hayes, you have taught me to run more analyses in R than I ever thought I would, and I hope to show the same excitement for statistics to my students as you have to me. Dr. Ronald Fisher, I have learned so much from you over the years and aim to make as big an impact on the legal psychology field as you have. And Dr. Jacqueline Evans, you have given me so many opportunities to grow and learn, and for that, I am glad to have had you as my advisor throughout my doctoral years. I would also like to thank Dr. Iris Blandón-Gitlin, who opened my eyes to the world of interrogation research, Dr. Deborah Goldfarb, who is one of the most supportive and knowledgeable advisors I know, and Dr. Timothy Luke, who offered me valuable advice on the development of this project.

Second, I would like to thank my family for always being excited about my work, and especially my parents. You have always been there for me, from helping me with high school math and writing assignments to watching me present this work as a doctoral candidate at a national conference. I would also like to thank all the graduate students for bringing sunshine into the lab, and especially Keith and Michelle, the best office mates one could have.

Third, I would like to thank all the research assistants who worked on this project. Caroline Griffin, I was incredibly lucky to have you manage this study and be there from the start, and same for you Sabrina Seijas as you took over. Audrey Barba, you managed the data impeccably, Guillermo Lauzurique, I was so grateful to have you run countless interrogations, Isabel Blanco, your dedication to the study and training was stellar, and

Kevin Martinez, your help on both studies truly contributed to everything running smoothly. And thank you to the all other research assistants who have worked on this study.

Fourth, I would like to thank the National Science Foundation, Law and Social Sciences Program, for funding this project (Award #1824011), as well as Florida International University's University Graduate School for awarding me the Dissertation Year Fellowship.

And most importantly, I would like to thank my husband and best friend, Patrick. You have been there for me through it all, and for that and all you do, I appreciate you more than you would ever know.

ABSTRACT OF THE DISSERTATION

ASSESSING THE STRATEGIC USE OF EVIDENCE USING A PSYCHOLOGICALLY REALISTIC PARADIGM: IMPROVING DIAGNOSTICITY OF ELICITED INFORMATION IN THE INTERROGATION ROOM

by

Amelia Susan Mindthoff

Florida International University, 2020

Miami, Florida

Professor Jacqueline Evans, Major Professor

The Strategic Use of Evidence (SUE) is an interrogation method that uses strategic timing (e.g., early vs. late disclosure) and framing of evidence disclosure to elicit verbal cues that can help interrogators discriminate between liars and truth-tellers. Despite mounting empirical support for its efficacy, there are gaps in the SUE literature that the present research addresses (e.g., studying SUE using a psychologically realistic interrogation paradigm). In Study 1, community members engaged in a supposed government-funded knowledge test. During testing, a research assistant posing as another participant prompted (guilty condition) or did not prompt (innocent condition) participants to cheat. An interrogator then accused both guilty and innocent participants of cheating and questioned them using either: early disclosure (evidence presented before questioning; antithesis of SUE), late disclosure (evidence presented after questioning; original SUE), SUE-Confrontation (SUE-C; evidence presented incrementally, with statements inconsistent with the evidence being pointed out), or SUE-Confrontation/Explain (SUE-C/E; same as SUE-C, but suspects are asked to explain

statement-evidence inconsistencies). The interrogation ended with a confession elicitation phase. The results revealed that guilty (vs. innocent) participants were more likely to confess, but evidence disclosure method did not influence confessions. When SUE-C or SUE-C/E was used, guilty (vs. innocent) participants' statements were more inconsistent with the evidence. To assess whether SUE-C and SUE-C/E are also effective in enhancing deception detection accuracy, Study 2 participants viewed Study 1 interrogation videos (one guilty, one innocent) in which either early disclosure, late disclosure, SUE-C, or SUE-C/E was used. After making initial veracity judgments, participants read that interviewees had confessed, and then rendered a second veracity judgment. Findings indicated that SUE-C/E and late disclosure resulted in deception detection accuracy that significantly exceeded chance responding. Furthermore, post-confession accuracy rates exceeded chance responding in only the SUE-C/E condition. Taken together, these studies' results lend support for the implementation of SUE, and especially SUE-C/E, methods during interrogations. Future studies should focus on continuing to refine SUE and developing effective training programs for law enforcement.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

ABBREVIATIONS AND ACRONYMS

EFM	Evidence Framing Matrix
CCTV	Closed-circuit television
SoS	Shift-of-Strategy
SUE	Strategic Use of Evidence
SUE-C	Strategic Use of Evidence (Confrontation)
SUE-C/E	Strategic Use of Evidence (Confrontation/Explain)
SUE-I	Strategic Use of Evidence (Incremental)
SUE-IPR	Strategic Use of Evidence (Introduce-Present-Respond)

I. INTRODUCTION

Suspect interrogations constitute an integral part of criminal investigations. The interrogation represents a pivotal point in the course of an investigation, as it is a time during which police investigators may determine whether or not to pursue further investigation of a suspect. However, some interrogation methods used by law enforcement have the potential to result in non-diagnostic information that can ultimately lead investigators to make incorrect classifications of suspects' culpability. Such errors are highly problematic, as they can result in the release of guilty perpetrators or the detention of innocent people. Research on interrogation methods has traditionally focused on assessing the pitfalls of such non-diagnostic methods, primarily in an effort to reduce the occurrence of false confessions and subsequent wrongful convictions. However, there has been a recent shift in the research narrative: There are increasing efforts to develop methods that can improve the diagnosticity of information elicited during an interrogation. This shift in research emphasis has stemmed from a push towards promoting methods grounded in "scientific knowledge" (i.e., methods stemming from empirical research) in place of methods derived from "customary knowledge" (i.e., methods stemming from experience and passed on via observation and word of mouth; Swanner, Meissner, Atkinson, & Dianiska, 2016). The current research sought to contribute to this push towards evidence-based methods by examining the effectiveness of a theoretically-grounded evidence disclosure technique at eliciting diagnostic information and confessions. Specifically, the current studies will examine the effectiveness of several variations of the Strategic Use of Evidence (SUE) technique.

The SUE technique was designed to improve investigators' judgments about the veracity of suspects' statements during questioning via strategically disclosing evidence to the suspect (e.g., presenting evidence after obtaining suspects' statements, as opposed to the more traditional practice of presenting evidence early in the interrogation). In past research, SUE has resulted in: (1) liars' statements being more inconsistent with the evidence than truth-tellers' statements, and (2) improved deception detection accuracy (e.g., Hartwig, Granhag, Strömwall, & Vrij, 2005; Hartwig, Granhag, Strömwall, & Kronkvist, 2006). Recently, two ecologically-valid SUE variations have been introduced: SUE-Confrontation (SUE-C) and SUE-Confrontation/Explain (SUE-C/E; Tekin et al., 2015; Tekin, Granhag, Strömwall, & Vrij, 2016). In addition to inducing liars to be more inconsistent with the evidence, SUE-C and SUE-C/E also have the potential to result in guilty suspects being increasingly forthcoming and likely to offer admissions of guilt.

Although these developments are encouraging, there are gaps in the research that need to be addressed before promoting SUE-C and/or SUE-C/E to law enforcement. The aim of the current studies was to address these gaps, thereby contributing to research on science-based criminal justice practice that is relevant to enhancing investigations. Specifically, Study 1 used an experimental paradigm high in psychological realism that enabled examination of SUE's effectiveness when questioning suspects who have (or have not) engaged in an intentional transgression. Using a psychologically realistic paradigm is new to the SUE literature, as only mock crime role playing scenarios have been used in past research (i.e., participants were explicitly told to commit a "crime" they knew would not result in serious repercussions). Study 1 therefore offers improved insight into how real-world suspects may react to SUE questioning. Additionally, Study 2

examined which SUE variation best aids observers in distinguishing between guilty and innocent suspects. In sum, the current research aimed to offer new empirical evidence for the effectiveness and suitability of SUE as a method that law enforcement officers can employ in order to achieve their goals of gathering information useful to investigations and obtaining true (vs. false) confessions from suspects.

Use of Evidence in Suspect Interrogations

Evidence disclosure plays a large role in suspect interviewing. Indeed, observational studies of real-life police interrogations reveal that the presentation of incriminating evidence against suspects (whether real or fabricated) is one of the most commonly used interrogation tactics (e.g., Kassin et al., 2007; Kelly, Redlich, & Miller, 2016; Leo, 1996). Considering interrogators' frequent use of evidence disclosure in the field, it is important to understand how different types of evidence presentation affect both interrogator decisions (e.g., regarding whether the suspect is guilty or innocent) and suspect decisions (e.g., whether to make incriminating statements or not).

The Role of Evidence in Interrogations

Types of Evidence Use

In the context of interrogations, *presentation of evidence* is comprised of a cluster of tactics that interrogators employ to indicate to a suspect[1] what the authorities know, or claim to know, in an effort to elicit more information (Kelly, Miller, Redlich, & Kleinman, 2013). In line with the definition of *presentation of evidence*, Kelly et al.

[1] Depending on the type of interrogations being examined (e.g., forensic interrogations or human-intelligence gathering interrogations), the interviewee is referenced to as 'source' or 'suspect.' For the present study, only the term 'suspect' will be used for the sake of simplicity, and because this study focuses on forensic interrogations.

(2013) identified eight common forms of evidence presentation in the interrogation-related literature. These tactics include: confronting the suspect with actual incriminating evidence; confronting the suspect with fabricated incriminating evidence; evidence bluffs (i.e., indicating that potentially incriminating evidence exists, such as untested DNA); highlighting contradictions in the suspect's statements; revealing evidence to the suspect to show the suspect that he or she is not able to disclose any more useful information until he or she eventually does so; presenting results from polygraphs or other physiological measures; sharing photos or witness statements specifically; and using visual aids. Any one of these tactics can be (and often are) employed in conjunction with one or more other evidence presentation tactics.

Overall, interrogators will use *evidence ploys* to convince suspects that there is incriminating and conclusive evidence against them (Ofshe & Leo, 1997). Evidence ploys are developed with the goal of leading suspects to believe that the case against them is highly compelling and that their only choice is to confess (Leo, 2008). To achieve their goal, interrogators will confront suspects with actual, existing incriminating evidence (i.e., true evidence ploys) or with fabricated evidence (i.e., false evidence ploys). As postulated by Ofshe and Leo (1997), the use of such evidence ploys likely results in guilty suspects "giving up" and in innocent suspects recognizing that there is little hope for their situation.

Prevalence of Evidence Use

Both self-report and observational studies demonstrate evidence use to be extremely common during suspect interrogations. For instance, a survey of U.S. law enforcement investigators revealed that interrogators often "confront the suspect with

evidence of his guilt" and sometimes "imply or pretend to have independent evidence of guilt" (Kassin et al., 2007). Interrogators also reported that on rare occasions they present suspects with crime scene or victim photographs, or conduct polygraph exams and tell the suspects that they failed the test. Kelly et al.'s (2015) international self-report survey revealed similar results, as presentation of evidence was cited as the third most-used interrogation method, following rapport and relationship building and confrontation/competition. Interrogators' self-reported use of presentation of evidence was empirically supported by Kelly et al.'s (2015) subsequent content analysis of 31 recorded interrogations retrieved from the Los Angeles Police Department's Robbery-Homicide Division, which revealed that presentation of evidence was used second-most frequently during questioning, following only rapport and relationship building.

Such self-report findings are further bolstered by field studies. In his seminal observational study of live and videotaped U.S. custodial interrogations, Leo (1996) noted that confronting suspects with actual incriminating evidence was one of the two most frequently used interrogation tactics across the 153 interrogations he observed, as it was used in 85% of the cases. Another field study similarly found that the tactic U.S. interrogators' most commonly employed was confronting suspects with evidence or with the allegation that there was a witness (Wald, Ayres, Hess, Schantz, & Whitebread II, 1967). Furthermore, Leo (1996) found that confronting suspects with fabricated incriminating evidence was also often used (observed in 30% of the cases). Leo also noted that true and false evidence ploys typically occurred at the onset of the interrogation. Overall, about 90% of Leo's observed interrogations involved the detective

confronting the suspect with real or fabricated evidence and then suggesting that it would be in the suspect's best interest to confess.

Theory Linking Evidence and Confession Behaviors

Evidence plays a large role in suspects' interrogation decision making. Indeed, in their meta-analysis assessing which psychological factors underlie guilty and innocent suspects' decision to confess in experimental contexts, Houston, Meissner, and Evans (2014) found that suspects' perceptions of the strength of evidence and the proof of their guilt emerged as a significant predictor of true confessions, but not false confessions. Furthermore, in his assessment of twelve studies on English police interrogations, Vrij (2003) found that strength of evidence was a commonly reported significant predictor of confession.

Such findings support theoretical constructs that have been developed to explain why suspects confess. Moston, Stephenson, and Williamson (1992) proposed that three factors contribute to suspects' decision to confess: (1) the suspect's and the crime's background characteristics; (2) the case's contextual characteristics; and (3) the interrogator's questioning techniques. In brief, this model posits that interrogators' questioning methods are influenced by their attitudes and beliefs, both of which are influenced by the case's background and contextual characteristics, and that the suspect's initial response during an interrogation is determined by the three aforementioned factors. Additionally, the suspect might change subsequent responses as a result of changes in the interrogator's line of questioning that are made based on the interrogator's interpretation of the suspect's strategy.

Most relevant to the present research are Moston et al.'s (1992) *contextual characteristics*, one of which is evidence. According to Moston et al., the strength of the existing evidence in a case is presumably a prominent predictor of suspect behavior and interrogator questioning method. As related to interrogators, when the interrogator has access to stronger rather than weaker evidence, he has a greater number of evidence-related strategies at his disposal. For suspects, gaining a sense of what pieces of evidence the police have is highly crucial in their decision to engage in denial or confession behaviors. For example, if it is unclear to a suspect that the police have strong incriminating evidence against him, then the suspect may initially deny involvement in the crime. Denial strategies can, however, backfire and ultimately lead the suspect to confess if the interrogator later reveals that he possesses strong evidence against the suspect. Indeed, after analyzing 1,067 cases, Moston et al. (1992) found that the number of admissions of guilt was high when the evidence was strong (66.7% of 430 cases), whereas the number of denials was high when the evidence was weak (76.6% of 274 cases).

Gudjonsson's (2003) cognitive-behavioral model of confessions also attests to the role that evidence plays in suspects' confession decisions. The model posits that the likelihood of a suspect confessing is determined by a number of factors (social, emotional, cognitive, situational, and physiological) and their relationship to short- and long-term consequences of confessing. Of interest to the present research are *cognitive events* (i.e., "the suspect's thoughts, interpretations, assumptions, and perceived strategies of responding to a given situation," p.127), as these events can have implications for suspect decisions in light of evidence. Importantly, these cognitive events do not involve

the interrogator's actual behaviors, but rather the suspect's subjective perception of the situation. For example, a suspect might be more likely to confess to the extent that he perceives the interrogator to hold strong evidence against him. Considering the implications of cognitive events, interrogators presumably have the possibility of employing tactics that play off of suspects' perceptions of the evidence known to interrogators.

Summary

Interrogators in the field often make use of evidence presentation techniques during suspect interrogations. Thus, it is crucial that the effects of evidence presentation methods on interrogation outcomes, specifically confession behaviors, are examined, especially since evidence strength is an important predictor of confession decisions and because evidence plays a role in confession decision making models. As such, the next section will review existing literature that explores the relationship between evidence-related interrogation techniques and confession behaviors.

Empirical Studies Examining Evidence Presentation Tactics and Confessions

Interrogation research has primarily focused on demonstrating that coercive interrogation methods (i.e., methods that minimize suspects' perceptions of the degree to which they have the choice to confess or not) increase the rate of false confessions (see Kassin et al., 2010, for the White Paper on false confessions). A number of these coercive methods involve the use of evidence. Empirical studies, as highlighted below, have demonstrated that some ways in which evidence is used have the potential to result in false confessions.

Empirically Examining False Confessions

Kassin and Kiechel (1996) examined the effect that false incrimination evidence has on false confession rates. In their study, an experimenter told participants that they were to type letters spoken either slowly or quickly by a confederate.[2] The experimenter warned participants to not touch the ALT key because doing so would result in the computer crashing and loss of data. Regardless of whether the ALT key was in fact pressed, the computer malfunctioned during the typing task. An experimenter then accused participants of having pressed the forbidden ALT key (i.e., of having committed a "crime"). All participants denied having pressed the forbidden key. The researchers then manipulated the use of false incriminating evidence. Specifically, the experimenter asked the confederate if she saw anything, to which the confederate either "admitted" to seeing the participants press the ALT key (false incriminating evidence) or stated that she did not see what had happened (no evidence). Results revealed that false confession rates were higher when participants were presented with false incriminating evidence compared to no evidence. The effect of the evidence ploy was further exaggerated for participants in the fast-typing condition, as 100% of these 17 participants falsely confessed to pressing the ALT key.

Even indicating that potentially incriminating evidence exists (i.e., evidence bluffs) can enhance false confession rates. Using Kassin and Kiechel's (1996) ALT key paradigm, Perillo and Kassin (2011) manipulated whether participants were presented with: false witness evidence (confederate saw participants hit the ALT key); an evidence

[2] A confederate is a research assistant who is pretending to be a participant. Actual participants are not aware that confederates are part of the research team.

bluff (keystrokes were recorded and could be checked, but had not been checked yet); a combination of the false witness evidence and the evidence bluff; witness affirmation of innocence (confederate stated that participants' hands were not near the ALT key); or no tactic (confederate stated that she did not see anything and the keystroke recording was not mentioned). The researchers found that, similar to presenting false evidence, evidence bluffs increased the rate of false confessions as compared to when no tactics were used. Specifically, 87% of participants in the bluff condition, 79% in the false evidence condition, and 77% in the combination condition falsely confessed, whereas a mere 27% of participants in the no tactics condition falsely confessed (note: the false confession rate in the innocence-affirmation condition was 36%, indicating the high-pressure nature of the paradigm).

While these findings emphasize the need to steer away from deceptive interrogation methods related to evidence presentation such as false evidence ploys and bluffs, they offer no insight on how to use actual incriminating evidence in a productive manner. This lack of insight is due to the fact that these findings stem from a paradigm that restricts researchers to assessing only false confession rates, and can be addressed by employing a paradigm that facilitates examination of *confession diagnosticity* (i.e., ratio of true to false confessions). By doing so, researchers can gain better insight regarding which evidence presentation approaches are most effective in increasing true confessions while simultaneously reducing the risk of false confessions.

Empirically Examining Confession Diagnosticity

While Kassin and Kiechel's (1996) ALT key paradigm enables researchers to examine false confession rates, it does not allow researchers to examine true confession

rates. Participants are generally unaware of whether or not they even committed the "crime" that they are being accused of. Furthermore, researchers employing this paradigm presume all participants to be innocent (i.e., the computer is programmed to *always* malfunction). Thus, the ALT key paradigm does not enable researchers to determine whether interrogation tactics are *diagnostic* (i.e., increase the likelihood that a true confession is elicited while simultaneously decreasing the likelihood that a false confession is elicited).

To address this pitfall of the ALT key paradigm, Russano, Meissner, Narchet, and Kassin (2005) introduced the "cheating paradigm." With this paradigm, an experimenter has the participant work with a confederate as a pair on a group task, during which the participant collaborates with the confederate to respond to the task questions. Upon completion of the group task, the experimenter has the participant and the confederate work on an individual task, explicitly telling them that they are to work alone and are not allowed to share answers. It is during this individual task that the culpability manipulation is introduced. In the guilty condition, the confederate asks the participant for the answer to one of the task questions. By responding to the confederate's request, the participant is effectively violating one of the experimental rules and is therefore guilty. This transgression does not happen in the innocent condition, as the confederate does not prompt the participant to cheat.

After the tasks, both guilty and innocent participants are accused of cheating, interrogated, and told that they may be charged with academic dishonesty. Overall, the cheating paradigm enables researchers to determine the diagnostic value of applied interrogation tactics in a psychologically realistic setting: Participants intentionally

commit a transgression on their own volition (or not) and are led to believe that there are potential consequences for their accused actions. The cheating paradigm further offers researchers with an ethical solution to studying suspect interrogations in a more externally valid manner than the ALT key paradigm.

Since its introduction, researchers have tailored the cheating paradigm to address important questions in the interrogation field. Most relevant to the present research is the variation introduced by Evans et al. (2013), in which a participant and a confederate complete what they believe to be a national assessment of college students' knowledge. They are monetarily incentivized to correctly answer the test questions. Before beginning the assessment, the participant and the confederate are told that they should complete the assessment as a pair, but that they should not use any additional resources to answer the questions.

In the guilty condition, the confederate uses a cheat sheet to answer some of the questions before pulling out a "forbidden" cellphone to call a friend and ask for answers to the remaining questions. The confederate then copies down the test questions to give to a friend that will be participating in the experiment in the future. In the innocent condition, the confederate does not use a cheat sheet or copy down test questions; however, the confederate uses a "forbidden" cellphone to call a friend and complain about the test's difficulty (but not to cheat). In the next phase of the paradigm, an experimenter informs the participant, regardless of culpability condition, that an unusually high number of the test questions were answered correctly, leading the experimenter to believe that cheating might have occurred. The interrogation of the participant then commences.

The primary benefit of the Evans et al. (2013) paradigm lies with the procedure's complexity. Specifically, the multiple actions performed by the confederate renders fewer restrictions on the amount of information that can be elicited from participants, as compared to Russano et al.'s (2005) original cheating paradigm. In the original cheating paradigm, the entire cheating incident involves a two-sentence exchange between the confederate and the participant. The complexity of the Evans et al. (2013) variation allows researchers to better examine the effectiveness of different interrogation methods in maximizing the amount of information elicited from an interviewee—a benefit that is highly advantageous to the further development of new evidence-based interrogation tactics. Furthermore, the increased number of actions performed by the confederate contributes to a higher number of possible pieces of evidence. The increase in the amount of evidence renders the Evans et al. variation of the cheating paradigm useful when empirically assessing evidence presentation tactics, and thus useful to the present research.

Informing the Development of Interrogation Methods with Science

The confession literature's existing strong focus on false confession research is not unwarranted, as actual wrongful conviction cases indicate that false confessions can, and do, occur (see Drizin & Leo, 2004). However, legal psychology researchers have begun to develop theory-derived interrogation methods that are effective at eliciting valuable information from withholding suspects—information that can help investigators better solve cases. As Vrij et al. (2017) discuss, mounting empirical evidence underlines the ineffectiveness of coercive interrogation tactics because of their tendency to elicit false confessions, and thus emphasize the need to employ evidence-based approaches that

improve assessments of suspects' credibility (i.e., are suspects offering true or false information). One such evidence-based technique is the Strategic Use of Evidence (SUE), a technique that is meant to (and indeed does) enhance the accuracy of investigators' credibility judgments (e.g., Hartwig et al., 2006; Luke et al., 2016).

The Strategic Use of Evidence: Perspectives from the Deception Detection Literature

At its core, the Strategic Use of Evidence (SUE) technique is a questioning framework that relies on the strategic handling of available evidence in an effort to elicit useful cues to deception (Hartwig, Granhag, & Luke, 2014). Specifically, SUE involves interrogators initially withholding incriminating evidence from suspects and only divulging the evidence at strategic moments throughout the interview, all the while posing free recall questions (e.g., "Can you tell me everything that happened on that day?") and evidence-specific questions that purposefully do not disclose the incriminating evidence (e.g., asking "Were you in the park?" when shoeprint evidence already suggests that the suspect was indeed in the park). Multiple variations of the SUE technique have been examined, and all studies point towards SUE's effectiveness in enabling observers to better distinguish between liars and truth-tellers on the basis of verbal deceptive cues (i.e., statement-evidence inconsistencies), as compared to more traditional evidence presentation methods (i.e., early evidence disclosure).

Theoretical Development of SUE

Triggered by professional interrogators' common practice of presenting all of the incriminating evidence they have at the onset of the interrogation (i.e., "early disclosure;" e.g., Leo, 1996), Pär Anders Granhag, Maria Hartwig, and their colleagues have pursued

a systematic line of research examining the effects of timing and mode of evidence presentation in suspect interviews. Presenting evidence early on during an interrogation can have detrimental effects for an interrogation's outcome. Read, Powell, Kebbell, and Milne (2009) posit that early evidence disclosure can result in higher levels of suspect aggression or denial, especially if a suspect believes that the interrogator has already made up his mind about what happened. Early evidence disclosure also enables suspects to mold a story that explains the incriminating evidence in an innocuous manner, as they are aware of the evidence against them before they provide their statements.

Such reactions by suspects to early evidence disclosure can be classified as counter-interrogation strategies, which are broadly defined as suspects' methods of successfully navigating an interrogation while maintaining their credibility (Granhag, Hartwig, Mac Giolla, & Clemens, 2015). As argued by Granhag et al. (2015), "the best approach to an interrogation is to understand the counter-interrogation techniques as practiced by suspects" (p. 293). By predicting which strategies suspects will implement during an interrogation, investigators can better plan how to successfully conduct their interrogations.

To gain a sense of the types of counter-interrogation strategies suspects might implement, it is useful to consider the matter from a self-presentational perspective. DePaulo (1992) defined self-presentation as managing one's own behavior in an effort to generate a particular impression of oneself on others. Applying this perspective to the interrogative setting, it is likely that guilty and innocent suspects both aim to create the same impression, namely that of someone honest and innocent. To achieve such a goal, suspects must engage in self-control in order to appear credible, and thus their strategies

can be understood from a self-regulation theoretical perspective (Granhag & Hartwig, 2008; Granhag & Luke, 2018).

Self-regulation theory posits that as people pursue their desired goals, they evade undesired outcomes by managing their behavior (Carver & Scheier, 2012). Granhag and Hartwig (2008) clearly delineated how self-regulation theory applies to the criminal questioning context. Specifically, all suspects perceive the impending interrogation as a threat because: (1) they lack knowledge regarding what, and how much, information the interrogator holds, and (2) there is a chance that they might be deemed guilty by the interrogator. Suspects' perception of threat subsequently initiates self-regulatory processes. In attempting to restore control over the situation, suspects can engage in either behavioral methods (e.g., remain silent during the interrogation) or cognitive methods (e.g., deciding how to construct their statement). Of specific relevance to studying evidence disclosure tactics are suspects' cognitive methods, namely, information control (i.e., gaining a sense of control via gathering information about an aversive event; Johnson, 1984) and decision control (i.e., gaining a sense of control via making a decision on how to handle the aversive event; Averill, 1973). Below, these cognitive methods are described in relation to the psychology of guilty and innocent suspects.

Psychology of the Guilty and the Innocent

As discussed, both guilty and innocent suspects presumably view an interrogation as a threat (although presumably to a lesser extent for innocent suspects). However, Granhag and Hartwig (2008) posit that because guilty and innocent suspects' knowledge of crime-relevant information differs, the problems that they face during interrogation

and the tactics that they use to address these problems also differ. Specifically, a guilty suspect is typically cognizant of all crime-relevant information since he is the perpetrator of the crime. As such, during interrogation, he is faced with the problem of not letting the interrogator discover this incriminating information, and thus seeks to conceal information. Conversely, an innocent suspect wants the interrogator to learn everything that he (the innocent suspect) knows, as this should ultimately support his innocence (note: there are instances in which innocent suspects too aim to conceal information, but this phenomenon is beyond the scope of the present research; Granhag & Luke, 2018).

Even though the types of problems guilty and innocent suspects face differ, they will both presumably engage in information control as they attempt to predict what will occur during the forthcoming interrogation. Extent of information control might be especially high for guilty suspects, who will likely try to determine what incriminating evidence the interrogator already possesses. The results of these information control activities will ultimately influence the counter-interrogation strategies suspects employ during their interrogation (Granhag & Hartwig, 2008).

Although they are similar in regard to information control, guilty and innocent suspects' actions do diverge in relation to how they decide to act during the interview (i.e., decision control). Specifically, guilty and innocent suspects have to decide what information they will present and conceal while making their statements (i.e., which information management strategies to employ; Granhag & Luke, 2018). Given that guilty suspects seek to conceal incriminating information, they will presumably engage in either avoidance (e.g., omitting details from their statements) or denial (e.g., opposing evidence presented to them) strategies. It is assumed that guilty suspects engage in such strategies

because they view the interrogator becoming knowledgeable of the concealed information as an aversive stimulus, which therefore leads them to seek avoidance of (preventing an encounter with) or escape from (eliminating) the aversive stimuli (Carlson, Buskist, & Martin, 2000).

In contrast, innocent suspects will presumably be more forthcoming. Innocents' forthcomingness can be explained by two phenomena (as delineated by Granhag & Hartwig, 2008, and Granhag & Luke, 2018). First, innocent suspects might foster a belief in a just world (Lerner, 1980). In response to undeserved outcomes threatening the maintenance of a belief in a just world, people will alter their perceptions of an individual so that the individual's fate shifts from being unjust to just. Put simply, people ultimately believe that others get what they deserve (Hafer & Sutton, 2016). Thus, an innocent suspect might believe that his statement will be deemed truthful and therefore makes little attempt to hide incriminating information, because being judged as a liar counters his belief that the world is just. Second, innocent suspects' forthcomingness can be explained by the illusion of transparency phenomenon, which refers to the tendency for people to believe that their internal thoughts and feelings are "seen" by others (Gilovich, Savitsky, & Medvec, 1998). Innocent suspects might therefore overly rely on their incorrect perception that the interrogator will see their innocence. In sum, innocent suspects believe that they have nothing to hide and that the interrogator will come to recognize their "transparent" innocence, all due to their perception that "the power of their own innocence [will] set them free" (Kassin, 2005, p. 218). Thus, innocent suspects will not likely try to cover up incriminating aspects of their stories.

Theoretically-Derived Predictions and Empirical Support for Suspects' Strategies. The psychology of guilty and innocent suspects lends support to five predictions about suspects' behaviors as set forth by Granhag and Hartwig (2008). The predictions are as follows:

(1) Guilty suspects will plan/strategize before an interrogation.

(2) Guilty suspects will attempt to avoid disclosing incriminating information.

(3) When unable to avoid disclosure, guilty suspects will engage in denial (i.e., deny knowing, or contradict, incriminating evidence when posed a direct question by the interrogator).

(4) Innocent suspects will be less likely to have a plan before being interrogated, as compared to guilty suspects.

(5) Innocent suspects will be forthcoming and only tell the truth.

Mounting empirical evidence lends support to these aforementioned predictions about suspects' strategies. For instance, to examine nonverbal and verbal differences between liars and truth tellers, Strömwall, Hartwig, and Granhag (2007) directed 30 participants to engage in either a mock crime (bought or sold fake drugs) or in an innocent act (buy a food item). Next, all participants were interrogated by in-training police officers who were told that the people they would be questioning were suspected of buying or selling drugs. Prior to the interrogation, those who engaged in the mock crime were told to lie and convince the interrogator that they bought a food item, and those who engaged in the innocent act were told to tell the interrogator the truth about their actions. Results revealed that a high percentage of liars' go-to verbal strategy was to 'keep it simple' (46.7%), whereas the majority of truth tellers' primary verbal strategy

was to ‘keep it real’ (50.0%). Furthermore, a smaller percentage of liars (11.1%) indicated that they had ‘no [verbal] strategy,’ as compared to truth tellers (30.8%; note: the authors did not indicate whether this was a statistically-significant difference).

Similarly, Hartwig, Granhag, and Strömwall (2007) had participants either commit a mock crime (liars) or not (truth tellers) and then had police trainees question participants. The researchers found that a significantly larger proportion of liars (60.5%) developed a strategy pre-interrogation, as compared to truth tellers (37.5%). Furthermore, liars exhibited a range of strategies pertaining to their verbal statements (i.e., told a detailed story, avoided lying, remained consistent with their story, offered an alibi, and/or presented an unrehearsed story). Truth-tellers, on the other hand, were cooperative and primarily sought to ‘tell the truth like it happened.’

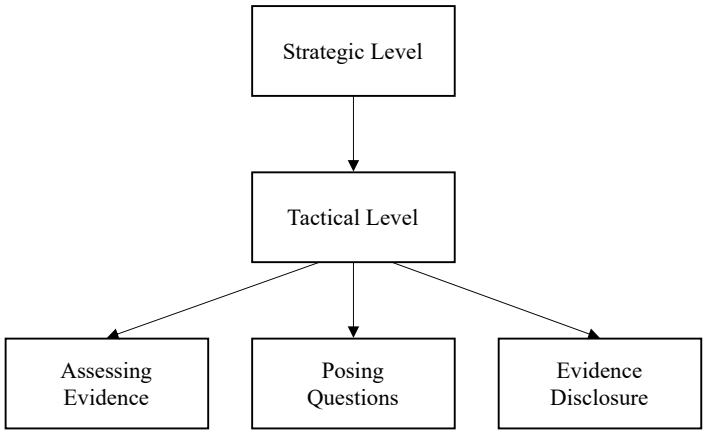
Taken together, these empirical findings lend support to Granhag and Hartwig’s (2008) five predictions regarding suspects’ interrogation strategies. The aforementioned findings regarding counter-interrogation strategies can thus be used to inform the development of best-practice interrogation methods. It is useful to know that innocent suspects are likely to be forthcoming; however, a goal of investigative interviewing is to have *all* suspects be forthcoming. Thus, shifting guilty suspects from denial and/or avoidance to forthcomingness is crucial for eliciting new critical information that can contribute to the progress of an investigation. One way to achieve this shift is to use the SUE technique to play on suspects’ perceptions of the evidence held by the interrogator.

The SUE Model

Originally proposed by Granhag (2010, as cited by Granhag & Hartwig, 2015) at a conference held by the High-Value Detainee Interrogation Group, the SUE model is a

higher order conceptualization of the practical SUE technique, and offers researchers and practitioners a conceptual framework to enhance understanding of the mechanisms by which SUE works and how to apply SUE to specific cases. Granhag and Hartwig (2015) extensively discuss the SUE model, which consists of two levels: strategic and tactical (see Figure 1). The *strategic level* consists of general principles by which SUE functions, and these principles apply to all interrogation cases. Stemming out of this level is the *tactical level*, which represents a toolbox of tactics that interrogators can employ on a case-by-case basis. These tactics are categorized into the following: *assessing the evidence* (evaluating the existing evidence pre-interrogation); *posing questions* (variations of questions potentially posed); and *disclosure of evidence* (how evidence is handled during questioning).

Figure 1. The SUE Model (Adapted from Granhag & Hartwig, 2015, p. 233)



The Strategic Level. The strategic level of the SUE model encompasses four principles. The first is *suspects' perception of the evidence*, which relates back to the previously discussed information control that suspects engage in as they attempt to determine how much information (or evidence) the interrogator actually possesses.

Suspects will presumably try to calibrate their perceptions of the evidence to match what evidence the interrogator holds. In doing so, suspects may over- or underestimate the amount of evidence held against them. Importantly, suspects' perceptions of the evidence against them can be influenced by tactics employed by interrogators, which can become useful when eliciting cues to deception (as will be further discussed).

The second principle involves *suspects' counter-interrogation strategies*, which are related to Granhag and Hartwig's (2008) previously discussed predictions about suspects' interrogation behaviors. In brief, innocent suspects are willing to tell the truth and be forthcoming, while guilty suspects engage in avoidance (e.g., being vague, omitting details) and escape (e.g., making denials) strategies. However, despite the difference in applied strategies, both innocent and guilty suspects seek to convince the interrogator that they are innocent.

The third principle is related to *suspects' verbal responses*. Suspects' statements can provide interrogators with a wealth of information, including verbal cues to deception and even new case-critical information. As related to SUE, the primary elicited deception cue is *statement-evidence inconsistencies* (i.e., the extent to which a suspect's statement is inconsistent with the evidence known to the interrogator). In their systematic review, Vredeveldt, van Koppen, and Granhag (2014) concluded that SUE tactics can be employed to effectively enhance the difference between the number of statement-evidence inconsistencies made by liars and truth tellers. Another relevant deception cue is *within-statement inconsistencies*, which relates to the extent that the details in a suspect's statement are contradictory or change over the course of the interrogation. Again, SUE tactics have been demonstrated to enhance the difference between liars' and truth tellers'

within-statement inconsistencies (Vredeveldt, van Koppen, & Granhag, 2014). Specific findings regarding statement-evidence and within-statement inconsistencies will be further discussed in the context of the tactical level.

The fourth and final principle involves *interrogators' perspective-taking* regarding the three aforementioned principles. The efficacy of SUE is enhanced when interrogators: consider how suspects might perceive the evidence against them; anticipate the counter-interrogation strategies that suspects might employ; and think about the types of verbal responses that will arise from suspects' employed counter-interrogation strategies. Ultimately, such perspective-taking can help interrogators formulate their own questioning strategy.

As Granhag and Hartwig (2015) conclude, these four strategic principles are all highly related. Specifically, a suspect will formulate a hypothesis regarding how much evidence he thinks the interrogator holds against him. The suspect's hypothesis will in turn dictate the counter-interrogation strategies that the suspect decides to employ, and these strategies will influence the verbal statements made by the suspect. Meanwhile, the interrogator that is cognizant of the suspect's engagement in strategizing will be better able to predict the suspect's strategies and the potential outcomes of those strategies, and as such, will be more effective in his decision of how and when to employ SUE tactics.

The Tactical Level. The purpose of implementing SUE tactics is to influence suspects' perceptions of the evidence, which will in turn affect their counter-interrogation strategies and, thus, verbal statements. As delineated by Hartwig et al. (2006), determining which SUE tactics to use requires the interrogator to identify accessible incriminating evidence prior to the interrogation. Such pre-interrogation planning enables

interrogators to develop a line of strategic questioning that addresses the known evidence without simultaneously disclosing it to suspects.

The SUE technique typically involves two types of questions: free recall and specific questions. *Free recall* refers to an open-ended prompt that is meant to elicit a general statement from suspects (e.g., "Please tell me everything that happened"). Such prompts are designed to trigger guilty suspects to employ avoidant counter-interrogation strategies (Granhag & Hartwig, 2015). *Specific questions* are questions that are formulated in a manner that is meant to elicit information about a specific crime-relevant detail (e.g., "Did you visit the library?" when the crime took place in the library). As posited by Granhag and Hartwig (2015), these specific questions are meant to exhaust any alternative explanations for the evidence that guilty suspects might concoct. Thus, in order to formulate effective specific questions, interrogators must predict possible explanations guilty suspects might have. Additionally, effective specific questions should address the evidence that the interrogator holds, but not cue suspects to the fact that the interrogator holds the evidence.

When used in conjunction with strategic evidence disclosure, free recall and specific questions can elicit verbal cues to deception. Specifically, Hartwig et al. (2011) had participants commit (guilty) or not commit (innocent) a mock crime, which they were later accused of committing. Participants were then questioned by an interrogator who used free recall, specific questions, or a combination of the two. In all conditions, the interrogator withheld the evidence against the suspect until all questions had been asked of and answered by participants. The results revealed that guilty participants omitted more evidence-related details during free recall responses than did innocent participants.

Additionally, guilty participants demonstrated more statement-evidence inconsistencies than did innocent participants, with the largest effect size for this difference emerging when participants were questioned with specific questions only. The difference was still significant, but smaller, for participants questioned with the combination of questions. The smallest (but still significant) difference emerged for those questioned with free recall only.

In addition to strategizing about lines of questioning, the SUE technique involves strategies on how and when to disclose evidence to suspects. There is a wide range of possibilities as to when during an interrogation evidence can be disclosed to suspects. The previously discussed *early evidence disclosure* technique represents the antithesis of SUE, and researchers generally employ the early evidence disclosure method as the control condition in SUE studies. Conversely, evidence can be disclosed after all questions have been posed (*late evidence disclosure*; note that some researchers have referred to this as SUE-Basic; e.g., Granhag, Strömwall, Willén, & Hartwig, 2013). Evidence can also be disclosed *gradually* throughout an interrogation, with evidence being disclosed one piece at a time after questions relevant to a given piece of evidence are posed and answered (e.g., Dando & Bull, 2011).

Timing of evidence disclosure differs from the manner in which the evidence is disclosed. As postulated by Granhag (2010), a single piece of evidence can be framed numerous ways along two dimensions: (1) *strength of the source of the evidence* (varies from weak to strong) and (2) *degree of precision of the evidence* (varies from vague to precise). Together, these dimensions form the Evidence Framing Matrix (EFM). To illustrate this construct, consider an example presented by Hartwig et al. (2014). The

evidence held by the interrogator is closed-circuit-television (CCTV) camera footage of the suspect in the Grand Central Terminal in New York City. In regard to the first dimension, the interrogator can disclose the source of his evidence (the CCTV footage) in a vague (e.g., "we have information that…") or precise (e.g., "we have CCTV footage that…") manner. Regarding the second dimension, the interrogator can manipulate the suspect's perception of how knowledgeable the interrogator is by directly stating he knows the suspect was at the Grand Central Station (specific) or by being more general (e.g., stating that he knows that the suspect was in Manhattan, in New York City, in the tri-state region, and so on). Overall, the EFM is meant to enhance the identification of liars, especially via increases in liars' within-statement inconsistencies. As discussed, liars will typically engage in avoidant strategies and offer little detail. Thus, when interrogators present evidence in an increasingly precise and specific manner, liars will presumably continually alter their statements so as to remain consistent with the evidence as presented by the interrogator, all in an effort to appear innocent.

Tying the Strategic and Tactical Levels Together. Granhag and Hartwig (2015) presented four associated phases (displayed in Table 1) in which SUE tactics are employed as dictated by the interrogator's perspective of the suspect's perceptions of the evidence and ensuing verbal statements. As the researchers note, these phases are very basic and are not "fixed stages" within the SUE technique. Rather, these phases conceptually represent how different SUE tactics can result in different outcomes. The researchers also note that SUE is not restricted to these phases. Overall, these phases are an ideal representation of how the discussed theory and principles come together to form the SUE technique.

Table 1. Basic Phases of the SUE Technique (Adapted from Granhag & Hartwig, 2015)

Phase 1	
1. Suspect's perception of the evidence	"They must have some information, but I am not sure how much and what.'
2. SUE tactics	Withhold the evidence; Ask for a free recall
3. S's perception of the evidence	"They do not mention any evidence; they might have less than I thought."
4. S's counter-interrogation strategy	Avoid providing information that might be incriminating.
5. S's verbal responses	Colored by *omissions*
Phase 2	
6. S's perception of the evidence	"Still not very clear how much and what information they hold."
7. SUE tactics	Withhold the evidence; Ask for a free recall; Exhaust alternative explanations; Ask specific questions
8. S's perception of the evidence	"They might have less information than I thought."
9. S's counter-interrogation strategy	Deny any incriminating actions
10. S's verbal responses	Colored by *statement-evidence inconsistencies*
Phase 3	
11. S's perception of the evidence	"Still not very clear how much and what information they hold."
12. SUE tactics	Withhold the evidence; Ask for a free recall; Exhaust alternative explanations; Ask specific questions; Disclosure according to the EFM
13. S's perception of the evidence	"They had more than I thought."
14. S's counter-interrogation strategy	"Need to alter my previous statement not to be inconsistent with the evidence presented to me."
15. S's verbal responses	Colored by *within-statement inconsistencies*
Phase 4	
16. SUE tactic	S is confronted with within-statement and/or statement-evidence inconsistencies
17. S's perception of the evidence	"They have more than I thought, better start providing them with the information they already have in order to avoid contradicting it."
18. SUE tactics	Introducing a new theme for which the interrogator lacks critical information
19. S's perception of the evidence	"I am sure they hold more information on this theme than they pretend."
20. S's counter-interrogation strategy	"Need to avoid being confronted with more inconsistencies, better tell what they already know."
21. S's verbal responses	Characterized by the suspect unintentionally telling information that is new to the interrogator

Moving from Theory to Practice: Empirical Studies

There is a growing amount of research being conducted on variations of the SUE technique. Numerous studies have examined differences in timing of disclosure, framing of the evidence, and methods by which interrogators address suspects' inconsistencies as evidence is presented throughout the course of the interrogation. Additional studies have been conducted to test the effectiveness of SUE trainings for law enforcement in terms of improving their deception detection accuracy. This section will review such studies, all of which ultimately attest to the SUE technique's efficacy in the context of suspect interrogations.

Timing and Framing of Evidence Disclosure

In the seminal SUE study, Hartwig et al. (2005) posited that the strategic manipulation of evidence presentation could function as an effective "tool" to detect deception because it should result in an increase in liars' statement-evidence inconsistencies during questioning. To test SUE's efficacy, the researchers employed a mock crime paradigm in which liars and truth-tellers were sent to a video store to pick up a DVD that was in fact not there. While searching for the DVD, both liars and truth tellers touched a briefcase with a wallet sticking out of it. Liars had been instructed to steal the exposed wallet, while truth-tellers had not. The mock crime employed created a scenario in which three pieces of evidence, all of which were incriminating but still ambiguous in terms of participants' guilt, could be obtained for both guilty and innocent suspects. Specifically, the evidence included the following: (1) a witness saw the suspect outside of the store; (2) a witness saw the suspect inside of the store; and (3) the suspect's fingerprints were on the briefcase.

All suspects were then accused of having stolen the wallet and were questioned by an interrogator who used either an *early disclosure* or *late disclosure* (the latter method represents SUE in its most basic form) technique. Prior to being interrogated, lying suspects were told to lie about having committed the mock crime, and truth telling suspects were told to offer a truthful account of what they had done; thus, as in all mock crime studies, liars are guilty suspects and truth-tellers are innocent suspects. In the *early disclosure* condition, after asking suspects whether they confessed to the crime, the interrogator presented all evidence he possessed, then allowed suspects to comment on the presented evidence before asking them to offer a free recall of what happened and then asking specific questions formulated to force suspects to address the evidence. Conversely, in the *late disclosure* condition, the interrogator asked for a free recall statement, asked the evidence-specific questions, asked for a confession, then presented all of the evidence at once, allowing suspects to subsequently comment on the presented evidence.

To determine the efficacy of late disclosure, Hartwig et al. (2005) measured the number of statement-evidence inconsistencies in suspects' statements. Specifically, the researchers coded suspects' free recall statement for whether they addressed what the interrogator knew based on the evidence, as well as coded for whether suspects' responses to the evidence-specific questions were inconsistent with the evidence (e.g., the suspect stated that he did not touch the briefcase, which contradicted the fingerprint evidence). Additionally, the researchers had observers view the video recorded interrogations and make a judgement regarding whether the suspect was lying or telling the truth (i.e., had the suspect in fact stolen the wallet).

The results were indicative of SUE's benefits. In line with the postulation that liars will make use of avoidant strategies, the researchers found that liars were more likely than truth-tellers to *not* address details suggested by the evidence during free recall in the late disclosure condition. The researchers also found that liars and truth-tellers exhibited similar rates of statement-evidence inconsistencies when responding to evidence-specific questions while being questioned with early disclosure; however, liars were significantly more inconsistent than truth tellers when questioned with late disclosure. Furthermore, observers who viewed late disclosure interrogations achieved significantly higher deception detection accuracy rates (61.7%) than those who viewed early disclosure interrogations (42.8%). While observers' accuracy rates did not differ between early and late disclosure interrogations for truth-tellers, observers were significantly more accurate when viewing liars in late disclosure interrogations than in early disclosure interrogations.

Jordan, Hartwig, Wallace, Dawson, and Xhihani (2012) also found that late evidence disclosure outperformed early evidence disclosure. After employing a mock crime paradigm similar to that employed by Hartwig et al. (2005), the researchers found that guilty suspects generally exhibited more statement-evidence inconsistencies that did innocent suspects. Notably, this difference was substantially larger when suspects were questioned with late disclosure than with early disclosure techniques.

Subsequent studies have further expanded upon disclosure timing conditions, as well as on how evidence was framed. For example, Dando, Bull, Ormerod, and Sandham (2015; first published 2013) deviated from the typical mock crime paradigm and instead used an interactive computer game in which participants either built part of an Olympic

stadium (truth-tellers) or blew up the stadium (liars). Participants were then interviewed with one of the following techniques: (1) early disclosure; (2) late disclosure; or (3) gradual disclosure (i.e., the evidence was revealed one piece at a time, with suspects having to address each piece of evidence as it was revealed and the interrogator challenging any contradictions). Unfortunately, the researchers did not assess verbal cues to deception, and only examined observers' deception detection accuracy. The results demonstrated that observers were more accurate when they viewed gradual interviews than either early or late interviews; however, the effect of disclosure method did not interact with suspect culpability.

Although Dando et al. (2015) did not assess how gradual disclosure affects guilty and innocent suspects' verbal statements, Sorochinski et al.'s (2014) study did address this topic. The researchers had participants engage (or not engage) in a mock crime and then submit to questioning on the suspicion that they were involved in a terrorist activity. Suspects were questioned with an early, late, or gradual disclosure technique. Note that the researchers' defined their gradual technique as starting with a free recall, asking evidence-specific questions, then presenting the relevant piece of evidence, repeating this process for each set of evidence pieces. Findings revealed that guilty suspects omitted more crime-relevant details during their free recall than did innocent suspects (the researchers did not assess if this differed by disclosure condition). Furthermore, guilty suspects demonstrated more statement-evidence inconsistencies than did innocent suspects when a late disclosure technique was used, but no differences between guilty and innocent suspects emerged for the early and gradual disclosure techniques.

The lack of difference between innocent and guilty suspects in Sorochinski et al.'s (2014) gradual condition is surprising. As the researchers noted, this was not the result of a decrease in guilty inconsistencies, but rather to an increase in innocent inconsistencies. As such, the researchers cautioned that innocent suspects questioned with a gradual disclosure technique realize that there is mounting evidence against them and thus begin to employ aversive strategies (e.g., try to distance themselves from the crime scene by contradicting known evidence), thereby acting more like guilty suspects. However, Sorochinski et al.'s (2014) finding diverged from Dando et al.'s (2015) finding that gradual disclosure was effective in increasing deception detection accuracy. Sorochinski et al. (2014) concluded that more research on matter is needed.

Although Dando et al.'s (2015) and Sorochinski et al.'s (2014) results attest to the effects of gradual disclosure, they did not offer insight on the effect of evidence *framing*. It is important to consider studies that address how framing evidence influences suspects' statements (i.e., Granhag et al., 2013; Luke et al., 2013). In Granhag et al.'s (2013) study, participants who had engaged in a mock crime (stole a book) or had not (checked the price of a book) were questioned with one of three techniques: (1) early disclosure; (2) late disclosure; or (3) incremental disclosure (SUE-Incremental, or SUE-I). Importantly, interrogators in this study had only one piece of evidence: incriminating, yet ambiguous, surveillance footage showing the suspect touching books on the shelf from which the book in question was stolen (but not actually showing whether the participant indeed stole the book). The SUE-I condition was characterized by incrementally presenting this one piece of evidence in three-steps, starting off vaguely (e.g., "We have information that you were there") and gradually moving to more specific (e.g., "We know you were there

because we have surveillance footage"). Examination of statement-evidence inconsistencies revealed that guilty suspects were more inconsistent than innocent suspects and late disclosure and SUE-I enhanced the difference between guilty and innocent suspects to a greater extent than did early disclosure. Additionally, guilty suspects demonstrated more within-statement inconsistencies than did innocent suspects, with this difference being largest when they were questioned with SUE-I, smaller (but still statistically significant) when they were questioned with late disclosure, and non-significant when they were questioned with early disclosure.

Luke et al. (2013) further expanded the evidence framing literature. Similar to Granhag et al.'s (2013) study, guilty and innocent suspects were questioned after committing or not committing a mock crime. Again, participants were questioned with either early, late, or gradual disclosure of a single piece of evidence (i.e., surveillance footage). However, novel to this study was that gradual disclosure either occurred in a two-step or four-step manner (for comparison, Granhag et al., 2013, presented the evidence in three steps). Not surprisingly, liars' statements contained more statement-evidence inconsistencies than truth-tellers' statements, regardless of interview style. However, the extent to which guilty and innocent suspects differed varied by disclosure method, as this difference was larger when the late disclosure and four-step methods were used compared to the early disclosure and two-step methods.

Summary. Taken together, the literature indicates that early disclosure is detrimental to interrogators' efforts to distinguish between guilty and innocent suspects. This is because liars and truth tellers behave similarly when evidence is presented early: They both integrate the evidence into their free recall statements and do not contradict the

evidence when they are asked evidence-specific questions. However, late evidence disclosure appears to consistently outperform early disclosure, both in terms of eliciting cues to deception in suspects' verbal statements and increasing observers' deception detection accuracy. Furthermore, gradual disclosure generally emerges as an effective method; however, there are some mixed results concerning this technique and more research is required.

Training SUE

Beyond being effective in enhancing the difference between guilty and innocent suspects, the SUE technique can also be effectively trained to law enforcement agents. For example, Hartwig et al. (2006) either trained or did not train 82 police trainees in methods of strategically using evidence during suspect questioning. The training session, which lasted three hours, included presentation of SUE's foundation and video examples of early and late disclosure methods. Additionally, trainees were guided on how to pinpoint incriminating information from a case file and how to use that information during questioning (i.e., develop evidence-specific questions). Lastly, trainees practiced implementing the SUE technique with one another using different case files.

Next, trained and untrained police trainees questioned suspects who had or had not committed a mock crime. Assessment of the interviews revealed that trained interviewers conducted their interviews in line with the SUE training that they had received. Furthermore, the difference between the number of statement-evidence inconsistencies made by guilty and innocent suspects was greater when suspects were questioned by a trained interviewer than an untrained interviewer. Liars specifically were more inconsistent when questioned by trained versus untrained interviewers.

Additionally, trained interviewers' veracity judgments were significantly more accurate than untrained interviewers' judgments. Specifically, trained interviewers' overall accuracy score was 85.4%, while untrained interviewers' score was 56.1%.

In another study, five experienced police officers were trained in three evidence disclosure techniques (early disclosure, late disclosure, and gradual disclosure), and then questioned mock suspects (Dando & Bull, 2011). Results revealed that when they used the gradual disclosure technique, interviewers achieved deception detection accuracy rates of 67% for liars and 74% for truth tellers. Accuracy rates were lower when late disclosure (54% for liars; 42% for truth tellers) and early disclosure (53% for liars; 47% for truth tellers) methods were used. Furthermore, the interviewers were more confident in their judgments and more likely to rely on suspects' verbal behaviors when they used the gradual technique than either the late or early technique.

Yet another training study attests to SUE's benefits. Luke et al. (2016) trained 59 law enforcement interviewers in the SUE technique, with a specific focus on applying the "funnel" structure of questioning. Specifically, this funnel method of strategic questioning involves interrogators first asking very broad questions (e.g., free recall) as related to held evidence, and becoming increasingly more specific in their subsequent evidence-specific questions (Hartwig et al., 2014). Results revealed that trained interviewers made greater use of a funnel-structure line of questioning and withheld evidence until after questioning to a greater extent than did untrained interviewers. Additionally, guilty suspects demonstrated more statement-evidence inconsistencies than did innocent suspects when they were questioned by trained interviewers—a finding that did not emerge when guilty and innocent suspects were questioned by untrained

interviewers. Furthermore, trained interviewers demonstrated higher deception detection accuracy (64.5%) than did untrained interviewers (42.9%). Considering these results in conjunction with those of Hartwig et al. (2006) and Dando and Bull (2011), it is apparent that the SUE technique can be effectively trained and can enhance interrogators' deception detection accuracy.

More Recent SUE Variations

Although the studies discussed so far attest to SUE's efficacy in enhancing the difference between liars and truth-tellers, they do not address an important goal of investigative interviewing: to have *all* suspects, both innocent and guilty, be forthcoming. As discussed previously, innocent suspects are generally forthcoming, but guilty suspects typically engage in avoidant and denial strategies. However, manipulations of guilty suspects' perception of the evidence can induce guilty suspects into becoming more forthcoming. Indeed, Luke, Dawson, Hartwig, and Granhag (2014) found that guilty suspects who were informed before their interview that the investigative team *might* have incriminating evidence were more forthcoming than uninformed guilty suspects, indicating that suspects' perceptions of the evidence against them are malleable.

Considering Luke et al.'s (2014) findings, Tekin et al. (2015) set out to extend the use of SUE to shifting guilty suspects from being avoidant to more forthcoming, and especially so in the context of offering *guilty admissions* (i.e., critical information that indicates a link between the suspect and the crime; note that this differs from a confession, in which a suspect explicitly takes responsibility for the crime). The researchers therefore developed the *SUE-Confrontation (SUE-C) technique*. Using the SUE-C technique, interrogators employ the basic SUE strategies (i.e., free recall,

evidence-specific questions) before disclosing the evidence and explicitly point out any inconsistencies between the suspect's statement and the evidence. Interrogators repeat this sequence for each piece of incriminating evidence that they hold. Tekin et al. (2015) predicted that this method would shift guilty suspects from avoidance/denial strategies towards forthcomingness as the suspects recognize that the interrogators do indeed have evidence against them and do not want to be inconsistent with this evidence at the risk of being perceived as liars. Consistent with Tekin et al.'s (2015) prediction, the researchers found that guilty suspects interviewed with SUE-C were more likely to offer self-incriminating information than those interviewed with either early or no evidence disclosure methods. The benefits of SUE-C were further bolstered by the finding that SUE-C elicited from guilty suspects more statement-evidence inconsistencies than did early disclosure.

Further building on SUE-C, Tekin et al. (2016) introduced *SUE-Confrontational/Explain (SUE-C/E)* and May, Granhag, and Tekin (2017) introduced *SUE-Introduce-Present-Respond (SUE-IPR*; note that both studies only testing these new variations on guilty, and not innocent, suspects). Unlike SUE-C, both SUE-C/E and SUE-IPR give suspects the chance to explain any inconsistencies between their statements and the evidence after the inconsistencies are pointed out to them, with the difference between these two methods being that there is a rapport-building introductory phase in SUE-IPR that is missing from SUE-C/E. Lending to support for SUE-C/E, Tekin et al. (2016) demonstrated that both SUE-C and SUE-C/E resulted in more statement-evidence inconsistencies than did early disclosure. However, only SUE-C was found to result in more guilty admissions, outperforming both SUE-C/E and early disclosure. The

researchers speculated that this finding could have been due to some of the suspects questioned with SUE-C/E believing that they had failed at achieving their goal of convincing the interrogator of their innocence on the basis of the fact that the interrogator stressed the detrimental effect any arising inconsistencies had on the suspects' credibility. As such, May et al. (2017) formulated SUE-IPR to give suspects the chance to comment on any inconsistencies, but in a non-guilt-presumptive manner (i.e., inconsistencies were not explicitly pointed out to be detrimental to the suspect's credibility, as they are in SUE-C/E). The researchers found that both SUE-C and SUE-IPR resulted in more statement-evidence inconsistencies elicited from guilty suspects, as compared to early disclosure. Additionally, SUE-IPR resulted in more new information being elicited from guilty suspects than did early disclosure, with the amount of new information elicited with SUE-C not differing from the amount elicited with SUE-IPR or early disclosure.

Overall, these studies highlight the possible benefits associated with having interrogators explicitly point out inconsistencies in suspects' statements (i.e., SUE-C, SUE-C/E, and SUE-IPR), especially in regard to enhancing statement-evidence inconsistencies. However, the extent to which allowing suspects to explain any inconsistencies is effective in shifting suspects to be more forthcoming remains ambiguous, as support for this is found with SUE-IPR but not SUE-C/E. Regardless, SUE-C/E should not be so quickly dismissed because, similar to SUE-IPR, it is characterized by greater *ecological validity* than is SUE-C (i.e., suspects are often given the opportunity by investigators to explain discrepancies in their statements; Walsh & Bull, 2015). Additionally, the existing research on SUE-C/E, and even SUE-C, is sparse, and methodological decisions made by Tekin et al. (2015; 2016) and May et al. (2017)

have rendered an incomplete picture of the benefits (and potential pitfalls) of the SUE confrontation variations. Thus, the present studies sought to further contribute to the research on these methods, as well as address past methodological concerns.

The Present Research

The reviewed literature indicates that the traditional method of early evidence disclosure is ineffective when employed during suspect interrogations when seeking to detect deception. Fortunately, the SUE technique presents itself as a valuable alternative to the early evidence disclosure method. However, the need remains for more externally-valid research on the SUE technique in the interrogative context and in relation to confession decisions, as past empirical interrogation and confession research has primarily focused on testing the effects of detrimental methods (e.g., evidence polys and bluffs). The present research aimed to address this need, as well as address other topics that have yet to be examined in the SUE literature.

Addressing Gaps in the Existing Research

While past research provides a solid foundation on which the present research was built, there are five key elements of the SUE research narrative that are missing and require empirical assessment. The goal of Study 1 and Study 2 was to address these gaps.

(1) Research on the relationship between the SUE technique and confession behaviors is lacking. Thus far, only two studies have examined confession rates elicited via the implementation of SUE. Specifically, Sellers and Kebbell (2009) manipulated evidence strength (weak or strong) and evidence disclosure timing (early or late) during the interrogations of guilty suspects who had committed a mock crime. Results revealed that confession rates were higher when the evidence was strong and when late disclosure

was used, as compared to when the evidence was weak or when early disclosure was used. Jordan et al. (2012) also examined both guilty and innocent mock suspects' confession behaviors in relation to early and late evidence disclosure. However, the resulting confession rates were extremely low (11.1%), and any further analyses of confession behaviors were not appropriate. Although these two studies did address confession behaviors, only early versus late disclosure conditions were examined, Sellers and Kebbell (2009) did not compare guilty and innocent suspects, and all participants in both studies were aware that they were involved in a mock crime/interview that was part of the experimental procedures. Thus, there is a need to examine confession behaviors while addressing the latter pitfalls in the past research. By examining confession rates as an outcome variable in the present research, findings can contribute to the achievement of societally relevant outcomes, as the implementation of an evidence-based interrogation method, such as SUE, should presumably lead to a decrease in false confessions, and thus, miscarriages of justice.

(2) None of the SUE variations have been tested with an interrogation paradigm characterized by psychological realism, which enables the examination of suspects who have (or have not) committed a transgression on their own volition (i.e., a transgression that is not known to be "part of the study"). All SUE studies to date have employed some variation of a mock crime paradigm: Participants were explicitly told by research teams to engage in guilty or innocent behaviors, guilty participants knew that the crime they were committing was fake (e.g. "go to the movie store and steal the wallet from the briefcase on top of the box of DVDs"), and participants were aware that no serious consequences would stem from their interrogations. Even though participants in such

paradigms reported being motivated to lie or tell the truth, findings arguably lack generalizability to criminal interrogation contexts. The lack of generalizability is due to the fact that "guilty" participants in these mock-crime studies, unlike real life suspects, do not engage in conscious transgressions with serious consequences. Thus, it is important that SUE be tested using a psychologically real paradigm, which will enhance the relevance of the findings to professional investigators and thereby facilitate cooperation between scientists and practitioners in the pursuit of effective, evidence-based interrogation methods.

(3) SUE-C and SUE-C/E have yet to be tested with innocent suspects in any context. Tekin et al. (2015; 2016) noted that they only studied guilty mock suspects because innocent suspects have been demonstrated to be forthcoming and consistent with the evidence, regardless of the type of evidence disclosure method used. However, the researchers neglected the fallibility of human memory, as even witnesses willing to offer information make memory errors or simply forget to mention certain details (Vrij, Hope, & Fisher, 2014). Thus, denying innocent suspects the chance to explain inconsistencies, as in the case with SUE-C (but not SUE-C/E), might be detrimental to the credibility of their statements because their statements may be rendered similar to those of guilty suspects (e.g., innocent suspects might have incorrectly recalled a detail that is inconsistent with the evidence, thus making them appear guiltier). Studying both innocent and guilty suspects can thus advance the understanding of how innocent and guilty suspects will (not) differentially behave in light of implementing SUE-C and SUE-C/E.

(4) SUE-C and SUE-C/E have yet to be compared to late disclosure, which represents SUE in its most basic form (i.e., all evidence is presented at once after

investigators ask all questions). Comparing SUE-C and SUE-C/E to late disclosure is important because implementation of these confrontation-based SUE variations in the field would require more training than would the relatively simple late disclosure variation. Should late disclosure perform on par with the other SUE variations, it would be preferable to law enforcement because of reduced training time commitments.

(5) Observers' culpability judgments of participants interrogated with SUE in a psychologically real interrogation experimental procedure have yet to be examined. Nor have their culpability judgments been assessed in any context when participant suspects are questioned using SUE-C and SUE-C/E. SUE has been shown to increase police trainees' accuracy rates to as high as 85.4% (Hartwig et al., 2006), as well as improve laypersons' deception detection accuracy (e.g., Hartwig et al., 2005; Jordan et al., 2012). However, it is unknown whether these high accuracy rates will be maintained when observers are viewing participant suspects who had been subjected to a psychologically real interrogation. Addressing this gap is particularly important because past SUE studies examining observers' deception detection accuracy lacked key elements associated with real-world criminal investigations: Mock crime suspects do not have a self-generated motive to lie and, if guilty, are instructed to lie rather than given the self-selected choice to lie or tell the truth (Levine, 2018). These two latter points may result in instructed lies not being reflective of lies rendered in the real-world and may have implications for observers' ability to make accurate culpability assessments (Levine, 2018). Additionally, it is unknown whether SUE-C or SUE-C/E will further enhance culpability accuracy rates. If SUE-C and/or SUE-C/E result in lower culpability accuracy rates, it will be

important that further evaluation of these methods is conducted before recommendations are made to law enforcement professionals.

II. STUDY 1

Study 1 examined four evidence disclosure methods (early disclosure, late disclosure, SUE-C, SUE-C/E) during the interrogations of participants who had or had not committed transgressions on their own volition during a modified version of Evans et al.'s (2013) interrogation cheating paradigm. Note that early disclosure served as the "control" condition, as it is a typically employed interrogation strategy (Leo, 1996); yet most, if not all, SUE studies have already demonstrated its inferiority in terms of eliciting reliable cues to deception. Study 1's research questions and associated hypotheses were as follows:

(1) *Is the SUE technique more diagnostic (as related to confessions) than the traditional early disclosure method (i.e., does SUE result in more true confessions and fewer false confessions, compared to early disclosure)?* I predicted that confessions would be more diagnostic in the three SUE conditions than in the early disclosure condition. Empirical examination of Gudjonsson's (2003) Cognitive-Behavioral model of confessions indicates that true confessions are related to perceived evidence strength—a perception that all SUE techniques aim to manipulate. Further, I predicted that SUE-C/E would result in even more diagnostic confessions than would the SUE-C and late disclosure conditions. This is because SUE-C/E's explanation component would offer an extra protective layer for innocent suspects whose memory might have failed them (this component is lacking in SUE-C). Additionally,

by the end of the interrogation, innocent suspects might not remember all of the inconsistencies they made throughout the interrogation (which they would have to explain all at once at the conclusion of questioning in the late disclosure condition). Instead, innocent suspects may have an easier time explaining inconsistencies when inconsistencies are incrementally presented (as in the SUE-C/E condition). I also predicted that late disclosure would outperform SUE-C in terms of diagnosticity, as suspects would still get the chance to explain inconsistencies after all of the evidence has been presented to them, thereby giving them a chance to reinforce their innocence.

(2) *Can SUE techniques yield statements with more inconsistencies compared to early disclosure in more realistic questioning scenarios?* I predicted that all three SUE methods would result in more statement-evidence inconsistencies than would early disclosure—a prediction that is in line with the reviewed research. Furthermore, I predicted that the SUE methods would result in guilty participants making more statement-evidence inconsistencies than would innocent participants; early disclosure would not, however, result in significant differences in statement-evidence inconsistencies between guilty and innocent participants. I also predicted SUE-C and SUE-C/E to perform similarly (as demonstrated by Tekin et al.'s, 2016, results). Lastly, I predicted that both SUE-C and SUE-C/E would result in overall fewer statement-evidence inconsistencies compared to late disclosure. This prediction is supported by self-regulation theory, as guilty suspects' avoidance/denial strategies were expected to shift to forthcoming strategies as they attempted to

regain their credibility in the SUE-C and SUE-C/E conditions. The expected shift to forthcoming strategies, which is in line with Tekin et al.'s (2016) findings and was not expected to occur in the late disclosure condition, would thus presumably result in fewer statement-evidence inconsistencies in the SUE-C and SUE-C/E conditions.

Method

Participant Recruitment and Sample Size

I recruited adult participants ($N = 219$) via flyers, online platforms, and snowball sampling. Research assistants posted the flyers throughout communities in Miami-Dade and Broward counties, and the Miami-Dade County public library system agreed to display flyers in all 50 of their branches. Additionally, I posted the recruitment notice on online platforms including Craigslist, Eventbrite, OfferUp, and Nextdoor. All advertisements denoted that the study was part of a national assessment of general knowledge, thereby masking the true nature of the study and facilitating psychological realism during the interrogation phase of the study. The advertisements instructed participants to either call or email the research team to set up a study appointment, as well as indicated that participants would be compensated $30, with the opportunity of receiving an extra $10 (for a total of $40; note: participants also received a $2 travel compensation). In reality, however, all participants received the full $42 (the falsely advertised $30 with the chance to earn more money was integral for the research design, as discussed below).

Participant Sample Characteristics

Eighteen participants either withdrew their consent for the usage of all their data (both paper and video) for the present study or did not complete the study (e.g., the study was stopped because the participant became too upset, the interrogator personally knew the participant); thus, the total sample size for all analyses (other than those analyses dependent on the video footage of the interrogation) was 201. The total sample size of the video recordings collected was 191, as there were no videos for 10 participants for the following reasons: some participants withdrew consent for use of their videos specifically; the recording system failed to record the interrogation session; or the recordings were corrupted.

Participants' ages ranged from 18 to 80-years-old ($M = 41$, $SD = 15$; $n = 198$[3]). Overall, the participant sample was diverse ($n = 196$): 37.8% participants identified as Hispanic/Latinx, 29.6% as Black, 24.0% as White, 1.5% as Asian, 1.5% as Native American, and 5.6% as Other. A little over half of participants were male (56.7%; $n = 194$). The majority of the participant sample were not students (79.8%; $n = 198$) and indicated that they were fluent in English (99.0%; $n = 198$). Additionally, a large portion of the sample indicated that they were currently not employed (46.5%; $n = 198$), with the remainder of the sample indicating that they were either employed less than 15 hours/week (8.1%), employed 15 to 35 hours/week (19.7%), or employed fulltime (25.8%).

[3] Not all participants chose to respond to demographic questions. The total number of participants who responded to each question is therefore listed with the corresponding results.

Arrest and Interrogation Experiences. Of the 198 participants who responded regarding whether they had ever been arrested, 51.5% indicated that they have previously been arrested. Table 2 displays the number of times participants indicated that they had been arrested. Additionally, participants who had been previously arrested indicated whether they had been interrogated by a police officer after being arrested.[4] Of the 67 participants who had been arrested and responded to this question, 61.2% indicated that they had indeed been interrogated by a police officer (Table 2 displays the number of times participants had been interrogated). Those participants who had been interrogated also indicated whether or not they had confessed during an interrogation and, if so, how many times they had truthfully and/or falsely confessed. Of the 41 participants who reported that they had been interrogated, 31.7% had confessed. Table 2 displays the breakdown of the number of times these confessors had made a truthful and/or false confession.

[4] I added interrogation detail questions to the demographic questionnaire after data collection had already begun; thus, not all participants who had been arrested responded to these interrogation-related questions.

Table 2. Participants' Arrest and Interrogation Self-Reported Details

Category	Number of Participants Who Responded	% of Total Participants Who Responded
Number of Times Arrested (n = 99)		
1 time	30	30.3%
2–5 times	43	43.4%
6–10 times	12	12.1%
More than 10 times	10	10.1%
No exact number*	4	4.0%
Number of Times Interrogated (n = 40)		
1 time	16	40.0%
2–5 times	19	47.5%
6–10 times	2	5.0%
More than 10 times	3	7.5%
Number of Times Made a Truthful Confession (n = 13)		
Never	1	7.7%
1 time	4	30.8%
2 times	3	23.1%
3 times	3	23.1%
4 times	2	15.4%
Number of Times Made a False Confession (n = 13)		
Never	6	46.2%
1 time	4	30.8%
2 times	1	7.7%
6 times	1	7.7%
More than 10 times	1	7.7%

Note. * These participants gave responses such as "don't remember" and "a lot."

Design

Study 1 implemented a 2 (*culpability:* guilty, innocent) x 4 (*evidence disclosure:* Early disclosure, Late disclosure, SUE-Confrontation, SUE-Confrontation/Explain) between-subjects design. Primary outcome variables included confession rate, statement-evidence inconsistencies, and participant subjective ratings for their perceptions of the evidence and the interrogation strategies that they employed.

Materials

Psychologically Realistic Cheating and Interrogation Paradigm

One participant was run per experimental session. Participants arrived at the lab along with a research assistant posing as another participant (i.e., a confederate[5]). An experimenter greeted them and proceeded with the consent process, which was consistent with the recruitment cover story in order to mask the true nature of the study. After obtaining consent, the experimenter directed the pair to leave all their belongings in the main lab room before leading them to the testing room where the remainder of the study took place. The testing room had a one-way mirror and was wired to covertly video and audio record everything that took place during the experimental session. Participants were unaware that they were being recorded at this point.

Next, I implemented a variation of the Evans et al. (2013) paradigm. Specifically, the experimenter informed the participant-confederate pair that they would be participating in a government-funded national research project headed by the Department of Educational Assessment,[6] and that the goal of the project was to determine the status of general knowledge among community members. The experimenter further indicated that the first part of the study consisted of answering 20 general knowledge questions, and that the pair "need[ed] to work alone on answering these questions" and "[could not] use any additional resources, such as cellphones, or help each other while working on these questions." Lastly, the experimenter informed the pair that they had the chance to

[5] Over the course of data collection, the confederate role was assumed by both male and female research assistants. The pronoun "he" is used here only to enhance readability.

[6] This was a fictional government department created for the purpose of this experiment.

receive up to an additional $10, as they each would be paid $0.50 for every question they answered correctly.[7] The experimenter then administered the test, provided the pair with some additional scrap papers, and left them alone in the room, stating that she would return in about 15 minutes to check on the pair's progress.

Guilty Condition. A couple of minutes after the experimenter left, a cellphone dinged, and the confederate retrieved it from his pocket and pulled up a photo of a paper with answers to the test questions. The confederate placed the phone on the desk and began to copy down the answers. He told participants that his friend had already participated in the study and gave him the questions. He then stated that he really needed all the money he could get so that he would be able to pay his bills this month, and thus he had looked up all the answers beforehand. After completing his test via copying the answers off his phone, the confederate pushed the phone in front of participants, asking participants if they wanted to use the "cheat sheet." Participants, assuming that they used the cellphone, were then guilty of engaging in a forbidden behavior (i.e., using an additional resource to cheat).[8] Once the confederate took the phone back, he took pictures of the test and stated that he sent the pictures to participants,[9] suggesting that participants could share the answers with friends so that they, too, could sign up for the study and make the extra $10.

---

[7] As previously noted, all participants received the full $40, regardless of their performance on the test.

[8] Not all participants in the guilty condition chose to use cheat using the phone. This fact is accounted for in the Results section.

[9] In reality, the confederate did not send participants the answers. If participants asked how the confederate got their contact information, the confederate indicated that their email address was listed on the study reminder email.

Innocent Condition. As in the guilty condition, after the experimenter left the testing room a cellphone dinged, and the confederate retrieved it from his pocket and placed it on the desk next to him. The confederate told participants that he really needed all the money he could get to pay his bills, so it was important that he figure out all the correct answers to the test questions. The pair then worked on the test questions while the phone remained on the desk. Upon completing the questions, the confederate picked up the phone and pulled up a photo of a piece of paper on which illegible text was scribbled. The confederate stated that his roommate had sent him a grocery list of items they needed at home, but that he could not decipher the list. Then, asking if they could read the grocery list and write it down on a piece of scrap paper, the confederate pushed the phone in front of participants. Thus, participants were innocent of engaging in any forbidden behaviors (i.e., they were not using additional resources to cheat), but critically, were still interacting with a cellphone and subsequently writing something down. After participants wrote down the grocery list, the confederate folded the paper and pocketed it. The confederate then took pictures of the test, indicating that he was sending the test questions to a friend that would be taking the study at a later point. Note that if participants in the innocent condition suggested using the phone to cheat at any time, the confederate would reiterate the experimenter's instructions.

Pre-Interrogation Phase. The experimenter returned after approximately 15 minutes (by which time the confederate had replaced the cellphone in his pocket) and stated that the pair would separately engage in an individual computer task. The experimenter then escorted the confederate out of the testing room, giving participants a questionnaire and new puzzle task to work on while they were waiting for their turn at the

computer task. Soon after, an interrogator,[10] blind to participant culpability, entered the testing room. He informed the participants that there was an issue with the test that they just turned in and that he would return shortly to discuss the issue. He then administered another questionnaire to participants in order to "keep the study going" and left. After five minutes, the interrogator returned, and the interrogation portion of the study commenced.

Interrogation Phase. The interrogator began the interrogation with the same scripted opening statement, regardless of whether or not participants had actually cheated. Specifically, the interrogator stated that he had a reasonable suspicion that the participant-confederate pair cheated on the test. The interrogator further stated that cheating on this government-issued test was a major issue, and that the administration protocol sent from the Department of Educational Assessment indicated that cheating was considered to be the provision of fraudulent responses, which could ultimately result in a monetary fine of $50–$500 depending on the severity of the cheating.

The interrogator then followed a randomly assigned protocol consistent with one of the four evidence presentation conditions (Early disclosure, Late disclosure, SUE-C, SUE-C/E). Note that all interrogators underwent extensive SUE training in line with past SUE training studies (e.g., Hartwig et al., 2006; Luke et al., 2016).[11] Specifically, interrogators were trained in the following: basic understanding of SUE's foundation;

[10] Over the course of data collection, the interrogator role was assumed by primarily male research assistants. The pronoun "he" is used here to enhance readability.

[11] I did not inform interrogators about any of my hypotheses, so as to prevent experimenter expectancy effects.

specifics about each of the four techniques and how to apply the techniques; and how lines of questioning are developed in relation to each piece of evidence that exists for a case. Additionally, interrogators learned the specifics about the study's case, including what evidence would exist and the types of questions that should be associated with the evidence, and then practiced the four techniques on mock participants until they were able to complete interrogations in a believable and effective manner.

The same four evidence pieces existed for both culpability conditions (Table 3).[12] Crucially, all evidence was incriminating but not absolutely indicative of cheating (i.e., guilt was ambiguous). Indeed, the evidence was consistent with both innocent and guilty participants' behaviors.

[12] In some participant cases, not all evidence pieces occurred. This happened primarily for Evidence Piece #4 when participants had strong negative reactions to the phone and would immediately push it out of their vicinity. Such instances are addressed in the analyses.

Table 3. Evidence Specifics and Related Specific Questions

Evidence Piece	Evidence Specifics	Related Specific Questions
#1	A research assistant from a different study (Witness 1) heard a phone go off in the testing room as she walked by.	Did you hear any unexpected sounds, for a test setting, while you were taking the test?
#2	Another research assistant (Witness 2) went into the room next door to the testing room to get a folder. When he was looking for the folder, he said that he happened to look through the one-way glass and saw a phone sitting on the table.	Other than the pen and scrap paper the experimenter gave you, were there any additional resources in this testing room that weren't supposed to be here?
#3	Because Witness 2 could not find the folder he was looking for, he asked another research assistant (Witness 3) to find it. Witness 3 said that when she entered the room next door to the testing room, she saw that the person sitting the farthest from the door was holding a phone.*	[Did you use or have access to / Were you in the vicinity of] any additional resources that weren't allowed at any point during the testing session?
#4	Witness 3 said that she found her folder after searching for a bit, and when she looked into the testing room again, she saw the participant writing something down while looking at the phone. She thought it was odd, considering the participant was in a testing setting, so she took a picture of the participant writing something down while looking at the phone.	Were you writing anything down while using or accessing any additional resources that weren't allowed?

Note. The witnesses are fictional. The experimenter observed the entire testing session and recorded the evidence as it occurred. Also, interrogators reworded the evidence-related questions according to participant responses so that questioning would not seem artificial, as well as in a manner to ensure that the evidence-related questions matched exactly how the evidence had played out during the testing session. *The participant was always the person to be sitting farthest from the door.

In the ***SUE-C/E condition*** (see Figure 2), the interrogator first asked a free recall question ("So, in as much detail as possible, please tell me everything that happened from the time you arrived here at the lab up until I entered the room."). Following participants' responses, the interrogator then posed Specific Question #1. If participants' responses

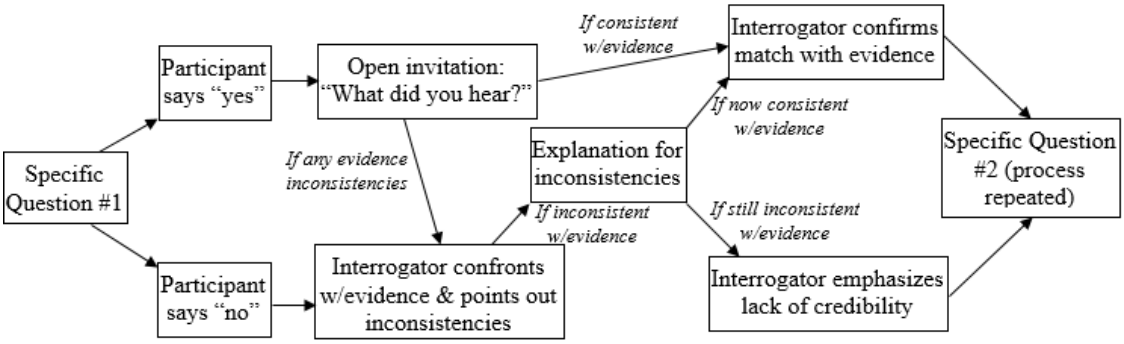
were characterized by *denial* ("I didn't hear anything"), the interrogator presented Evidence Piece #1 and, in a non-accusatory manner that did not indicate presumption of guilt, stressed the seriousness of concealing information[13] ("We have a witness saying that she heard a phone in the room. Considering this, it seems to me that you're not matching up with what the research assistant told me, which is a bit strange."). Conversely, if participants' responses were characterized by *complete agreement* with the evidence ("I heard a phone go off"), the interrogator would confirm a match ("You say you heard a phone ding, and we have a witness confirming this"). For instances in which participants indicated *vague agreement* with the evidence (e.g., "yes" or "I heard something"), the interrogator would follow up with an open invitation ("What did you hear?"). The interrogator confirmed a match if participants' responses to this open invitation were evidence-consistent. However, if any information was omitted or inconsistent in response to the open invitation, the interrogator would confront participants with the evidence and stress the gravity of the situation in the same manner as if the participant had denied hearing anything at all.

After any instance of evidence confrontation that resulted in statement-evidence inconsistencies, the interrogator would ask participants to explain the inconsistencies ("Why aren't you matching up with the evidence that I have?"). The interrogator would confirm participant explanations that were consistent with the evidence, but if participant explanations remained inconsistent with the evidence ("I really didn't hear a phone"), the interrogator would again stress the severity of persisting inconsistency ("You're still not

[13] This method of confrontation is similar to that employed in SUE-IPR; however, I was not testing SUE-IPR, which involves an additional introductory phase that was not tested in this study.

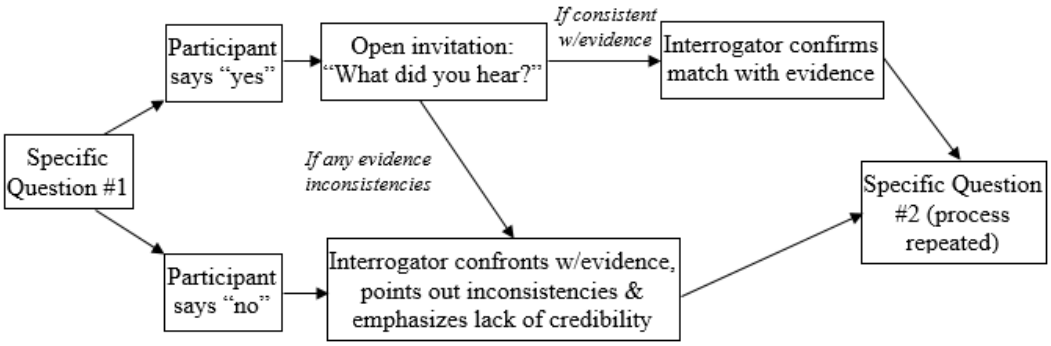
matching up with the evidence, and that doesn't make sense"). This process was repeated for all four evidence pieces.

Figure 2. The SUE-C/E Questioning Process



The ***SUE-C condition*** was identical to the SUE-C/E condition, except that participants were *not* given the opportunity to address any statement-evidence inconsistencies that were pointed out by the interrogator. After pointing out any inconsistencies, the interrogator immediately proceeded to the next specific question (see Figure 3). In the ***late evidence disclosure condition***, the interrogator first asked participants the free recall question, followed by the specific questions from Table 3. Then, the interrogator presented all of the evidence at once in a narrative format and asked participants an open follow-up question ("Can you explain this?"), giving participants the chance to respond to the evidence. In the ***early evidence disclosure condition***, the interrogator presented all of the evidence (in narrative format) to participants *before* asking any questions. Then, the interrogator asked the aforementioned free recall question and the specific questions displayed Table 3.

Figure 3. The SUE-C Questioning Process



Confession Elicitation. Once questioning concluded, the interrogator exited the room for three minutes to "review his notes" (in line with Tekin et al., 2015; 2016). Upon reentry, the interrogator asked an open-ended question about whether anything else happened that he should know about that had occurred after the phone was on the table. Next, the interrogator told participants that he reviewed the protocol that the Department of Educational Assessment had sent along with the test and, according to the protocol, he needed to explicitly ask participants whether or not they had cheated on the test. He then presented participants with an official form on which the following statement was written: "Participant used additional resources to provide fraudulent responses."

As he presented the form, the interrogator indicated that if participants signed the form, they could "wrap this up." However, if they refused to sign the form, the interrogator indicated that the protocol stated that an affiliate from the Department of Educational Assessment would have to "continue looking into [the] matter more in-depth." The interrogator concluded by stating, "So, if you cheated, I would like you to admit to it and sign this form." If participants refused to sign the form during this first prompt, the interrogator reiterated the outcomes of signing/not signing the form and again

asked for a written confession. The confession elicitation ended if participants refused to sign on the second prompt. The interrogator ended the interrogation by stating that he did not think participants would be in any trouble (to minimize participant distress) and left.

Participant Debriefing. As soon as the interrogator exited the testing room, the experimenter entered and commenced debriefing. First, the experimenter asked participants what they thought the study was about (i.e., a manipulation check to assess the paradigm's believability), and whether the experimental team was trying to trick or deceive them in anyway. Participants then completed the post-study questionnaire (described below).[14] After they completed the questionnaire, the experimenter told participants what the true nature of the study was and informed participants that the confederate and the interrogator were actually part of the research team. The experimenter also repeatedly assured participants that they were not in any trouble whatsoever, as there was no real government-issued test or monetary fine.

The experimenter then told participants that the entire portion of the study that occurred in the testing room had been covertly video-recorded. At this point, the experimenter presented participants with the final consent form and the video consent form. The final consent form gave participants the option to allow me to use their data or to completely withdraw their consent for their data. Regarding the video consent form, participants had the option to: (1) allow me to use their video footage for the present study, as well as potential future studies; (2) allow me to use their video footage for the

[14] Although some of the post-interrogation questions may have suggested to participants that the accusation and interrogation were part of the study, withholding the official debriefing until after participants responded to these questions should have minimized their knowledge about the true intent of the research while they completed the questionnaire.

present study only; or (3) deny permission to use their video footage for research and have it deleted. The experimenter immediately deleted the videos for participants who responded with the latter option.

Lastly, the experimenter emphasized to participants the importance of maintaining psychological realism for future participants by not spreading word about the true nature of the study. Participants then signed a confidentiality agreement indicating that they would not discuss study details with potential participants and that they would tell potential participants who asked about to the study that they just had to answer questions and fill out some paperwork. Finally, participants were compensated their full payment ($42).

Post-Debriefing Questionnaire

The post-study questionnaire's function was to garner insight into participants' interrogation experiences and to assess Houston et al.'s (2014) proposed underlying psychological predictors of confessions. First, participants responded to questions regarding their perceptions of how much evidence the interrogator had supporting the cheating accusation both for when the interrogator first accused them of cheating and when the interrogator had finished questioning them (1 = *the interrogator knew nothing about what happened during the testing session* and 7 = *the interrogator knew about everything that happened during the testing session*). Participants next responded to questions about the extent to which they perceived the evidence against them to be strong both at the start of the interrogation and by the end of the interrogation (1 = *not strong at all* and 7 = *extremely strong*). Participants also indicated the extent to which they aimed to implement each of the following strategies during the interrogation (1 = *not at all* and

7 = *extremely*): to be withholding (e.g., not give information); to be in denial (e.g., reject all allegations made); and to be forthcoming (e.g., voluntarily give information).

Regarding Houston et al.'s (2014) proposed psychological predictors of confessions, participants responded on a seven-point Likert-type scale (1 = *not at all* and 7 = *extremely*) to items addressing: how much pressure they felt to tell the interrogator that they cheated; how anxious they were at the start of the interrogation; how guilty they felt about their behavior during the testing session; how severe they thought the consequences would be if they did tell the interrogator what happened and if they did not tell the interrogator what happened; and whether they felt that the interrogator was pushing them to confess. The post-study questionnaire ended with demographic questions as well as questions regarding criminal justice system experiences (e.g., prior arrests, prior interrogation experiences).

Procedure

Participants, along with the confederate, met the experimenter in the psychology building foyer before being led to the initial lab station. Once there, the experimenter consented participants (and the confederate), and had them complete some filler tasks (i.e., personality questionnaires). After moving to the testing room, the cheating and interrogation paradigm outlined above commenced. Following the interrogation's conclusion, the experimenter extensively debriefed participants, and administered post-debriefing questionnaires and compensated participants, thus ending the study.

Pilot Testing

Before running actual community-member participants, the entire experimental procedure as delineated above was tested and refined using a student participant sample.

The only procedural deviation was that of participant payment and incentive: Instead of $30, the student study was advertised to compensate one research credit, and instead of a chance to win an extra $10, student participants had the chance to win an extra research credit (which all participants ultimately received). The primary goals of this pilot were to: (1) determine whether the 15-minute interaction time between the participant and the confederate was too long or too short; (2) allow interrogators to practice the questioning protocol; (3) test out different prompts confederates could use to get participants to cheat / write out the grocery list; and (4) note and fix any unexpected problems that arose with the protocol and script. I pilot tested the procedure with 28 student participants. During pilot testing, I made minor refinements to the procedures and scripts. These refinements primarily involved improving the clarity of instructions throughout the experiment and interrogator statements during the interrogation.

Coding for Statement-Evidence Inconsistencies

If participants granted consent to use their videos for the present study, I sent the audio recordings of their interrogations to a third-party company for transcription. Transcriptions included participants' interrogation questioning session as well as the confession elicitation that followed. Culpability and evidence disclosure conditions were not listed on the transcriptions. Coders were, however, able to identify the evidence disclosure technique used based on the questioning format. Additionally, coders may have been able to infer participant guilt depending on the types of responses participants made (e.g., discussion of a grocery list was almost always brought up by innocent participants).

To examine the extent to which participants' statements did not match the evidence against them, two independent coders, who were blind to the study hypotheses, reviewed participant final statements made in response to the interrogator's specific questions. In the SUE-C and SUE-C/E conditions specifically, participants may have provided multiple responses to the specific questions (e.g., a response to the open invitation, a response to the interrogator's request for an explanation). In such instances, the coders rated all of participants' responses but for purposes of analyses, only participants' *final* response to each specific question was considered.

The coders evaluated "the extent to which participants' response to a specific question [was] consistent with the piece of evidence corresponding to that question." Participant responses received one of three possible scores: (1) *consistent* (the participant's response to the specific question was fully consistent with the evidence related to the specific question); (2) *not fully consistent* (the participant's response to the specific question *may* have been consistent with the evidence related to the specific question, meaning that there was a chance that the participant's response was not consistent with the evidence [e.g., the response was vague; the participant was hedging the specific question]); or (3) *inconsistent* (the participant's response to the specific question was *not* consistent with the evidence related to the specific question, meaning that the participant did not acknowledge the evidence at all, denied the evidence, or made an "I don't know" remark).

The two coders reviewed the same 45 (23.6%) transcripts so that I could assess interrater reliability. I then divided the remaining transcripts between the two coders for scoring. Overall, interrater reliability, as assessed by the Cohen kappa statistic, was either

"moderate" or "almost perfect" (McHugh, 2012).[15] Table 4 presents kappa values related to each of the specific questions.

Table 4. Kappa Scores for Coding Statement-Evidence Inconsistencies

Item	*n*	Kappa	95% CI
Specific Question #1	44	1.00	*
Specific Question #2	45	0.91	[0.78, 1.03]
Specific Question #3	45	0.69	[0.51, 0.87]
Specific Question #4	40	0.71	[0.53, 0.90]

Note. The interrogator may not have asked all specific questions (i.e., this happened in instances in which an evidence piece did not occur, the interrogator forgot to ask a question, or the participant did not provide a substantial response that was codable). Hence, the sample sizes presented here differ. CI = confidence interval. *No CI was calculated since there was perfect overlap between scores.

Results

Of the 201 participants who completed the study, I excluded 52 from all subsequent analyses. The 52 excluded participants represent participants who had been assigned to the guilty condition but had not cheated. These non-cheaters were therefore not guilty, but instead had self-selected to be innocent. However, I did not collapse these non-cheaters into the innocent condition for three reasons: (1) non-cheaters experienced a completely different situation than did innocent participants (e.g., non-cheaters had the chance to cheat and watched the confederate cheat, whereas innocent participants did not have the chance to cheat and dealt with a grocery list); (2) adding the non-cheaters to the innocent condition would have resulted in highly unbalanced guilty and innocent subsample sizes; and (3) adding non-cheaters as a third level to the culpability

[15] Interpretations for kappa statistics was derived from McHugh's (2012) interpretations of Cohen's kappa: 0–.20 (none); .21–.39 (minimal); .40–.59 (weak); .60–.79 (moderate); .80–.90 (strong); and above .90 (almost perfect).

independent variable would have changed the controlled nature of the experiment by introducing a quasi-experimental component.

Because I excluded non-cheaters, the overall sample size was reduced to 149 participants (and 140 participants for analyses dependent on transcripts of the recorded interrogations). In terms of assessing participants' statement-evidence inconsistencies using a 2 (culpability) x 4 (evidence disclosure) ANOVA, a post-hoc power analysis in G*Power (Faul et al., 2007) revealed that the sample size rendered enough power (.80) to detect a small effect size (Cohen's $f^2 = 0.27$).

Confession Data

Less than one-third of participants signed the confession form: Of 149 participants, 28.2% signed the written confession. Table 5 displays the number of participants who confessed in each of the experimental cells. Guilty participants constituted the majority of confessors, with just over half (52.9%) of guilty participants confessing. Very few of the confessions elicited were false confessions, as a mere 6.3% of innocent participants confessed.

Table 5. Number of Participants Who Confessed by Culpability and Evidence Disclosure Condition

Culpability Condition	Evidence Disclosure Condition				
	Early	Late	SUE-C	SUE-C/E	Total
Guilty	50.0% (9/18)	47.6% (10/21)	57.1% (8/14)	58.8% (10/17)	52.9% (37/70)
Innocent	6.3% (1/16)	14.3% (2/14)	3.4% (1/29)	5.0% (1/20)	6.3% (5/79)
Total	29.4% (10/34)	34.3% (12/35)	20.9% (9/43)	29.7% (11/37)	28.2% (42/149)

Note. Numbers in parentheses represent the number of participants who confessed out of the total number of participants in each experimental cell.

To examine whether there was an effect of culpability and evidence disclosure conditions on confession rates, I carried out a hierarchical logistic regression, with culpability and evidence disclosure entered in the first step and the interaction between these two terms entered in the second step. The overall model was significant at the first step, $\chi^2(4) = 43.30$, $p < .001$, $R^2_{Nagelkerke} = .36$. Table 6 displays all regression results. Only culpability condition emerged as a significant predictor of confession. Specifically, the odds that a guilty participant confessed were 16.64 times greater than the odds that an innocent participant confessed. The odds of confession remained stable across the evidence disclosure conditions ($p = .98$). The second step of the regression, which included the interaction term, also emerged as non-significant, $\chi^2(4) = 2.35$, $p = .50$, $R^2_{Nagelkerke} = .38$.

Table 6. Results for Logistic Regression Analysis Examining the Impact of Culpability and Evidence Disclosure on Confessions

Predictor	*B* (*SE*)	Wald χ^2	*p*	OR	95% CI
Step 1					
Guilty vs. Innocent	2.81 (.53)	28.27	< .001	16.64	[5.90, 46.90]
Evidence Disclosure	--	0.17	.98	--	--
Late vs. Early	.10 (.59)	0.03	.87	1.11	[0.35, 3.53]
SUE-C vs. Early	.05 (.62)	0.01	.93	1.05	[0.31, 3.57]
SUE-C/E vs. Early	.23 (.61)	0.15	.70	1.26	[0.38, 4.16]
Late vs. SUE-C	-.04 (.60)	0.004	.95	0.96	[0.30, 3.15]
SUE-C/E vs. SUE-C	.18 (.62)	0.09	.77	1.20	[0.36, 4.00]
SUE-C/E vs. Late	.22 (.59)	0.14	.71	1.25	[0.39, 3.94]
Step 2					
Culpability * Evidence	--	2.47	.48	--	--

Note. The second evidence disclosure condition listed in each comparison served as the reference category for the regression analysis. The model was run three times, with a different reference category each time, in order to make all comparisons across the evidence disclosure conditions.

Psychological Predictors of Confessions

In line with Houston et al.'s (2014) postulation, I examined whether five different psychological factors (affect, consequences, evidence, guilt, and pressure) predicted confessions. To address this question, I conducted a logistic regression with confession as the outcome variable and the following seven variables as predictors: (1) how guilty participants felt about their behavior during the test; (2) how much evidence participants thought the interrogator had when they were first accused of cheating; (3) how much pressure participants felt to tell the interrogator that they cheated; (4) extent to which participants felt like the interrogator was pushing them to confess; (5) how anxious participants were when the interrogation began; and (6) how severe participants thought the consequences were if they (a) did and (b) did not tell the interrogator what happened. Table 7 displays means and standard deviations for these seven variables. Note that the sample size for this analysis was 145, as some participant responses for the predictor variables were missing.

Table 7. Means and Standard Deviations for Potential Variables Representing Psychological Predictors of Confessions

Item	All Confessions	True Confessions
Anxiety: How anxious were you when the interrogation began?	3.43 (2.09)	3.64 (2.21)
Consequences (telling the researcher): How severe did you think the consequences would be if you *did* tell the interrogator what happened?	2.92 (2.14)	3.36 (2.14)
Consequences (*not* telling the researcher): How severe did you think the consequences would be if you *did not* tell the interrogator what happened?	3.75 (2.32)	4.09 (2.17)
Evidence: How much information do you think the interrogator had supporting the cheating accusation when you were first accused of cheating?	4.73 (2.10)	4.84 (2.05)
Guilt: When you were being questioned, how guilty did you feel about your behavior during the testing session?	2.58 (2.04)	3.38 (2.36)
Pressure: How much pressure did you feel to tell the interrogator that you cheated?	2.73 (2.14)	3.07 (2.21)
Pushy interrogator: I felt like the interrogator was pushing me to confess.	4.72 (2.34)	4.59 (2.44)

Note. Due to missing data, $N = 145$ for the variables listed in this table. Descriptives in the True Confession column are based on a sample of 69 true confessors. All items were measured on a seven-point Likert-type scale (1 = *not at all* and 7 = *extremely*).

The overall model was significant, $\chi^2(7) = 46.59$, $p < .001$, $R^2_{Nagelkerke} = .40$ (Table 8 displays all regression findings). Two variables significantly predicted confession. Specifically, every one-unit increase of feelings of guilt resulted in odds of confession 1.71 times greater. A one-unit increase in participants' perception that the interrogator was pushing them to confess, however, resulted in odds of confession only 0.74 as large. All other predictors were non-significant.

To determine whether these same confession predictors emerged for true confessions specifically, I ran a second logistic regression using only the sample of guilty participants ($n = 69$; see Table 7 for descriptives); I did not run this analysis for innocent

participants because the sample of false confessions was too small (n = 5). Again, the overall model was significant, $\chi^2(7) = 19.73$, $p = .006$, $R^2_{Nagelkerke} = .33$ (Table 8 displays regression findings). Only feelings of guilt emerged as a predictor of true confessions. Specifically, every one-unit increase in guilty feelings resulted in odds of confession 1.64 times greater. No other variable significantly predicted true confessions, although perceptions that the interrogator was pushing for a confession was still trending towards significance (p = .08) and resulted in an odds ratio similar to the interrogator pushiness odds ratio observed in the overall confession model.

Table 8. Results for Logistic Regression Analysis Examining Five Psychological Factors (Affect, Consequences, Evidence, Guilt, and Pressure) on Confessions

Predictor	*B* (*SE*)	Wald χ^2	*p*	OR	95% CI
All confessions model					
Anxiety	-.18 (.15)	1.43	.23	0.84	[0.62, 1.12]
Consequences (telling the researcher)	.19 (.12)	2.44	.12	1.21	[0.95, 1.52]
Consequences (*not* telling the researcher)	-.03(.12)	0.05	.82	0.97	[0.76, 1.24]
Evidence	-.04 (.12)	0.12	.73	0.96	[0.76, 1.21]
Guilt	.54 (.15)	13.38	< .001	1.71	[1.28, 2.28]
Pressure	.15 (.12)	1.71	.19	1.17	[0.93, 1.47]
Pushy interrogator	-.30 (.10)	8.33	.004	0.74	[0.61, 0.91]
True confessions model					
Anxiety	-.17 (.19)	0.80	.37	0.84	[0.58, 1.23]
Consequences (telling the researcher)	.16 (.16)	1.02	.31	1.18	[0.86, 1.62]
Consequences (*not* telling the researcher)	-.002 (.17)	< 0.001	.99	1.00	[0.71, 1.40]
Evidence	-.06 (.15)	0.13	.72	0.95	[0.70, 1.28]
Guilt	.50 (.21)	5.56	.02	1.64	[1.09, 2.49]
Pressure	-.07 (.19)	0.14	.70	0.93	[0.65, 1.34]
Pushy interrogator	-.24 (.14)	3.09	.08	0.79	[0.60, 1.03]

Note. The outcome confession variable was coded as such: 0 = no confession; 1 = confession.

Statement-Evidence Inconsistencies

Table 9 displays the number of participants who made fully consistent (participant's response fully consistent with the evidence), not fully consistent (participant's response *may* be consistent with the evidence), and inconsistent (participant's response not consistent with the evidence) statements in response to the evidence-specific questions. Overall, the number of participants who made a not fully consistent statement was low: A mere 5.0% of participants made a not fully consistent statement in response to Specific Question #1, 6.8% in response to Specific Question #2, 8.2% in response to Specific Question #3, and 9.8% in response to Specific Question #4. Considering these low percentages, as well as the definition of a not fully consistent statement, I chose to collapse not fully consistent statements into inconsistent statements for the purposes of the present analyses.

Table 9. Number of Participants Who Made Consistent, Not Fully Consistent, and Inconsistent Statements, By Culpability Condition and Evidence Disclosure Condition

Specific Question	Culpability Condition	Evidence Disclosure Condition	Level of Statement-Evidence Consistency		
			Consistent	Not Fully Consistent	Inconsistent
Question #1	Guilty	Early	8	1	8
		Late	1	1	16
		SUE-C	4	0	9
		SUE-C/E	10	1	6
		Total	23	3	39
	Innocent	Early	12	0	3
		Late	5	1	8
		SUE-C	14	1	12
		SUE-C/E	11	2	5
		Total	42	4	28
	Total	Early	20	1	11
		Late	6	2	24
		SUE-C	18	1	21
		SUE-C/E	21	3	11
		Total	65	7	67
Question #2	Guilty	Early	8	1	8
		Late	8	2	7
		SUE-C	4	0	10
		SUE-C/E	10	2	4

Specific Question	Culpability Condition	Evidence Disclosure Condition	Level of Statement-Evidence Consistency		
			Consistent	Not Fully Consistent	Inconsistent
		Total	30	5	29
	Innocent	Early	5	1	8
		Late	10	0	3
		SUE-C	17	0	8
		SUE-C/E	14	3	0
		Total	46	4	19
	Total	Early	13	2	16
		Late	18	2	10
		SUE-C	21	0	18
		SUE-C/E	24	5	4
		Total	76	9	48
Question #3	Guilty	Early	8	1	7
		Late	7	1	9
		SUE-C	8	0	6
		SUE-C/E	11	1	4
		Total	34	3	26
	Innocent	Early	5	3	7
		Late	4	1	9
		SUE-C	13	3	9
		SUE-C/E	15	1	1
		Total	37	8	26
	Total	Early	13	4	14
		Late	11	2	18
		SUE-C	21	3	15
		SUE-C/E	26	2	5
		Total	71	11	52
Question #4	Guilty	Early	10	1	6
		Late	6	2	8
		SUE-C	6	1	7
		SUE-C/E	6	5	5
		Total	28	9	26
	Innocent	Early	8	1	6
		Late	7	2	3
		SUE-C	18	0	5
		SUE-C/E	8	0	2
		Total	41	3	16
	Total	Early	18	2	12
		Late	13	4	11
		SUE-C	24	1	12
		SUE-C/E	14	5	7
		Total	69	12	42

Creation of Inconsistency Variables

I analyzed statement-evidence inconsistencies using two different inconsistency scores: raw inconsistency scores and mean composite inconsistency scores. The raw

inconsistency score represented the total number of evidence pieces that the participant's statement was inconsistent with. As there were four pieces of evidence (with one specific question asked per evidence piece), the maximum score was 4, with higher scores indicating more statement-evidence inconsistencies. However, as previously mentioned, not all 140 participants responded to all four specific questions (i.e., some responded to less than four specific questions; $n = 23$). Thus, it was not appropriate to use a raw inconsistency score for all participants, as has been done in past SUE research (e.g., Tekin et al., 2016).

To be able to compare statement-evidence inconsistencies across all participants despite the number of questions they responded to, I calculated a mean composite inconsistency score for each participant. This composite score represented the proportion of inconsistent statements across each of the statements elicited from the participant in response to the specific questions that they were asked (i.e., the number of specific questions asked was the denominator). Inconsistency composite scores therefore ranged from 0 to 1, with higher scores indicating a higher rate of statement-evidence inconsistencies. Below, I present the results of the analyses for the raw inconsistency score, followed by the results of the analyses for the composite inconsistency score.

Analysis with the Raw Statement-Evidence Inconsistency Score

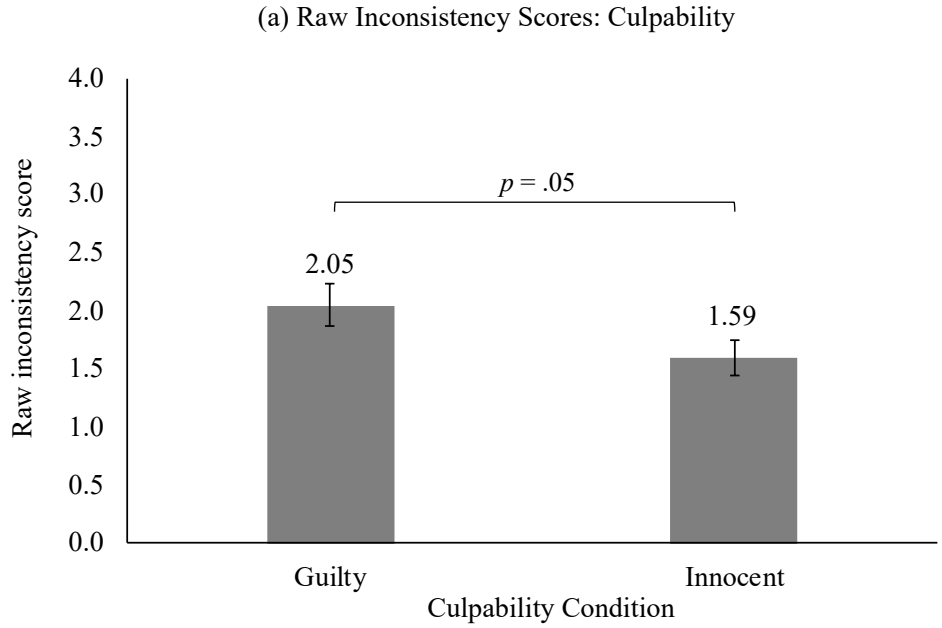
To determine whether there was an effect of culpability condition and evidence disclosure condition on raw statement-evidence inconsistency scores, I ran a 2 (culpability) x 4 (evidence disclosure) between-subjects ANOVA. Figures 4a through 4c present mean raw inconsistency scores across conditions. A main effect for evidence disclosure condition emerged, $F(3, 109) = 3.33$, $p = .02$, $\eta_p^2 = 0.08$, 90% CI [0.01, 0.16].

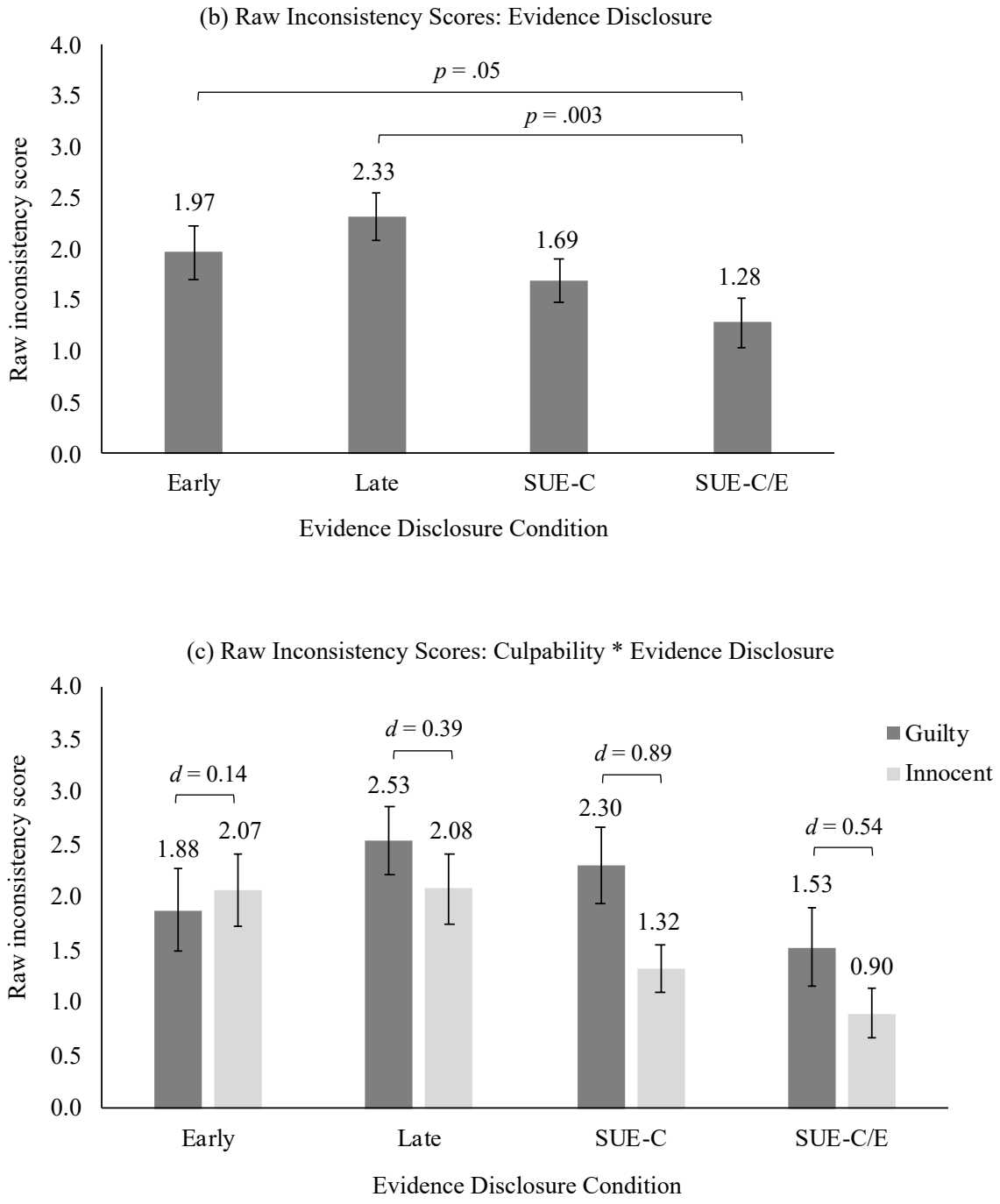
LSD post-hoc comparisons revealed that participants in the Late condition made more statement-evidence inconsistencies compared to participants in the SUE-C/E condition (p = .003). Similarly, participants in the Early condition made more statement-evidence inconsistencies than did participants in the SUE-C/E condition (p = .05). No other comparisons were significant (all ps > .22). A main effect of culpability also emerged, $F(1, 109) = 3.91$, $p = .05$, $\eta_p^2 = 0.04$, 90% CI [0.01, 0.17], whereby guilty participants made more statement-evidence inconsistencies than did innocent participants.

There was no significant interaction between culpability and evidence disclosure technique, $F(3, 109) = 1.20$, $p = .31$, $\eta_p^2 = 0.03$, 90% CI [0.00, 0.08]. However, because I hypothesized a priori that the differences in statement-evidence inconsistencies between guilty and innocent participants would be greater in the three SUE conditions than in the early disclosure condition, I carried out four independent-samples t-tests (Bonferroni correction: critical p = .013) to address this hypothesis. The results revealed that the significance-level associated with the difference in the raw number of statement-evidence inconsistencies between guilty and innocent participants in the SUE-C condition nearly exceeded the critical cut-off, $t(33) = 2.46$, $p = .02$, $d = 0.89$, 95% CI [0.52, 1.26]. Specifically, there was a trend such that guilty participants questioned with SUE-C made more evidence-inconsistent statements than did innocent participants in the same interrogation condition. Statement-evidence inconsistency comparisons between guilty and innocent participants in the other three evidence disclosure conditions were non-significant: Early, $t(28) = .38$, $p = .71$, $d = 0.14$, 95% CI [-0.35, 0.64]; Late, $t(25) = .96$, $p = .35$, $d = 0.39$, 95% CI [-0.05, 0.83]; and SUE-C/E, $t(21.82) = 1.43$, $p = .17$, $d = 0.54$, 95% CI [0.08, 1.00]. However, it is notable that a medium effect size emerged in the

SUE-C/E condition, especially compared to the small effect size rendered in the Early condition. Also notable is that when descriptively reviewing the results, I found that statement-evidence inconsistency means were higher for guilty than innocent participants in all three SUE conditions but not in the Early disclosure condition (statement-evidence inconsistencies were descriptively higher for innocent than guilty participants in the Early condition).

Figure 4. Mean Raw Inconsistency Scores By (a) Culpability Condition, (b) Evidence Disclosure Condition, (c) Culpability by Evidence Disclosure





Note. Error bars represent standard errors. Effect sizes are displayed in Figure 4c.

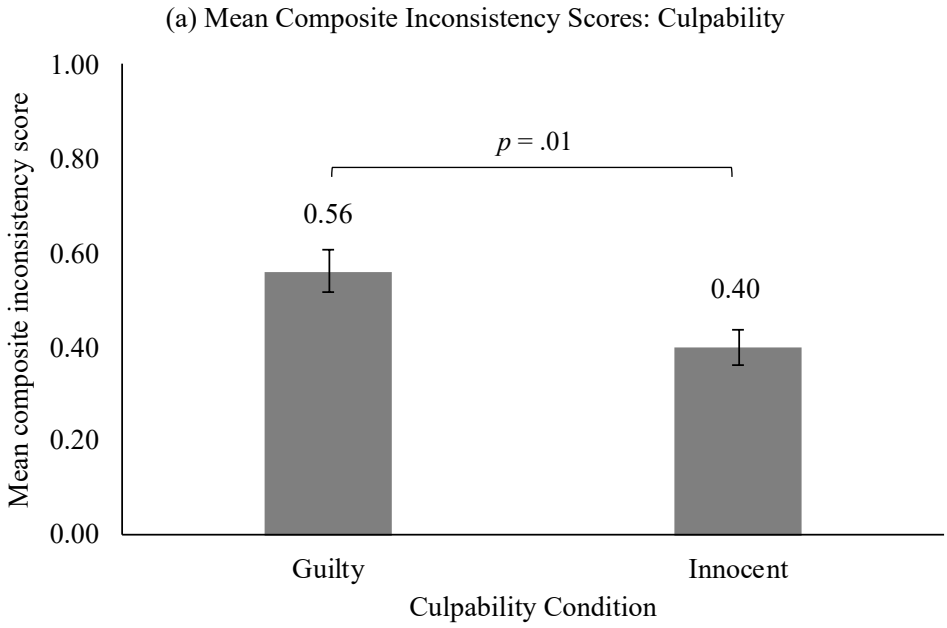
Analysis with the Mean Composite Statement-Evidence Inconsistency Score

I conducted a 2 (culpability) x 4 (evidence disclosure) between-subjects ANOVA to assess the effect of my independent variables on the mean composite inconsistency score. Figures 5a through 5c display composite scores. The pattern of results was similar to the results derived from the raw inconsistency score analysis. Specifically, a main effect emerged for culpability, $F(1, 132) = 7.25$, $p = .01$, $\eta_p^2 = 0.05$, 90% CI [0.01, 0.12], as well as for evidence disclosure, $F(3, 132) = 3.74$, $p = .01$, $\eta_p^2 = 0.08$, 90% CI [0.01, 0.14]. Concerning evidence disclosure techniques, just as with raw scores, LSD post-hoc comparisons revealed that composite inconsistency scores were significantly higher in the Late condition ($p = .001$) and the Early condition ($p = .05$), as compared to the SUE-C/E condition; no other comparisons were significant (all ps > .06). As with raw scores, guilty participants' composite inconsistency scores were significantly higher than those for innocent participants.

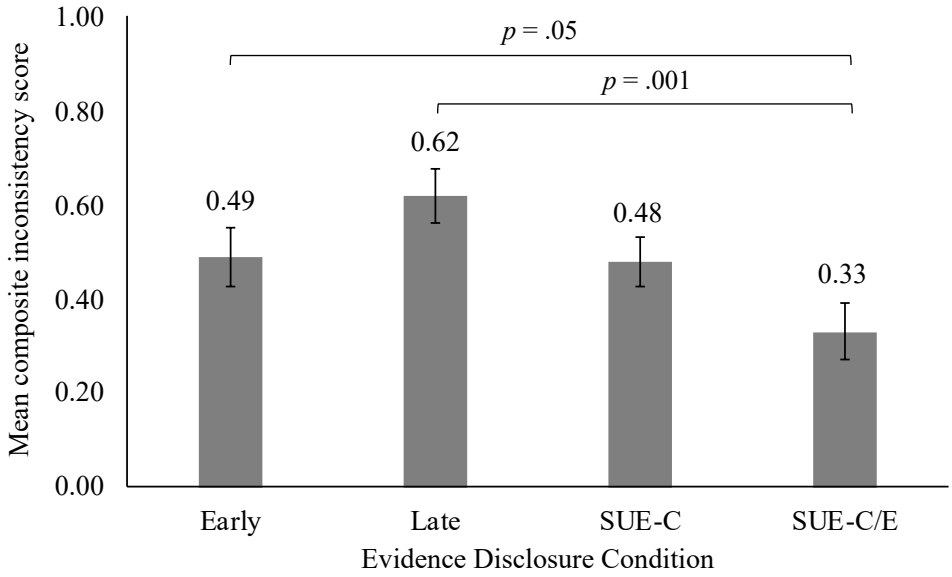
Additionally, there was no significant interaction between culpability and evidence disclosure on composite inconsistency scores, $F(3, 132) = 0.68$, $p = .57$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.04]. As with the raw scores, I examined the extent to which guilty and innocent participants differed in each evidence disclosure condition in terms of their mean composite inconsistency scores in order to address my a priori hypothesis. Independent-samples t-tests (Bonferroni correction: critical $p = .013$) indicated a trend towards significance in the SUE-C/E condition, as guilty participants demonstrated a higher mean composite inconsistency score than did innocent participants, $t(27.57) = 2.08$, $p = .047$, $d = 0.73$, 95% CI [0.63, 0.84]. Significant differences did not emerge in the other three evidence disclosure conditions: Early, $t(30) = .13$, $p = .90$, $d = 0.05$, 95%

CI [-0.08, 0.17]; Late, $t(30) = 1.54$, $p = .14$, $d = 0.57$, 95% CI [0.46, 0.67]; and SUE-C, $t(39) = 1.82$, $p = .08$, $d = 0.61$, 95% CI [0.52, 0.71]. Notable, however, are the medium effect sizes demonstrated in the Late and SUE-C conditions, especially as compared to the effect size of nearly zero in the Early condition.

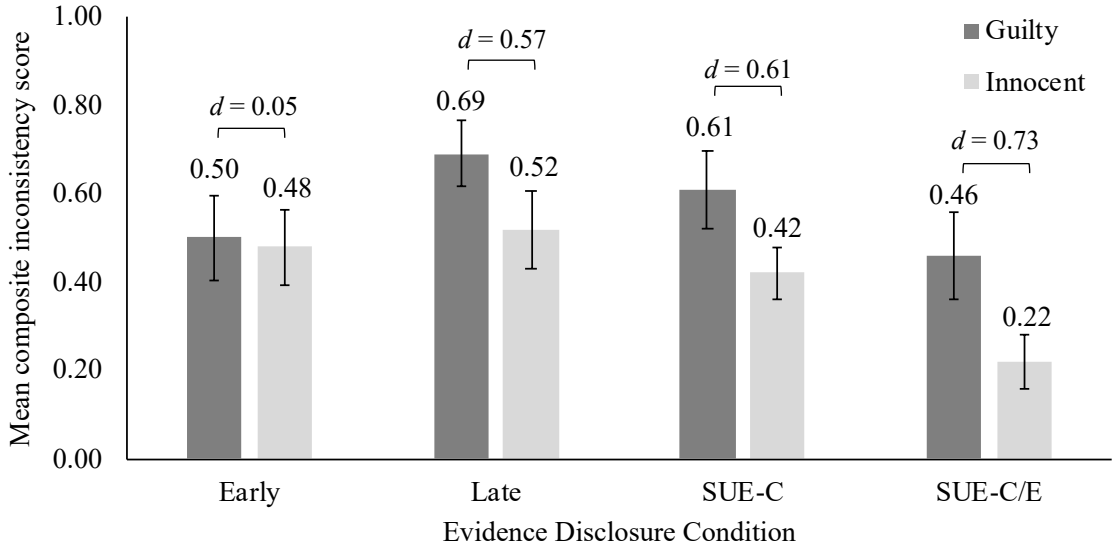
Figure 5. Mean Composite Inconsistency Scores By (a) Culpability Condition, (b) Evidence Disclosure Condition, (c) Culpability by Evidence Disclosure



(b) Mean Composite Inconsistency Scores: Evidence Disclosure



(c) Mean Composite Inconsistency Scores: Culpability * Evidence Disclosure



Note. Error bars represent standard errors. Effect sizes are displayed in Figure 5c.

Participants' Subjective Ratings

Perceptions of the Evidence

To determine whether participants' perceptions of the amount of evidence the interrogator had against them changed from the start to the end of the interrogation, and whether this change varied by culpability and evidence disclosure conditions, I carried out a 2 (interrogation start, interrogation end) x 2 (culpability) x 4 (evidence disclosure) mixed-design ANOVA, with interrogation timing varying within-subjects (see Table 10 for descriptives; note that both of these ratings were provided on the debriefing questionnaire, they were not provided at the start and end of the interrogation respectively). The results revealed a main effect for interrogation timing, Greenhouse-Geisser $F(1, 138) = 4.39$, $p = .04$, $\eta_p^2 = 0.03$, 90% CI [0.001, 0.09]. Not surprisingly, participants believed that the interrogator had more incriminating evidence by the time the interrogation ended compared to when the interrogation began. There was no main effect of culpability, $F(1, 138) = 0.93$, $p = .34$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.05], nor of evidence disclosure condition, $F(3, 138) = 1.71$, $p = .17$, $\eta_p^2 = 0.04$, 90% CI [0.00, 0.08]. Additionally, there was no significant interaction between interrogation timing and culpability condition, Greenhouse-Geisser $F(1, 138) = 0.78$, $p = .38$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.04], between interrogation timing and evidence disclosure condition, Greenhouse-Geisser $F(3, 138) = 0.59$, $p = .62$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.04], between culpability and evidence disclosure conditions, $F(3, 138) = 1.36$, $p = .26$, $\eta_p^2 = 0.03$, 90% CI [0.00, 0.07], nor between all three variables, Greenhouse-Geisser $F(3, 138) = 0.02$, $p = .997$, $\eta_p^2 < 0.01$.

I also examined whether participants' perceptions of the strength of the evidence against them changed from the start to the end of the interrogation, and whether this varied by culpability and evidence disclosure conditions (see Table 10 for descriptives; note that both of these ratings were provided on the debriefing questionnaire, they were not provided at the start and end of the interrogation respectively). A 2 (interrogation start, interrogation end) x 2 (culpability) x 4 (evidence disclosure) mixed-design ANOVA, with interrogation timing serving as the within-subjects variable, showed that guilty participants perceived the evidence against them to be stronger than did innocent participants, $F(1, 138) = 17.50$, $p < .001$, $\eta_p^2 = 0.11$, 90% CI [0.04, 0.20]. Perceptions of evidence strength did not significantly vary from the start to end of the interrogation, Greenhouse-Geisser $F(1, 138) < 0.01$, $p = .999$, $\eta_p^2 < 0.01$, nor by evidence disclosure condition, $F(3, 138) = 0.89$, $p = .45$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.05]. There was no significant interaction between interrogation timing and culpability condition, Greenhouse-Geisser $F(1, 138) = 0.73$, $p = .39$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.04], between interrogation timing and evidence disclosure condition, Greenhouse-Geisser $F(3, 138) = 0.99$, $p = .40$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.06], between culpability and evidence disclosure conditions, $F(3, 138) = 1.76$, $p = .16$, $\eta_p^2 = 0.04$, 90% CI [0.00, 0.08], nor between all three variables, Greenhouse-Geisser $F(3, 138) = 0.22$, $p = .89$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.01].

Table 10. Means and Standard Deviations for Participants' Perceptions of the Amount and Strength of Incriminating Evidence at the Start and End of the Interrogation

Variable	Culpability Condition	Evidence Disclosure Condition	Interrogation Time	
			Start	End
Amount of evidence	Guilty	Early	4.72 (2.24)	5.61 (1.46)
		Late	5.05 (2.09)	5.24 (2.19)
		SUE-C	4.00 (2.00)	4.36 (1.95)
		SUE-C/E	5.44 (1.71)	6.13 (1.86)
		Total	4.84 (2.05)	5.36 (1.95)
	Innocent	Early	5.06 (1.65)	5.56 (1.63)
		Late	4.54 (2.11)	4.46 (2.33)
		SUE-C	4.54 (2.13)	4.68 (2.34)
		SUE-C/E	4.50 (2.61)	4.80 (2.62)
		Total	4.84 (2.05)	4.86 (2.28)
	Total	Early	4.88 (1.97)	5.59 (1.52)
		Late	4.85 (2.08)	4.94 (2.24)
		SUE-C	4.36 (2.08)	4.57 (2.20)
		SUE-C/E	4.92 (2.27)	5.39 (2.38)
		Total	4.73 (2.10)	5.10 (2.14)
Strength of evidence	Guilty	Early	5.11 (1.78)	5.28 (1.84)
		Late	4.19 (2.23)	4.38 (2.36)
		SUE-C	3.50 (2.07)	3.36 (1.87)
		SUE-C/E	4.13 (2.73)	4.37 (2.53)
		Total	4.28 (2.24)	4.41 (2.23)
	Innocent	Early	2.56 (2.31)	2.25 (1.77)
		Late	3.00 (2.38)	3.00 (2.45)
		SUE-C	3.00 (2.13)	2.50 (1.90)
		SUE-C/E	3.00 (2.34)	3.35 (2.46)
		Total	2.91 (2.23)	2.75 (2.13)
	Total	Early	3.91 (2.43)	3.85 (2.35)
		Late	3.74 (2.29)	3.85 (2.45)
		SUE-C	3.17 (2.09)	2.79 (1.91)
		SUE-C/E	3.50 (2.55)	3.81 (2.51)
		Total	3.55 (2.33)	3.53 (2.32)

Note. N = 146 for amount of evidence and strength of evidence analyses. Standard deviations are displayed in parentheses. Amount of evidence (1 = *the interrogator knew nothing about what happened during the testing session* and 7 = *the interrogator knew about everything that happened during the testing session*) and strength of evidence (1 = *not strong at all* and 7 = *extremely strong*) were measured on seven-point Likert-type scales.

Perceptions of Strategies

I ran three 2 (culpability) x 4 (evidence disclosure) between-subjects ANOVAs (critical $p = .017$)[16] to assess three variables concerning participants' perceptions of the extent to which they aimed to be: (1) withholding; (2) in denial; and (3) forthcoming. Table 11 displays means and standard deviations for these three outcome variables. In terms of being withholding, the only significant finding that emerged was for the main effect for culpability, $F(1, 138) = 10.76$, $p = .001$, $\eta_p^2 = 0.07$, 90% CI [0.02, 0.15]; there was no significant main effect of evidence disclosure condition, $F(3, 138) = 0.54$, $p = .66$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.04], nor a significant interaction, $F(3, 138) = 0.99$, $p = .40$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.06]. Similarly, the only significant finding for the strategy of being in denial was the main effect for culpability, $F(1, 138) = 7.42$, $p = .007$, $\eta_p^2 = 0.05$, 90% CI [0.01, 0.12]; there was no significant main effect of evidence disclosure condition, $F(3, 138) = 0.28$, $p = .84$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.02], nor a significant interaction, $F(3, 138) = 1.04$, $p = .38$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.06]. Lastly, as with the other outcomes, the only main effect to emerge for the forthcomingness variable was that for culpability, $F(1, 138) = 17.21$, $p < .001$, $\eta_p^2 = 0.11$, 90% CI [0.04, 0.20]. There was neither a main effect for evidence disclosure condition, $F(3, 138) = 0.47$, $p = .70$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.03], nor a significant interaction, $F(3, 138) = 0.12$, $p = .95$, $\eta_p^2 < 0.01$, 90% CI [0.00, 0.001]. Overall, the results demonstrated that guilty participants' ratings for aiming to be withholding and in denial were higher than innocent participants'

[16] I did not use a MANOVA to examine these three variables, as the strategy of being forthcoming is conceptually different than those strategies of being withholding and being in denial.

ratings, while innocent participants reported higher forthcomingness scores than did guilty participants.

Table 11. Means and Standard Deviations for Participants' Ratings of the Extent to Which They Aimed to Be Withholding, In Denial, and Forthcoming

Variable	Evidence Disclosure Condition	Culpability Condition		
		Guilty	Innocent	Total
Withholding	Early	2.17 (1.79)	1.19 (0.54)	1.71 (1.43)
	Late	2.14 (1.85)	1.54 (0.78)	1.91 (1.55)
	SUE-C	2.29 (1.77)	1.96 (1.53)	2.07 (1.60)
	SUE-C/E	2.69 (2.18)	1.20 (0.70)	1.86 (1.69)
	Total	2.30 (1.87)	1.53 (1.11)	1.90 (1.56)
In Denial	Early	3.17 (2.26)	3.63 (2.90)	3.28 (2.55)
	Late	2.86 (2.29)	4.31 (2.78)	3.41 (2.55)
	SUE-C	1.93 (1.33)	4.18 (2.75)	3.43 (2.58)
	SUE-C/E	3.00 (2.37)	3.50 (2.84)	3.28 (2.61)
	Total	2.78 (2.14)	3.91 (2.77)	3.38 (2.55)
Forthcoming	Early	5.61 (1.58)	6.63 (1.26)	6.09 (1.51)
	Late	5.05 (2.36)	6.54 (0.97)	5.62 (2.06)
	SUE-C	5.14 (1.75)	6.29 (1.21)	5.90 (1.50)
	SUE-C/E	5.12 (2.13)	6.25 (1.77)	5.75 (1.99)
	Total	5.23 (1.97)	6.39 (1.34)	5.84 (1.76)

Note. $N = 146$ for analyses of participants' reports of the extent to which they aimed to be withholding, in denial, and forthcoming. Standard deviations are presented in parentheses. Each of the three outcomes were measured on a seven-point Likert-type scale (1 = *not at all* and 7 = *extremely*).

Discussion

Confession Findings

The overall confession rate in the present study was low (28.2%). This low confession rate does not come as a surprise, as the majority of the evidence disclosure conditions were characterized by SUE and the primary purpose of the SUE technique is not to elicit confessions, but rather to elicit information and diagnostic cues to deception (e.g., Granhag & Hartwig, 2015). Indeed, Jordan et al. (2012) elicited very few confessions overall (11.1%). Yet, it is important to note that Jordan et al. (2012)

employed a mock crime paradigm to test SUE, whereas the present study employed a psychologically realistic paradigm. The present participants believed that if they confessed, they faced the possibility of a $50–500 fine; thus, even with such a seemingly severe consequence in place, it is notable that the present study's confession rate was at nearly 30%. Contrarily, Sellers and Kebbell (2009) achieved a high overall confession rate (65.3%), but the researchers had used a mock crime paradigm and their entire sample was guilty (they did not look at false confession rates). Comparably, the confession rate for guilty participants in the current study was over 50%.

Examination of potential underlying psychological predictors of confession, as proposed by Houston et al. (2014), offer some insight on why participants in the present study confessed. As demonstrated by the present findings, the strongest predictor for confession was participants' feeling of guilt: The more guilty participants reportedly felt, the more likely they were to confess (a finding consistent with Houston et al., who examined student samples). Notably, perceptions of the evidence did not significantly predict confessions. One potential reason for evidence perceptions' lack of predictive power could be attributed to the perceived strength of the evidence used in the present paradigm. Indeed, participants' ratings for evidence strength were low (all means fell below the mid-point on the scale). It is possible that if the evidence presented to participants was stronger (e.g., showing participants the confederate's test with identical answers to participants' tests), then there may have been an increase in overall confession rate. Indeed, Sellers and Kebbell (2009) found that evidence strength significantly predicted confessions. Future studies should therefore consider examining the impact of evidence strength on confession decisions in the context of a psychologically realistic

paradigm, as this can contribute to a better understanding of how perceptions of the evidence can impact confession decisions.

Despite the present study's low confession rate, the pattern of true versus false confessions specifically was similar to patterns seen in past research. For example, when no coercive interrogation tactics were used, Russano et al. (2005) achieved a 46% true confession rate and a 6% false confession rate—rates similar to the 52.9% and 6.3%, respectively, achieved across all evidence disclosure conditions in the present study. Furthermore, the present results suggest that strategic evidence disclosure techniques, at least when used in isolation, do not necessarily decrease the risk of false confessions beyond the rates found when early disclosure of evidence is used. This finding could be attributed to the fact that early disclosure of evidence itself did not result in a high number of false confessions (i.e., there was a floor effect for false confessions).

The lack of false confessions in the early disclosure condition is an important finding to note, as the interrogation literature has traditionally tied early evidence disclosure to the risk of false confession (e.g., Leo, 1996). Early evidence disclosure is, however, often incorporated with other tactics to form interrogation methods that, as a whole, have been demonstrated to increase the risk of false confession. For example, the Reid technique promotes the use of early evidence disclosure (Inbau, Reid, Buckley, & Jayne, 2013), along with other undiagnostic tactics such as false evidence ploys and implicit promises of leniency. Considering the present findings, however, it seems that early evidence disclosure alone is not necessarily predictive of false confessions.

Moreover, the majority of findings in the SUE literature attest to SUE's efficacy in eliciting reliable, diagnostic cues to deception (e.g., statement-evidence

inconsistencies; Hartwig et al., 2005; Hartwig et al., 2014). Yet, only two other studies (Jordan et al., 2012; Sellers & Kebbell, 2009) examined confession rates. These studies were limited to mock crime paradigms with low confession rates (Jordan et al., 2012) or true confessions exclusively (Sellers & Kebbell, 2009). The present study addressed this gap, as the current confession findings can assure law enforcement that the implementation of the SUE technique, whether it be late disclosure, SUE-C, or SUE-C/E, does not negatively impact true confession rates, nor does it increase false confession rates. As such, the present findings can help researchers better promote the SUE technique to law enforcement. Overall, the existing and current research suggest that the implementation of SUE (in isolation) does not directly impact confession decisions.

Statement-Evidence Inconsistencies

As expected, guilty participants demonstrated more statement-evidence inconsistencies than innocent participants (although both culpability groups were relatively high in terms of statement-evidence inconsistencies). This finding is in line with the postulation that guilty suspects are less forthcoming than innocent suspects. Past studies have found that guilty suspects typically engage in denial and avoidant behaviors to mitigate interrogators' perceptions of their guilt, whereas innocent suspects aim to tell the truth (Granhag & Hartwig, 2008; Granhag & Luke, 2018). My results were consistent with this: Guilty participants in the present study reported being more withholding and in denial than did innocent participants, and innocent participants reported being more forthcoming than did guilty participants. The present findings therefore contribute to the robust literature indicating that guilty suspects usually engage in counter-interrogation

strategies characterized by avoidance and innocent suspects are generally more forthcoming.

Unexpectedly, main effect analyses of evidence disclosure technique on statement-evidence inconsistencies showed that the three SUE conditions did not result in significantly more statement-evidence inconsistencies than did the early disclosure condition. In fact, SUE-C/E demonstrated the fewest statement-evidence inconsistencies across all of the evidence disclosure conditions, with the comparison to the early and late disclosure conditions being statistically-significant. This latter finding concerning the comparison between the SUE-C/E and late disclosure conditions is especially relevant, as the past studies examining SUE-C/E and SUE-C (Tekin et al., 2015; 2016) did not compare these disclosure methods to the late disclosure method, which is the original and most studied of the SUE variations. When revisiting past SUE literature, we typically see a main effect of evidence disclosure on statement-evidence inconsistencies, by which the late disclosure condition renders more statement-evidence inconsistencies than does the early evidence disclosure condition (e.g., Jordan et al., 2012). Thus, in terms of eliciting statement-evidence inconsistencies, the late disclosure condition appears to excel above other SUE variations.

However, just because SUE-C/E did not elicit many statement-evidence inconsistencies overall is no cause for its dismissal as a valid and effective questioning strategy. One potential reason for the lower number of statement-evidence inconsistencies seen in SUE-C/E, and especially so for guilty participants, could be due to the manner in which I coded participants' statements. Specifically, participants' *final* statement to each specific question was coded for inconsistency in order to enable

examination of the effect of SUE-C (participants can clarify their initial responses) and SUE-C/E (participants can clarify their initial responses as well as explain any evidence-inconsistencies) on participants' overall response to each specific question (i.e., if I coded only initial responses, the data would not reflect the effect of the core component of SUE-C/E, being the explanation, on participants' statements). I therefore postulate that participants questioned with SUE-C/E may have shifted their counter-interrogation strategy within their responses to a single specific question from avoidance to forthcomingness after being confronted with any inconsistencies. Essentially, participants used their opportunity for explanation to be forthcoming rather than to try to explain away any inconsistencies—an opportunity that participants in the SUE-C and late disclosure conditions did not have.

Furthermore, the differences in timing of evidence disclosure between the SUE-C and C/E conditions and the late disclosure condition may contribute to the differences in statement-evidence inconsistencies. The interrogator presented participants with the evidence incrementally in the SUE-C and C/E conditions, along with corrective feedback (i.e., pointing out statement-evidence inconsistencies), whereas the interrogator presented all evidence at once in the late disclosure condition with no corrective feedback. By presenting the evidence incrementally along with feedback, I postulate that the interrogator employing SUE-C and SUE-C/E essentially gave participants the chance to shift their strategy to one of being more forthcoming once they recognized that their avoidant strategy was not working (i.e., the interrogator was able to keep pointing out the inconsistencies and participants were therefore able to be more dynamic in their statement strategies from one specific question to the next). Contrarily, in the late

disclosure condition, participants presumably maintained their avoidant strategy throughout the questioning, as they were not aware of any of the evidence that the interrogator held and did not receive corrective feedback throughout the questioning session.

When considering the difference in statement-evidence inconsistencies between evidence disclosure conditions depending on participants' guilt status, the results did partially support my hypothesis. Although *p*-values fell below the Bonferroni-corrected significance level, I found that the effect sizes representing the difference between guilty and innocent participants' statement-evidence inconsistencies ranged from medium to large in the SUE-C and SUE-C/E conditions, regardless of how I measured statement-evidence inconsistencies (raw or mean composite scores). Unexpectedly, the effect size that emerged in the late disclosure condition ranged from small (raw score) to medium (mean composite score). Guilty suspects may therefore be more distinguishable from innocent suspects on the basis of statement-evidence inconsistencies when questioned with SUE-C or SUE-C/E than with the late disclosure method. Yet, all three SUE methods appear to be more effective than early disclosure, which resulted in essentially no difference between guilty and innocent participants' statement-evidence inconsistencies (as measured by the composite score) or even more inconsistencies by innocent than guilty participants (as measured by the raw score).

Although Study 1 findings are interesting, statement-evidence inconsistencies do not necessarily represent a practical outcome measure for law enforcement officers who are considering the benefits of employing SUE tactics, as they may not have the resources to code their interrogations for statement-evidence inconsistencies. As such, I

pursued Study 2 to determine what implications the differences in participant statements, which derived from the SUE technique employed, had for observers' ability to accurately judge participants' culpability.

III. STUDY 2

Study 2 examined observers' ability to accurately determine whether interviewees questioned with different evidence disclosure techniques were guilty or innocent of having committed a transgression. Specifically, student participants viewed a random selection of Study 1 videotaped interrogations that employed one of the four types of evidence disclosure methods (early disclosure, late disclosure, SUE-C, or SUE-C/E). Study 2's research questions and associated hypotheses were as follows:

(1) *Can SUE techniques help observers better distinguish between guilty and innocent suspects?* I hypothesized that participants would demonstrate higher culpability judgment accuracy rates[17] when viewing interviewees being questioned with one of the three SUE methods, compared to being questioned with the early disclosure method. This prediction fell in line with past research (e.g., Hartwig et al., 2005).

(2) *Can implementing the SUE technique protect the credibility of innocent suspects who ultimately confess?* I hypothesized that participants would be less likely to misidentify innocent interviewees who were questioned with a SUE technique as guilty even after they confessed, as compared to those

[17] I did not measure participants' deception detection accuracy per se. Specifically, participants in the present study did not make decisions about whether interviewees were lying or telling the truth, but rather about whether interviewees had or had not cheated on the test. I therefore used the term "culpability judgment" in the present study. I do, however, draw parallels between the present findings and deception detection findings in the SUE literature, as these concepts are similar.

questioned with early disclosure. Such a pattern of results would presumably be due to the greater number of evidence-consistent statements that innocent suspects offer when questioned with SUE techniques.

Method

Participants

I recruited participants via the psychology department's participant pool and awarded participants one research participation credit for partaking in the study. A total of 342 participants consented to and completed the study. Of participants who consented, 46 participants either did not attempt or failed the sound test[18] and were thus excluded from the sample. An additional 47 participants quit the study prior to viewing the interrogation videos and one participant quit before viewing the second video; thus, I excluded these 48 participants.

From the remaining 248 participants, I excluded additional participants based on the following criteria (exclusions were made in this order): missing one or both of two attention check questions ("For this question, would you please select the number seven?" and "For this question, can you select the number that two plus two equals?;" n = 11); missing the interviewee gender attention check question (n = 6); and taking longer than two standard deviations above the mean study duration to complete the study (i.e., longer than 946.17 minutes; n = 5).

Participants responded to two additional attention check questions related to the content of what interviewees said during the interrogation ("Did the Participant in the

[18] The sound test was implemented at the start of the study to ensure that participants could hear and understand verbal content. The Procedure section includes details concerning the sound test.

video mention writing anything down while looking at the phone?" and "Did the Participant in the video explicitly mention that the *other participant* had a phone during the testing session?"), with each of these questions being posed twice (once after the guilty video and once after the innocent video). Each correct response earned one point. The maximum score participants could therefore receive was four. The proportion of participants who received each possible score on these manipulation check questions (i.e., the total number of correct responses provided by participants) is as follows: score of 0 (4.3%); score of 1 (10.8%); score of 2 (29.0%); score of 3 (36.4%); and score of 4 (19.5%). Thus, a large proportion of participants did not correctly answer all interrogation content manipulation questions. Rather than take an extremely conservative route and exclude all participants who did not achieve a perfect score, which would have resulted in a total sample size of 45 participants, I only excluded participants who missed more than one interrogation content manipulation check question (n = 102).

One-hundred and twenty-nine participants remained in the sample after I made the aforementioned exclusions.[19] A post-hoc power analysis in G*Power (Faul et al., 2007) revealed that this sample size rendered enough power (.80) to find a medium effect (Cohen's $f^2 = 0.34$) using a one-way ANOVA. The majority of participants were female (85.3%) and Hispanic/Latinx (71.9%), with the remaining participants identifying as Black (14.8%), White (7.8%), Asian (1.6%), and multiracial/other (3.9%). Participants were, on average, 23-years-old (SD = 6), with ages ranging from 18 to 54-years-old.

[19] One participant did not respond to demographic questions; 128 participants total responded.

Design

Study 2 was characterized by a 2 (*interviewee culpability:* guilty, innocent) x 4 (*type of evidence disclosure:* Early, Late, SUE-C, SUE-C/E) mixed design, with interviewee culpability varying within-participants and type of evidence disclosure varying between-participants. Cell sizes for the between-participants variable were as follows: Early (n = 24); Late (n = 34); SUE-C (n = 36); and SUE-C/E (n = 35).

Materials

Interrogation Videos

The interrogation videos used in the present study were randomly selected from the sample of videos collected in Study 1. To ensure stimulus sampling, I selected three videos for each experimental cell; however, regarding the Innocent-Early disclosure cell, I selected only two videos because many of the videos pertaining to this cell did not have Study 1 participant consent for use. Thus, the video sample included a total of 23 interrogation videos. I assessed all 23 included videos to ensure that: (1) there was a clean culpability manipulation (i.e., guilty interviewees had used the phone to cheat and innocent interviewees had used the phone to write down the grocery list); (2) all four evidence pieces had occurred (see Table 3 in Study 1); (3) the interrogator had correctly followed the interrogation script related to the evidence disclosure condition; and (4) interviewees in the videos did not explicitly admit to copying answers directly off of the phone.

All videos were characterized by an equal-focused camera view (i.e., participants viewed the profile of both the interrogator and interviewee). The videos began when the interrogator introduced the problem ("…after looking into your tests more, I have a

reasonable suspicion that you and the other participant cheated on the test…") and ended after interviewees responded to the interrogator's final question. Notably, participants did not view the confession elicitation portion of interviewees' interrogations. The videos lasted on average 5.97 minutes (*SD* = 1.19).

Post-Video Questionnaire

The post-video questionnaire followed each interrogation video that participants watched. Participants indicated whether they believed that the interviewee in the video had or had not cheated on the test (dichotomous culpability judgment: the participant cheated on the test; the participant did not cheat on the test). Then, participants rated: how confident they were in their judgment of whether or not the participant had cheated (Likert-type scale; 1 = *not at all confident* and 10 = *extremely confident*); how likely it was that the participant cheated on the test (Likert-type scale; 1 = *definitely did not cheat* and 10 = *definitely did cheat*); and the extent to which they relied on nonverbal (e.g., eye gaze, body movements), verbal (e.g., participant contradicting self), and paraverbal cues (e.g., the way the participant spoke words, tone of voice) when making their decision about whether or not the participant had cheated (Likert-type scale; 1 = *not at all* and 10 = *a lot*).

Next, participants read that following the questioning session they just viewed, the same interviewee had ended up admitting to using a phone to cheat on the test and signed a document stating this. Participants received this information regardless of whether the interviewee had actually confessed or not in Study 1.[20] After receiving this

[20] I tested confession influence in this manner because I had expected that not all types of evidence disclosure methods would result in confessions, particularly from innocent suspects (this was indeed the case as seen in Study 1). The described design proactively addressed this potential issue.

information, participants were again asked whether they believed that the interviewee had cheated or not on the test, how confident they were in their judgment about whether the interviewee had cheated, and how likely they thought it was that the interviewee had cheated. Participants also indicated the level of influence of the interviewee's admission on their ultimate decision about the interviewee's culpability (Likert-type scale; 1 = *not at all influential* and 10 = *extremely influential*). Lastly, participants used a slider to indicate which of the following they weighted more in their final decision about whether or not the interviewee was guilty of cheating: (1) the statements that the interviewee made in the video of their questioning session; or (2) the written admission of cheating that the interviewee made after the questioning session. Negative values up to –50 indicated more weight placed on the interviewee's statements, positive values up to +50 indicated more weight placed on the interviewee's written admission of cheating, and 0 indicated equal weight.

Finally, participants responded to attention check questions. To ensure that participants had actually viewed the interrogation videos, I asked them: (1) to indicate the sex of the interviewee; (2) whether the interviewee in the video explicitly stated that they were writing something down while looking at the phone, and if so, what did they say that they were writing;[21] and (3) whether the interviewee had explicitly mentioned that the other participant had a phone during the testing session. Participants also responded to an additional attention check question in this block of questions (e.g., for this question, would you please select the number seven?).

[21] I did not exclude participants based on the free response portion of this attention check question.

Procedure

The study took place entirely online. After consenting to participation, participants read that they would be watching videos as part of the study and should ensure that they were in a private setting without distractions and would be able to hear audio. I then presented participants with a sound test, which required them to select the numbers they heard in an audio clip. Participants had to retake this sound test as many times as it took them to select the correct numbers, thus ensuring that participants would be able to hear the interrogation videos presented later in the study.

Upon successfully passing the sound test, participants read that the videos they were about to watch were recorded during a recent study in which the researchers assessed the types of statements that interviewees made when being accused of cheating on a test. Participants further read that the true purpose of this recent study was to understand people's reactions to being interrogated; however, the interviewees were not aware that this was the true purpose of the study and instead believed the study that they were a part of was examining the community's level of general knowledge. I then emphasized to participants that the reactions and statements that they would see in response to the accusation are real, as the interviewees in the videos believed that they were in trouble.

Next, participants read more about the assessment that the interviewees had taken part in (i.e., each assessment session involved two participants answering 20 general knowledge questions in the same testing room and were told that they needed to work alone on answering these questions, could not use any additional resources, such as cellphones, or help each other while working on the questions). Participants then read

that during the testing session, some participants were given the opportunity to cheat, while others were not given the opportunity to cheat; thus, some participants did cheat, and some did not cheat.

At this point, I told participants that they were now about to watch a random selection of videos that depicted the questioning of interviewees from different testing sessions, and that all interviewees, both who had or had not cheated, had been accused of cheating. I emphasized to participants that all pieces of evidence presented to the interviewees in the videos were true (i.e., the researcher did not make them up), and that the pieces of evidence did not exclude the possibility that interviewees did not cheat, as there was a 50% chance that any given interviewee was innocent and a 50% chance that any given interviewee was guilty.

I presented participants with a total of two videos (one randomly selected guilty and one randomly selected innocent; order counterbalanced) associated to the evidence disclosure condition to which they were randomly assigned.[22] After watching the first video, participants completed the post-video questionnaire. Participants then read a reminder that the pieces of evidence did not exclude the possibility that the interviewee did not cheat, as there was a 50% chance that any given interviewee was innocent and a 50% chance that any given interviewee was guilty. The second interrogation video followed this reminder, after which participants completed the post-video questionnaire

[22] In deception detection studies, participants typically view more than two interview videos. However, given the nature of Study 1 from which the videos were drawn, this was not possible. Participants would have presumably seen a pattern in interviewees' responses, as all innocent interviewees in the sample used for Study 2, except one, mentioned the grocery list, whereas none of the guilty participants mentioned a grocery list.

for the second interrogation they viewed. Lastly, participants completed a demographic questionnaire and were debriefed.

Results

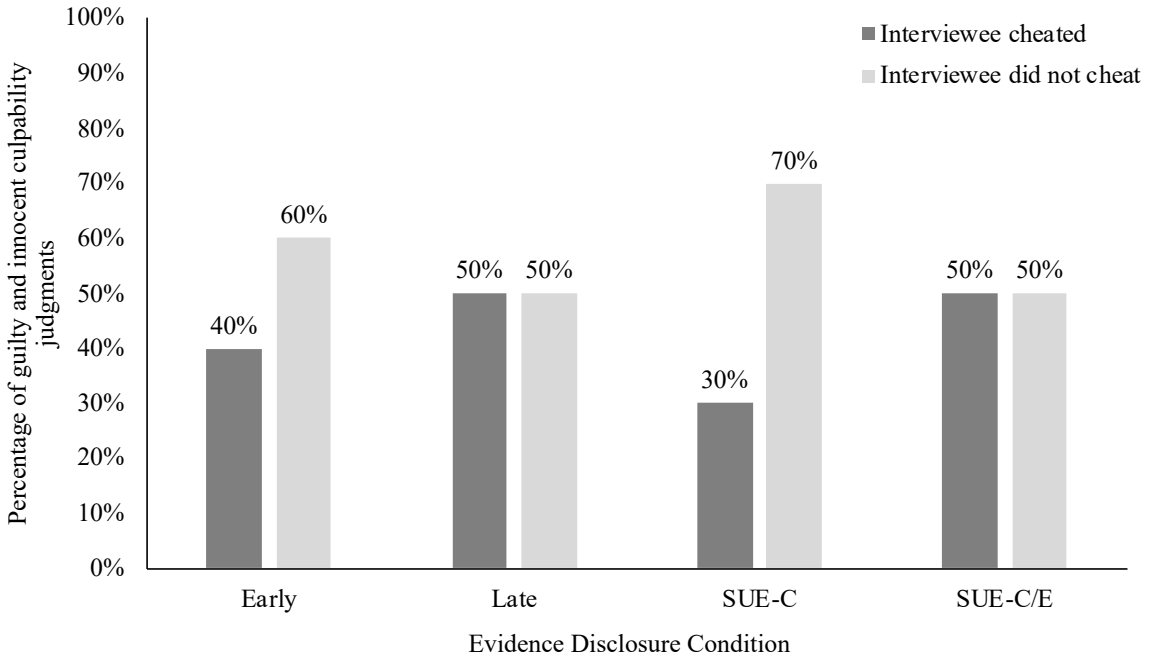
Before conducting the main analyses, I assessed whether participant responses significantly varied depending on the specific video stimuli they viewed. Four one-way ANOVAs (one for each evidence disclosure condition; Table 12 displays results) indicated that participants' pre-confession overall culpability judgment accuracy scores did not significantly differ across the three separate guilty interrogation videos within each of the conditions. Pre-confession overall accuracy scores also did not significantly differ across the individual innocent interrogation videos within each of the evidence disclosure conditions (see Table 12).

Table 12. One-way ANOVA Results Testing Differences in Pre-Confession Overall Culpability Judgment Accuracy Scores Across Individual Video Stimuli

Culpability	Evidence Disclosure	Results
Guilty Videos	Early	$F(2, 21) = .21$, $p = .81$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.11]
	Late	$F(2, 31) = 2.11$, $p = .14$, $\eta_p^2 = 0.12$, 90% CI [0.00, 0.27]
	SUE-C	$F(2, 33) = 1.71$, $p = .20$, $\eta_p^2 = 0.09$, 90% CI [0.00, 0.23]
	SUE-C/E	$F(2, 32) = .02$, $p = .98$, $\eta_p^2 < 0.01$, 90% CI [0.00, 0.00]
Innocent Videos	Early	$F(1, 22) = .26$, $p = .61$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.16]
	Late	$F(2, 31) = .11$, $p = .90$, $\eta_p^2 = 0.01$, 90% CI [0.00, 0.04]
	SUE-C	$F(2, 33) = .04$, $p = .96$, $\eta_p^2 < 0.01$, 90% CI [0.00, 0.00]
	SUE-C/E	$F(2, 32) = .26$, $p = .78$, $\eta_p^2 = 0.02$, 90% CI [0.00, 0.09]

Additionally, to ensure that participants did not have a bias towards rendering either a guilty or innocent culpability judgment, I descriptively assessed the overall percentage of guilty and innocent judgments made in each evidence disclosure condition. As demonstrated in Figure 6, there did not appear to be any ceiling effects in terms of guilty or innocent culpability judgments. Furthermore, SUE-C/E and Late appeared to not lead to any bias whatsoever.

Figure 6. Percentage of Pre-Confession Guilty and Innocent Culpability Judgments by Evidence Disclosure Condition



Note. As each participant made two culpability judgments, the total number of culpability judgments rendered in each evidence disclosure condition were as follows: Early (n = 48), Late (n = 68), SUE-C (n = 72), and SUE-C/E (n = 70).

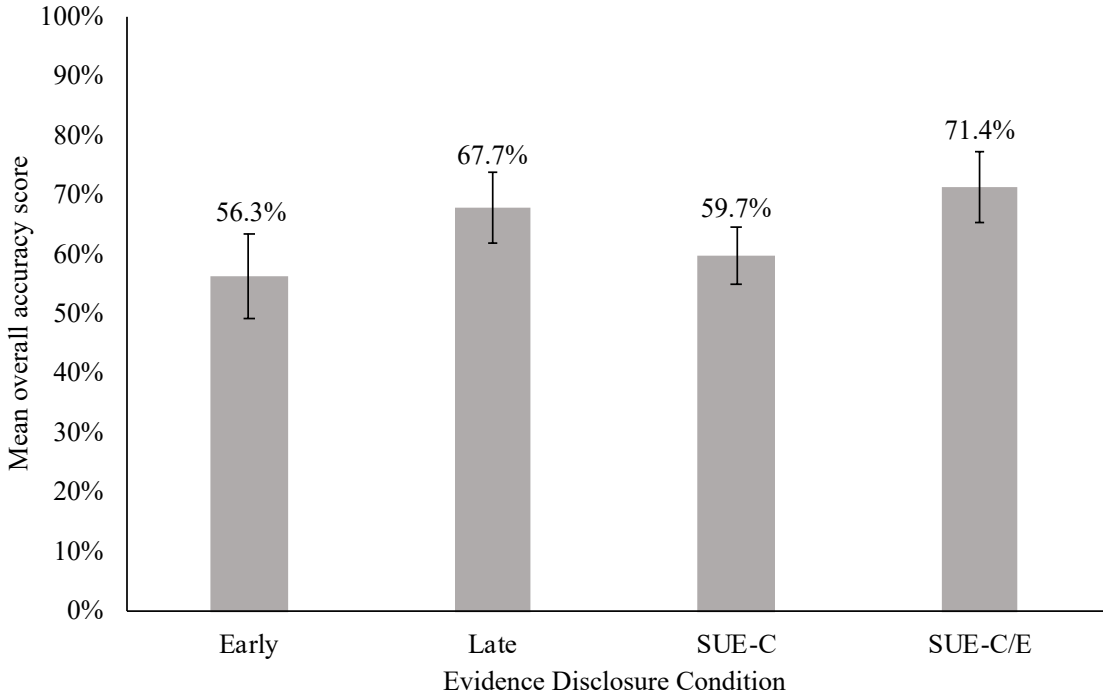
Pre-Confession Judgments

Overall Accuracy

To examine participants' culpability judgment accuracy, I created an overall accuracy variable. As previously mentioned, all participants made two culpability

judgments: one for a guilty interviewee and one for an innocent interviewee. If participants accurately indicated that the guilty interviewee had cheated and that the innocent interviewee had not cheated, participants received an overall accuracy score of 1. If participants only correctly judged one of the interviewees who they viewed, they received a score of 0.5. Participants received a score of 0 if they incorrectly judged both the guilty and innocent interviewees. Across all evidence disclosure conditions, overall accuracy averaged at 64.3% (*SD* = 33.2%). Figure 7 displays participants' overall accuracy scores by evidence disclosure condition. A one-way ANOVA revealed that overall accuracy scores did not significantly differ across the evidence disclosure conditions, $F(3, 125) = 1.36$, $p = .26$, $\eta_p^2 = 0.03$, 90% CI [0.00, 0.08].

Figure 7. Mean Overall Accuracy Scores by Evidence Disclosure Condition



Note. Error bars represent standard errors.

Overall Accuracy Compared to Chance. A one-sample *t*-test revealed that the overall accuracy score (64.3%) was significantly higher than chance (i.e., 50%), $t(128) = 4.91$, $p < .001$, $d = 0.43$, 95% CI [0.25, 0.61]. To determine the extent to which accuracy differed from chance at each level of evidence disclosure, I next conducted four one-sample *t*-tests (critical $p = .013$). Participants who viewed SUE-C/E interrogations demonstrated overall accuracy (71.4%) that significantly exceeded that of chance, $t(34) = 3.63$, $p = .001$, $d = 0.61$, 95% CI [0.25, 0.97]. A similar pattern emerged for participants who viewed Late disclosure interrogations (67.7%), $t(33) = 2.98$, $p = .005$, $d = 0.51$, 95% CI [0.15, 0.87]. Overall accuracy scores in both the Early disclosure (56.3%), $t(23) = 0.90$, $p = .38$, $d = 0.18$, 95% CI [-0.22, 0.59], and SUE-C (59.7%), $t(35) = 2.02$, $p = .05$, $d = 0.34$, 95% CI [-0.001, 0.67], conditions did not significantly differ from chance responding.

Identifying Innocence and Guilt

Table 13 displays the percentage of participants who correctly identified innocent interviewees as having not cheated (i.e., accurately identifying innocence), as well as the percentage of participants who correctly identified guilty interviewees as having cheated (i.e., accurately identifying guilt). A one-tailed, two-proportions *z*-test[23] indicated that the proportion of participants who correctly identified innocence exceeded the proportion of participants who were able to correctly identify guilt, $z = 1.82$, $p = .03$, 95% CI [-0.01, 0.22].

[23] I conducted a one-tailed, rather than two-tailed, *z*-test because I expected participants to demonstrate a truth bias, which is a prevalent finding in the deception detection literature (e.g., Bond & DePaulo, 2006). Additionally, I used a *z*-test here because I was comparing proportions.

Table 13. Percentage of Participants Who Accurately Identified Innocence and Who Accurately Identified Guilt

	Evidence Disclosure Condition				
	Early *n* = 24	Late *n* = 34	SUE-C *n* = 36	SUE-C/E *n* = 35	Total *N* = 129
Innocence Accuracy	62.5%	64.7%	77.8%	71.4%	69.8%
Guilt Accuracy	50.0%	70.6%	41.7%	71.4%	58.9%

Comparing Culpability Judgment Accuracy Across Evidence Disclosure Conditions. I carried out two logistic regressions to examine whether the evidence disclosure method used during interviewees' interrogations predicted participants' ability to accurately identify innocence and guilt (see Table 14 for relevant statistics). The results showed that there was a significant impact of evidence disclosure on participants' ability to accurately identify guilt, $\chi^2(3) = 9.46$, $p = .02$, $R^2_{Nagelkerke} = .10$. Notably, the odds that participants accurately identified guilt were 3.50 times higher if the interrogator had employed SUE-C/E as opposed to SUE-C. A similar pattern emerged when comparing the Late and SUE-C conditions: The odds that participants would accurately identify guilt were 3.36 times higher in the Late condition compared to the SUE-C condition. Despite influencing participants' ability to accurately detect guilty interviewees, evidence disclosure method did not differentially impact participants' ability to accurately identify innocence, $\chi^2(3) = 2.19$, $p = .54$, $R^2_{Nagelkerke} = .02$.

Table 14. Results for Logistic Regression Analyses Examining the Impact of Evidence Disclosure on Participants' Ability to Accurately Identify Innocence and Guilt

Predictor	*B* (*SE*)	Wald χ^2	*p*	OR	95% CI
Identifying Innocence					
Late vs. Early	.10 (.55)	0.03	.86	1.10	[0.37, 3.26]
SUE-C vs. Early	.74 (.58)	1.63	.20	2.10	[0.67, 6.57]
SUE-C/E vs. Early	.41 (.56)	0.52	.47	1.50	[0.50, 4.53]
Late vs. SUE-C	-.65 (.54)	1.44	.23	0.52	[0.24, 2.09]
SUE-C/E vs. SUE-C	-.34 (.55)	0.38	.54	0.71	[0.18, 1.50]
SUE-C/E vs. Late	.31 (.52)	0.36	.55	1.36	[0.49, 3.77]
Identifying Guilt					
Late vs. Early	.88 (.56)	2.49	.12	2.40	[0.81, 7.13]
SUE-C vs. Early	-.34 (.53)	0.40	.53	0.71	[0.25, 2.02]
SUE-C/E vs. Early	.92 (.55)	2.74	.10	2.50	[0.84, 7.40]
Late vs. SUE-C	1.21 (.51)	5.74	.02	3.36	[1.25, 9.06]
SUE-C/E vs. SUE-C	1.25 (.50)	6.17	.01	3.50	[1.30, 9.40]
SUE-C/E vs. Late	.04 (.53)	0.01	.94	1.04	[0.37, 2.95]

Note. The second evidence disclosure condition listed in each of the comparisons served as the reference group for the regression analyses. The model was run three times, with a different reference category each time, in order to make all comparisons across the evidence disclosure conditions.

Post-Confession Judgments

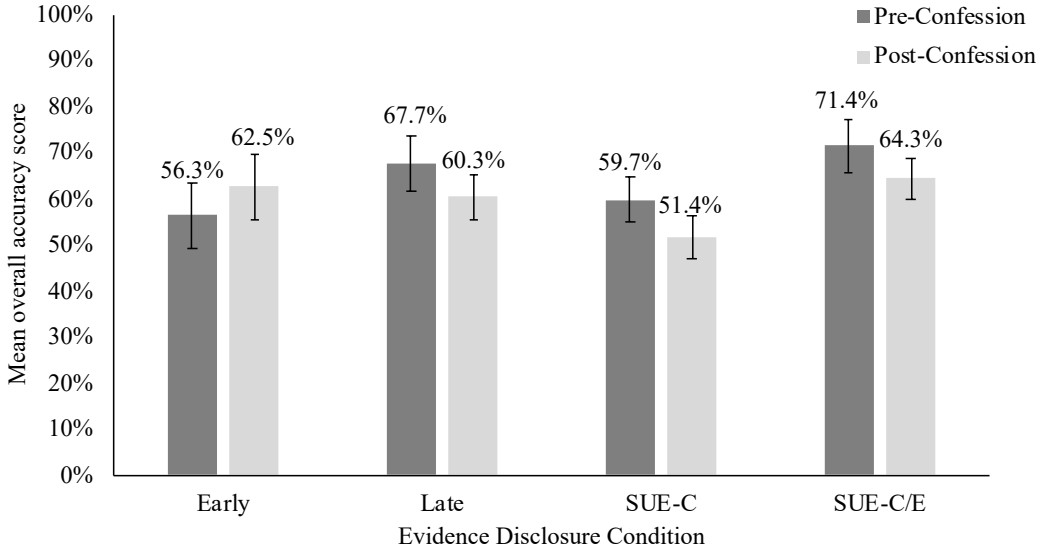
Overall Accuracy

I calculated overall accuracy scores in the same manner as I did for pre-confession overall accuracy scores. Across all evidence disclosure conditions, participants' mean accuracy was 59.3% (*SD* = 29.2%). Similar to the pattern seen for pre-confession scores, post-confession overall accuracy scores did not significantly differ across the evidence disclosure conditions, as revealed by a one-way ANOVA, $F(3, 125) = 1.35$, $p = .26$, $\eta_p^2 = 0.03$, 90% CI [0.00, 0.08] (see Figure 8 for post-confession scores, compared against pre-confession scores).

Overall Accuracy Compared to Chance. As revealed by a one-sample *t*-test, participants' mean accuracy score (59.3%) exceeded chance responding (i.e., 50%),

$t(128) = 3.62$, $p < .001$, $d = 0.32$, 95% CI [0.14, 0.50]. Additionally, four one-sample *t*-tests (Bonferroni-correction: critical $p = .013$) showed that post-confession scores exceeded chance responding in the SUE-C/E condition, $t(34) = 3.26$, $p = .003$, $d = 0.55$, 95% CI [0.19, 0.90], but did not significantly exceed chance in the other three evidence disclosure conditions: Early, $t(24) = 1.81$, $p = .08$, $d = 0.37$, 95% CI [-0.05, 0.78]; Late, $t(33) = 2.03$, $p = .05$, $d = 0.35$, 95% CI [-0.001, 0.69]; nor SUE-C, $t(35) = .30$, $p = .77$, $d = 0.05$, 95% CI [-0.28, 0.38].

Figure 8. Mean Overall Accuracy Scores by Evidence Disclosure Condition (Post-Confession Scores Presented Against Pre-Confession Scores)



Note. Error bars represent standard errors.

Comparing Pre- and Post- Confession Overall Accuracy. A paired-samples *t*-test revealed a significant drop in participants' overall accuracy scores from 64.3% to 59.3% after reading that interviewees had confessed, $t(128) = 1.96$, $p = .05$, Hedge's $g =$

0.16, 95% CI [-0.002, 0.32]. However, a series of four paired-samples *t*-tests[24] (critical *p* = .013) comparing pre- and post-confession scores at each level of evidence disclosure indicated that pre- and post-confession accuracy scores did not significantly differ for any of the evidence disclosure groups: Early, $t(23) = 1.00$, $p = .33$, Hedge's $g = 0.18$, 95% CI [-0.18, 0.55]; Late, $t(33) = 1.71$, $p = .10$, Hedge's $g = 0.22$, 95% CI [-0.04, 0.49]; SUE-C, $t(35) = 1.78$, $p = .08$, Hedge's $g = 0.29$, 95% CI [-0.04, 0.62]; nor SUE-C/E, $t(34) = 1.30$, $p = .20$, Hedge's $g = 0.28$, 95% CI [-0.12, 0.58].

Identifying Innocence and Guilt

Figures 9 and 10 display the percentage of participants who, after reading about a confession: (a) correctly identified innocent interviewees as having not cheated (i.e., accurately identifying innocence), and (b) who correctly identified guilty interviewees as having cheated (i.e., accurately identifying guilt). These percentages are displayed against the related pre-confession percentages, as well as by evidence disclosure condition. The proportion of participants who correctly identified guilt post-confession (80.6%) surpassed the proportion of participants who correctly identified innocence (38.0%), as indicated by a one-tailed, two-proportion *z*-test, $z = 6.97$, $p < .001$, 95% CI [0.31, 0.53].

Change in Accurate Innocence Identifications. I ran five McNemar chi-square analyses[25] (Bonferroni-correction: critical $p = .010$; see Table 15 for McNemar findings), one for each evidence disclosure condition and one for the total, to determine whether the

[24] I decided to run a series of *t*-tests with a Bonferroni correction rather than a repeated measures ANOVA because there was no significant interaction between interviewee veracity and evidence disclosure, meaning that practical insight into the effect of the different evidence disclosure techniques on observers' overall accuracy would not be fully explored.

[25] The McNemar chi-square accounts for the fact that I was analyzing a repeated measure.

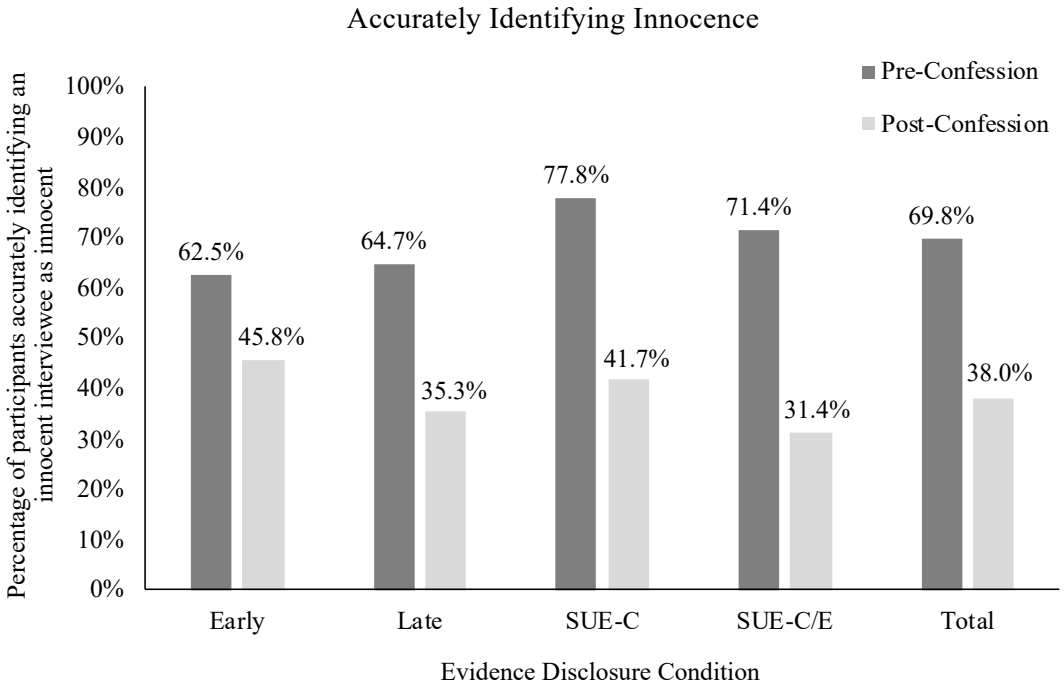
number of participants who correctly identified innocence changed from prior to reading about the confession to after reading about the confession (see Figure 9). Overall, the proportion of participants who correctly judged innocent interviewees as innocent pre-confession (69.8%) was greater than the portion of accurate participants post-confession (38.0%). The pattern of a decrease from pre- to post-confession in the number of participants who accurately identified innocence held across all evidence disclosure conditions, except for the Early disclosure condition in which the decrease was non-significant (p = .05) after applying the Bonferroni-corrected critical p-value.

Table 15. Results from the McNemar Chi-Square Analyses Comparing the Proportion of Participants Pre- and Post-Confession Who Accurately Identified Innocence

Evidence Disclosure Condition	$\chi^2_{McNemar}$	p	95% CI
Early disclosure	4.00	.05	[-0.2%, 31.6%]
Late disclosure	10.00	.002	[12.3%, 43.5%]
SUE-C	13.00	< .001	[18.4%, 50.4%]
SUE-C/E	14.00	< .001	[21.3%, 54.2%]
Total	41.00	< .001	[23.3%, 39.4%]

Note. The confidence interval for all McNemar chi-square tests represents the 95% CI for the difference in paired proportions. Also note that I used the uncorrected McNemar test, as the corrected test is overly conservative.

Figure 9. Percentage of Participants Who Accurately Identified Innocence (Pre- and Post-Confession Across Evidence Disclosure Conditions)



Note. For all evidence disclosure conditions other than Early, the percentage of participants who were accurate was significantly lower at post-confession compared to pre-confession.

Change in Accurate Guilt Identifications. To examine if the number of participants correctly identifying guilt changed from pre- to post-confession (see Figure 10), I again ran five McNemar chi-square analyses (Bonferroni-correction: critical p = .010), one for each evidence disclosure condition and one for the total (see Table 16 for McNemar findings). Across all evidence disclosure conditions, I found that the percentage of participants who correctly identified a guilty interviewee as being guilty increased from pre-confession (58.9%) to post-confession (80.6%). Similarly, the number of accurate participants was higher post- than pre-confession in the Early, SUE-C, and SUE-C/E evidence disclosure conditions. The percentage of accurate participants did not

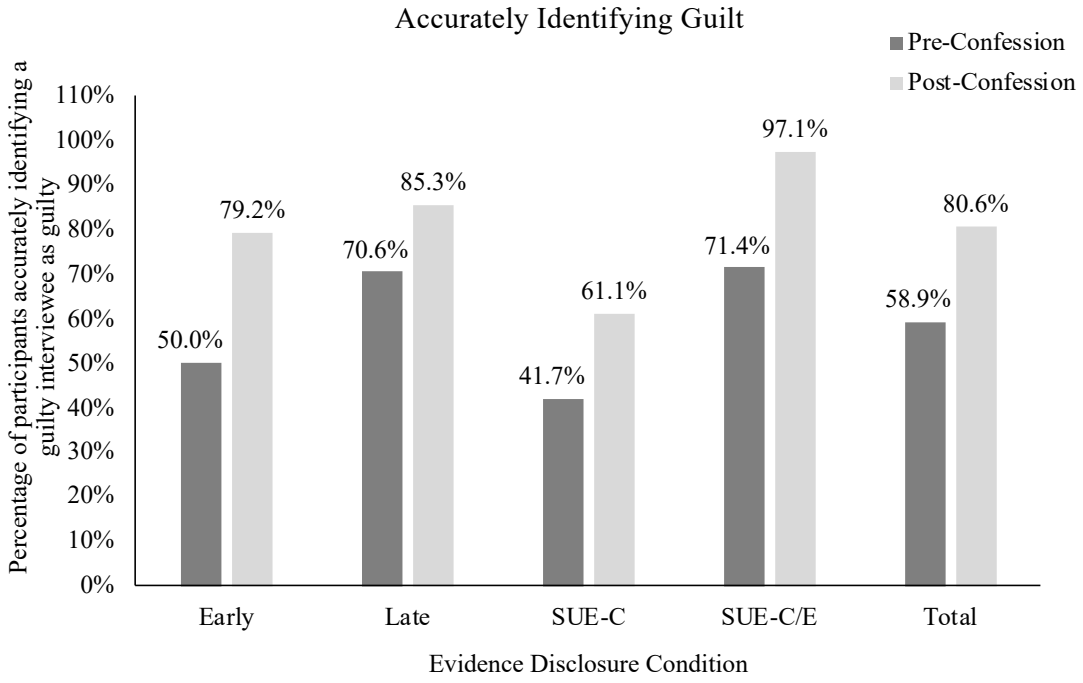
significantly increase in the Late condition after applying the Bonferroni-corrected critical *p*-value.

Table 16. Results from the McNemar Chi-Square Analyses Comparing the Proportion of Participants Pre- and Post-Confession Who Accurately Identified Guilt

Evidence Disclosure Condition	$\chi^2_{McNemar}$	*p*	95% CI
Early disclosure	7.00	.008	[-46.6%, -8.4%]
Late disclosure	5.00	.03	[-28.7%, -1.2%]
SUE-C	7.00	.008	[-42.2%, -9.5%]
SUE-C/E	9.00	.003	[-42.2%, -9.5%]
Total	28.00	< .001	[-28.9%, -14.4%]

Note. The confidence interval for all McNemar chi-square tests represents the 95% CI for the difference in paired proportions. Also note that I used the uncorrected McNemar test, as the corrected test is overly conservative.

Figure 10. Percentage of Participants Who Accurately Identified Guilt (Pre- and Post-Confession Across Evidence Disclosure Conditions)



Note. The percentage of participants who correctly identified guilt was higher post-confession compared to pre-confession for the Early, SUE-C, and SUE-C/E conditions. The pre- and post-confession percentages did not significantly differ in the Late condition.

Culpability Judgment Accuracy Across Evidence Disclosure Conditions. To determine whether the presence of a confession impacted the influence of evidence disclosure method on participants' ability to accurately judge interviewees' culpability, I carried out two logistic regressions (one for identifying innocence and one for identifying guilt) examining post-confession judgments (see Table 17 for regression statistics). The model for identifying guilt was significant, $\chi^2(3) = 16.70$, $p = .001$, $R^2_{Nagelkerke} = .19$. After reading that the interviewee had confessed, the odds that participants who viewed an interrogation characterized by Late disclosure would accurately identify guilt were 3.69 times greater than for participants who viewed a SUE-C interrogation. Similarly, the odds that participants viewing a SUE-C/E interrogation would make an accurate guilt judgment were 21.64 times greater than the odds for participants viewing a SUE-C interrogation and 8.95 times greater than the odds for participants viewing an Early interrogation. Yet, the type of evidence disclosure method employed in the interrogations did not predict participants' ability to accurately identify innocence, $\chi^2(3) = 1.58$, $p = .66$, $R^2_{Nagelkerke} = .02$.

Table 17. Results for Logistic Regression Analyses Examining the Impact of Evidence Disclosure on Participants' Ability to Accurately Identify Innocence and Guilt After Learning that the Interviewee Had Confessed

Predictor	*B* (*SE*)	Wald χ^2	*p*	OR	95% CI
Identifying Innocence					
Late vs. Early	-.43 (.55)	0.65	.42	0.65	[0.22, 1.88]
SUE-C vs. Early	-.17 (.53)	0.10	.75	0.84	[0.30, 2.39]
SUE-C/E vs. Early	-.61 (.55)	1.25	.26	0.54	[0.19, 1.59]
Late vs. SUE-C	-.27 (.49)	0.30	.58	0.76	[0.29, 2.01]
SUE-C/E vs. SUE-C	-.44 (.50)	0.80	.37	0.64	[0.24, 1.70]
SUE-C/E vs. Late	-.17 (.51)	0.12	.73	0.84	[0.31, 2.29]
Identifying Guilt					
Late vs. Early	.42 (.70)	0.37	.55	1.53	[0.39, 5.99]
SUE-C vs. Early	-.88 (.61)	2.11	.15	0.41	[0.13, 1.36]
SUE-C/E vs. Early	2.19 (1.13)	3.75	.05	8.95	[0.97, 82.32]
Late vs. SUE-C	1.31 (.59)	4.85	.03	3.69	[1.16, 11.80]
SUE-C/E vs. SUE-C	3.07 (1.07)	8.25	.004	21.64	[2.65, 176.41]
SUE-C/E vs. Late	1.77 (1.14)	2.48	.12	5.86	[0.65, 53.09]

Note. The second evidence disclosure condition listed in each of the comparisons served as the reference group for the regression analyses. The model was run three times, with a different reference category each time, in order to make all comparisons across the evidence disclosure conditions.

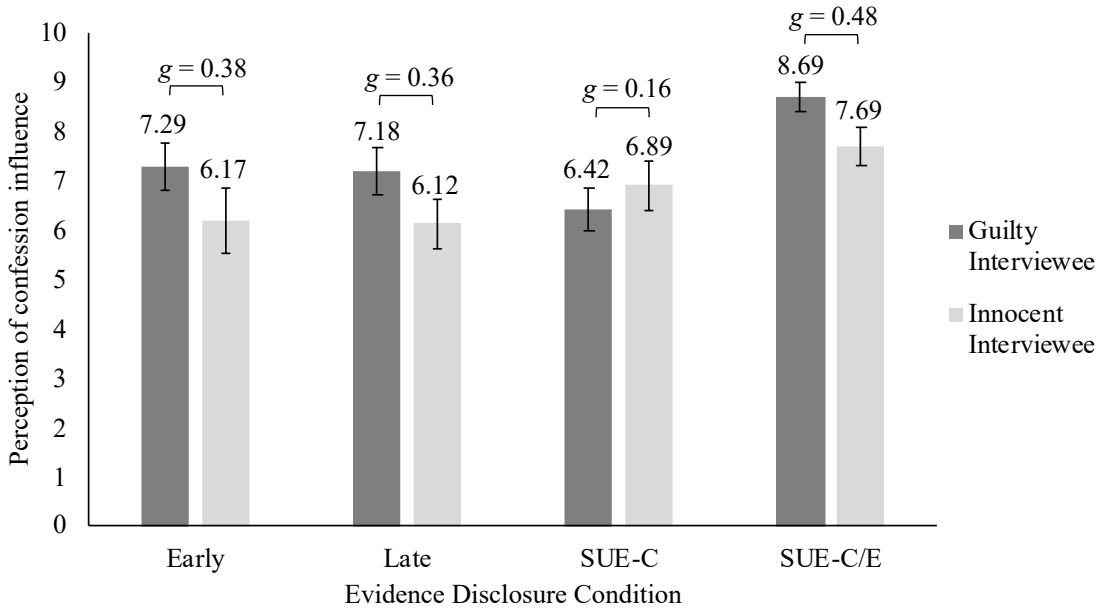
Participants' Subjective Perception of the Confession's Impact on Their Decisions

Participants indicated how influential they believed interviewees' confessions were on their ultimate decisions about interviewees' culpability (see Figure 11 for means). To assess this outcome variable, I carried out four paired-samples *t*-tests[26] (Bonferroni-correction: critical $p = .013$) comparing participants' responses about guilty interviewees to their responses about innocent interviewees. The results indicated that participants in the SUE-C/E condition found guilty interviewees' confessions to be more influential on their ultimate guilt decision than they found innocent interviewees'

[26] Again, I ran a series of *t*-tests with a Bonferroni correction rather than a repeated measures ANOVA because there was no significant interaction between interviewee culpability and evidence disclosure, and the analysis presented here offers more concrete insight into which evidence disclosure technique results in less weight being placed on false confessions.

confessions to be, $t(34) = 2.76$, $p = .009$, Hedge's $g = 0.48$, 95% CI [0.12, 0.85]. A similar pattern emerged for the Late condition, $t(33) = 2.46$, $p = .019$, Hedge's $g = 0.36$, 95% CI [0.06, 0.67], albeit the significance level did not exceed the Bonferroni cut-off value. Confession influence scores did not vary by interviewee culpability for the Early disclosure, $t(23) = 1.51$, $p = .15$, Hedge's $g = 0.38$, 95% CI [-0.14, 0.92], and SUE-C, $t(35) = .95$, $p = .35$, Hedge's $g = 0.16$, 95% CI [-0.18, 0.51], conditions.

Figure 11. Participants' Mean Scores for How Influential They Perceived Interviewees' Confessions to Be on Their Ultimate Decision About Interviewees' Culpability



Note. Error bars represent standard errors. Effect sizes represent Hedge's *g*.

Discussion

Culpability Judgment Accuracy

Across all evidence disclosure conditions, the participants in the present study demonstrated an overall accuracy rate (64.3%) that significantly exceeded chance. This accuracy rate is relatively high when considering the accuracy rate that is typically seen in the tangential deception detection literature (e.g., in their meta-analysis, Bond &

DePaulo, 2006, found an average of 54% accuracy rate). However, the present accuracy rate seems to be driven by the contribution of the SUE-C/E (71.4% overall accuracy) and late disclosure (67.7% overall accuracy) questioning methods to increasing culpability judgment accuracy. The benefit of SUE-C/E and late disclosure seemingly stem from observers being especially successful at identifying guilty interviewees who are being questioned by these two techniques. Notably, participants in the SUE-C/E and late disclosure conditions significantly outperformed participants in the SUE-C condition in regard to accurately identifying guilt specifically. This latter finding is best characterized by a postulation made by Hartwig et al. (2005, p. 480): The SUE technique is "mainly a technique for detecting lies rather than truths." Indeed, the present findings fall in line with past SUE research that indicates that deception detection accuracy is improved, primarily via enhanced detection of liars, when the late disclosure method is used (e.g., Hartwig et al., 2005).

The present findings also contribute new insight to the SUE literature, as there are no known studies that examine the effects of SUE-C and SUE-C/E on culpability judgment accuracy. Interestingly, using SUE-C resulted in the lowest number of participants who were able to accurately identify guilt, whereas using SUE-C/E resulted in the highest number of participants who accurately identified guilt. One potential reason for this finding may rest with the line of questioning that is associated with each of these techniques. Both SUE-C/E and SUE-C involve the interrogator pointing out interviewees' evidence-inconsistent statements; however, it is only in SUE-C/E that interviewees are able to explain why they were inconsistent with the evidence. It is therefore possible that guilty interviewees' explanations for inconsistencies were

perceived as weak by observers, consequently highlighting the interviewees' guilt. Indeed, guilty interviewees may have not been prepared to respond to a request for an explanation, which represents a sort of unanticipated question. When formulating responses prior to an interview, liars do not typically consider the need to prepare for unanticipated questions and thus when put on the spot during an interview, they fail to provide adequate responses (Vrij & Granhag, 2012). As such, guilty interviewees questioned with SUE-C/E may have highlighted their guilt via their responses to explanations for inconsistencies, thus contributing to the high guilt-accuracy achieved by observers in the SUE-C/E condition. Relatedly, it is possible that SUE-C/E's explanation component forced guilty interviewees to more often change their stories (i.e., make within-statement inconsistencies)—a cue to deception that observers in the present study may have relied on. However, given that I did not code for within-statement inconsistencies in Study 1, this is a postulation that future analyses can examine.

The Impact of Confessions

The presence of a confession resulted in a decrease in overall culpability judgment accuracy across all evidence disclosure conditions (decrease from 64.3% to 59.3%), as well as a general decrease in the number of participants who accurately identified innocence. This drop in accurate innocence identifications, which was even seen in the overall well-performing SUE-C/E, falls in line with a common phenomenon in the confession literature; namely, confessions are powerful pieces of evidence that typically sway evaluators to make guilty judgments (e.g., Kassin & Neumann, 1997). Regarding the drop in accurate innocence identifications in the SUE-C/E condition specifically, one potential explanation may be due to SUE-C/E's explanation component.

Specifically, any pre-confession statement inconsistencies may have served to highlight innocent interviewee's innocence (e.g., they had viable explanations for inconsistencies); however, a confession may have undermined this positive effect by leading participants to rely on inconsistencies as indicators of lying.

Confessions did, however, improve participants' accuracy in identifying guilty interviewees, and especially so for the SUE-C/E and late evidence disclosure conditions. Furthermore, the post-confession accuracy rate achieved in the SUE-C/E condition remained significantly above chance—a finding that did not emerge in any of the other three evidence disclosure conditions. This effect appeared to be driven by nearly all participants (a little over 97%) in the SUE-C/E condition accurately identifying guilty interviewees after being told about a confession.

One potential explanation for this latter finding rests with the notion that SUE-C/E is particularly effective in highlighting guilty interviewees' guilt. Indeed, guilt identification accuracy in the SUE-C/E condition was already high (over 70%) prior to participants being told that interviewees had confessed. Due to the aforementioned potential reasons for this result (i.e., inadequate explanations, within-statement inconsistencies), there may have also been a considerable number of participants who were near the threshold of rendering a guilty judgment for guilty interviewees. Given that they were near the threshold of judging the interviewees as guilty, learning that the interviewees had confessed may have then had a greater effect on leading these participants to shift their inaccurate pre-confession judgment of innocence to an accurate post-confession judgment of guilt. Indeed, a descriptive assessment of participants'

reported extent to which a guilty interviewee's confession had influenced their culpability judgment was highest in the SUE-C/E condition.

IV. GENERAL DISCUSSION

The goal of the current research was to address critical gaps that exist within the SUE literature. In doing so, the results from the present studies offer a better understanding of the SUE technique, as well as provide insight as to which variations of SUE are most effective in aiding law enforcement officers in achieving their investigative goals.

Studying the SUE Technique with a Psychologically Realistic Paradigm

All past SUE studies have employed mock crime paradigms. Mock crime studies are extremely important to the advancement of the interrogation field. However, it is important that findings from these studies are replicated using psychologically realistic paradigms given a number of limitations that are tied to mock crime paradigms. Although findings stemming from these mock crime paradigm studies do offer valuable insight on the types of statements guilty and innocent suspects make during an investigative interview, the findings are not highly generalizable to actual criminal contexts in which suspects have (or have not) committed crimes on their own volition. Furthermore, participants in mock crime studies, although motivated to lie, are not in a high stakes situation and do not typically face potential punishment. For instance, in the original SUE study, Hartwig et al. (2005) told participants that they would receive two lottery tickets for being convincing truth-tellers. This incentive is in stark contrast to the stakes the current participants faced: a fine of $50–500 for committing (or not) a transgression of their own volition. Additionally, guilty participants in mock crime studies are sometimes

explicitly told to lie about the fact that they had committed the crime (e.g., Hartwig et al., 2005), offered financial incentives for confessing and no financial reward for making unbelievable denials (e.g., Sellers & Kebbell, 2009), or explicitly told to convince the interrogator that they are innocent (e.g., Luke et al., 2013). In the present study, participants were not given any instruction regarding what goal they should pursue during the interrogation. As Luke et al. (2013) posit, lack of a concrete instruction for guilty participants to lie during the interview enables researchers to gain a better understanding of the types of decisions suspects will make without any external guidance.

Despite these aforementioned differences between mock crime and psychologically realistic paradigms, Study 1's results generally fell in line with findings from past SUE studies (Hartwig et al., 2014). Namely, SUE techniques resulted in guilty suspects being more inconsistent with the evidence as compared to innocent suspects, with the addition of SUE-C and SUE-C/E inducing guilty suspects to being forthcoming. The fact that the present findings correspond to those stemming from past mock crime studies lends further support for the efficacy of the SUE technique in the context of suspect interrogations.

Confessions and the SUE Technique

There is very little work that has directly examined the effect of evidence disclosure on suspect confession decisions, and the two known empirical studies that do address this topic have key limitations. Specifically, these two studies employed mock crime paradigms (the limitations of which have already been discussed) and either resulted in extremely low confession rates (Jordan et al., 2012) or only examined guilty participants' confession decisions (Sellers & Kebbell, 2009). Thus, Study 1 explicitly

measured true and false confession decisions as one of its primary outcome variables and did so using a psychologically realistic paradigm that presumably resulted in participant decisions that would be more reflective of decisions made in real world interrogation settings.

Study 1's confession findings did not conform to my hypothesis that confessions would be most diagnostic in the SUE-C/E condition, followed by the late disclosure and the SUE-C conditions, with all three SUE techniques outperforming the early disclosure condition. Instead, the results demonstrated that confession rates remained relatively stable across evidence disclosure conditions, and no one of the evidence disclosure conditions rendered a high number of false confessions (i.e., there was only one or two false confessors in each of the evidence disclosure conditions). The implications of these findings are twofold. First, the usage of early evidence disclosure did not result in a high number of false confessors. Although early evidence disclosure may be detrimental to investigators' ability to distinguish between liars and truth-tellers, as past (e.g., Hartwig et al., 2006) and the present research demonstrate, it is seemingly not detrimental in regard to increasing the risk of false confession. Second, SUE techniques do not seem to further reduce the risk of false confession as compared to the early evidence disclosure method. However, this finding could be due to the overall floor effects that emerged for false confessions in Study 1, or to the fact that participants rated the evidence against them as weak.

Taken together, the present findings suggest that evidence disclosure techniques used in isolation of other interrogation tactics do not strongly impact suspects' decisions to confess. However, a single interrogation tactic is rarely employed over the course of an

interrogation; instead, multiple tactics are employed throughout suspect interrogations at various timepoints (e.g., Kelly et al., 2016; Leo, 1996). It is therefore important to consider what the effects of different evidence disclosure methods are on suspect confession decisions when these disclosure tactics are paired with other interrogation tactics—an avenue that future research should pursue.

Testing SUE-C and SUE-C/E with Innocent Suspects and Against Late Disclosure

Past studies examining the efficacy of SUE-C (May et al., 2017; Tekin et al., 2015; 2016) and SUE-C/E (Tekin et al., 2016) did not include an innocent suspect sample. Although the counter-interrogation strategy literature suggests that innocent suspects are typically forthcoming (Granhag & Hartwig, 2008) and should therefore be more consistent with the evidence, it is still highly important to include an innocent suspect sample when determining the efficacy of any interrogation tactic (e.g., information diagnosticity can be measured). The SUE literature consistently shows that late evidence disclosure results in innocent suspects making fewer statement-evidence inconsistencies than guilty suspects; thus, it would be expected that SUE-C and SUE-C/E would also demonstrate a similar pattern since suspects are also not presented with all of the evidence at once. However, what remained unknown in the literature was whether SUE-C and SUE-C/E would enhance this difference to a greater extent than late evidence disclosure had been shown to do.

Indeed, Study 1's statement-evidence inconsistency results did highlight the importance of including an innocent suspect sample as well as a comparison to the late disclosure condition. The effect size of the difference in statement-evidence inconsistencies between guilty and innocent suspects was overall greatest in the SUE-C

and SUE-C/E conditions, thus suggesting that these two variations of SUE may be preferable over late disclosure in terms of enhancing the difference in deceptive verbal cues between liars and truth-tellers.

Examining Culpability Judgments in Light of SUE-C and SUE-C/E

Overall, SUE-C/E outperformed the other evidence disclosure methods in terms of aiding observers to better distinguish between guilty and innocent interviewees in Study 2. The goal of the SUE-C/E is to move guilty suspects from engaging in avoidant strategies to becoming more forthcoming during questioning (Tekin et al., 2016). Guilty interviewees questioned with SUE-C/E in Study 1 were indeed more forthcoming, as demonstrated by the overall low number of statement-evidence inconsistencies seen in this condition. The enhanced forthcomingness in the SUE-C/E condition presumably resulted in observers in Study 2 being more easily able to pinpoint guilty interviewees. Interestingly, SUE-C, which is also meant to induce guilty suspects into being more forthcoming (Tekin et al., 2015), did not effectively aid observers in rendering accurate culpability judgments (the lowest accuracy rates were demonstrated in the SUE-C condition in Study 2). As mentioned in Study 2's Discussion, this could have been due to the differences in the line of questioning used in SUE-C versus SUE-C/E: Interviewees in SUE-C did not get the chance to explain any evidence-inconsistent statements as did interviewees in SUE-C/E. Guilty interviewees in the SUE-C/E condition may have not anticipated this additional request for explanations and thus may have provided a weak response and/or a response that was inconsistent with their prior responses, thereby making their guilt more apparent.

Additionally, one potential reason for why participants who viewed late disclosure interrogations outperformed those who viewed SUE-C interrogations rests with the fact that guilty interviewees in late disclosure interrogations received no corrective feedback as they made statement-evidence inconsistencies. Due to the lack of feedback, late disclosure guilty interviewees were presumably not as likely to shift to a forthcoming strategy but instead remained avoidant, which essentially emphasized their guilt to observers. Contrarily, guilty interviewees questioned with SUE-C presumably shifted to a more forthcoming strategy but considering the lack of the benefits rendered by explanations as in SUE-C/E, observers in the SUE-C group were less able to identify guilty interviewees as compared to observers in the SUE-C/E group. Thus, SUE-C failed to adequately assist observers in detecting guilty interviewees as compared to late disclosure and SUE-C/E.

Limitations

Study 1

Although the interrogation paradigm that I employed was highly believable, it did result in many participants randomly assigned to the guilty condition not actually engaging in the cheating transgression. There are two potential reasons for this. First, the act of cheating from a cellphone may have been perceived as an overly extreme form of cheating. Many past studies employed lesser forms of cheating (e.g., Russano et al., 2005, had participants cheat by verbally sharing one answer with a confederate who requested help), and perhaps the use of a cellphone with a cheat sheet on it seemed to be a more severe form of breaking the experimental rules (note: Evans et al., 2013, also made use of a cheat sheet on a cellphone, but participants themselves did not have access to the cheat

sheet or cellphone and only observed the confederate using them). Second, the present study served as the first time a psychologically realistic interrogation paradigm was used with a sample other than college students. It is possible that community member participants take experimental rules more seriously than do college student participants, thus resulting in lower rates of cheating among community member participants. Furthermore, the differences in demographic variables between college student and community member samples could lend themselves as predictors of willingness to engage in cheating behaviors. Considering these possibilities, a future study should directly compare college student and community member samples regarding their cheating and confession behaviors. Findings from such a study would have strong implications for how future interrogation research is conducted.

Another limitation to Study 1 is a limitation that is inherent to all laboratory-based interrogation research: Participants in the study were not suspects of real crimes. In order to examine how guilty and innocent suspects behave during interrogations, researchers must make a trade-off, sacrificing external validity (participants are not real crime suspects) to maximize internal validity (researchers know the ground truth in terms of suspect culpability). Despite this limitation, the present study used a psychologically realistic paradigm (guilty participants cheated on their own accord and both guilty and innocent participants believed that they were in trouble and that there were financial consequences to be faced).

Additionally, the interrogators in the present study were research assistants. Although the interrogators received SUE training that is comparable to what actual law enforcement may receive, the interrogators had no experience in questioning real

suspects, nor did they have experience in conducting investigative interviews, or investigations, beyond the context of this study. This limitation does raise the need for future research to involve law enforcement officers in examining the effects of various interrogation techniques. Such studies would, for example, employ procedures such as those used by Hartwig et al. (2006) and Luke et al. (2016): Law enforcement officers would constitute the participant sample, as they would be trained in different evidence disclosure techniques, conduct interrogations on mock suspects, and make interrogation outcome decisions. Researchers may take this line of work one step further by subjecting participants to a psychologically realistic paradigm (as used in Study 1) and then having law enforcement who were trained (or untrained) in SUE methods conduct the questioning. Although such a study may push ethical boundaries, there are potential solutions to render the study ethically-sound. For instance, law enforcement officers could pretend to be part of the research team so that participants would not believe that they are in trouble with real police.

Study 2

One limitation that emerged in Study 2 was that of participant attention. Given that the study was fully online, I was not able to control participants' engagement in viewing the interrogation videos. I attempted to address this by including a sound check test that participants had to pass in order to continue with a study, thus ensuring that they could, at the bare minimum, hear the interrogation audio. Additionally, I included attention check questions. As evident by the exclusions I made on the basis of these attention check questions, many participants did not seem to pay attention to the content of what was being said during the interrogation videos. Although failure to correctly

respond to attention check questions could have been due to memory errors or inability to retain details discussed during the interrogation, there was no way to rule out that failures were due to participants not paying attention, as the study was completed fully online. Furthermore, the attention check questions were based on highly relevant information discussed during the interrogation (e.g., whether the interviewee discussed writing something while looking at the forbidden phone). It is therefore hard to believe that participants, who had been instructed to determine whether interviewees had or had not cheated on the test, could not correctly answer at least 75% of the attention check questions correctly if they had been paying attention. To resolve this issue, future studies should consider having participants come into the lab to watch interrogation videos and make culpability judgments. This procedure can help researchers better ensure that participants are able to properly hear the audio and are completing the study in an environment without any distractions. Future researchers could also consider incentivizing participants to make accurate culpability judgements, which may encourage participants to pay closer attention to the study stimuli.

Another limitation to Study 2 was the quality of the video stimuli. Although the audio was clear, the visual quality of the video stimuli was not perfectly clear. Specifically, participants could see interviewees' gross movements, but not minute movements (e.g., eye gaze). This limitation, however, could have resulted in participants relying more on verbal cues to deception rather than nonverbal cues, and thus be a potential underlying cause for the relatively high culpability judgment accuracy rate achieved by participants in Study 2. Indeed, reliance on nonverbal cues to deception (e.g., eye gaze) does not result in successful deception detection (e.g., Hartwig & Bond, 2011),

with research instead showing that interviewing techniques focusing on eliciting verbal cues to deception facilitate deception detection accuracy (a tangential measure to the present culpability judgment measure; Vrij & Granhag, 2012).

Practical Implications and Future Directions

The combination of results stemming from Study 1 and Study 2 suggest that SUE-C/E may be the most effective method of evidence disclosure (of those tested in the present research) to employ during suspect interrogations. Regarding confessions, SUE-C/E (like the other SUE methods) does not reduce true confession rates, nor does it inflate false confession rates. Furthermore, SUE-C/E seemingly triggers guilty suspects to be more forthcoming when being questioned (a replication of past findings; Tekin et al., 2016), as well as results in differential rates of statement-evidence inconsistencies between guilty and innocent suspects—a cue that law enforcement officers report considering when detecting deception (Deeb et al., 2018). Lastly, SUE-C/E is most effective in enhancing observers' culpability judgment accuracy. For these reasons, I recommend that the underlying theory and practical concepts of SUE-C/E be explored during trainings in which researchers promote evidence-based interrogation methods to law enforcement.[27]

Regarding evidence-based trainings, we see in prior studies that SUE trainings are effective: Law enforcement officers who are trained in SUE can effectively apply SUE techniques to increase the accuracy of their interrogation outcome decisions (Hartwig et

[27] It is important to highlight during trainings some caveats of the SUE technique. First, the SUE technique works only when evidence exists. Second, since the SUE technique relies on verbal cues to deception, it is effective for cases in which the interviewee is cooperatively speaking with the interrogator. Third, proper implementation of SUE requires preparation time in order to effectively structure evidence-related questions.

al., 2006; Luke et al., 2016). However, Hartwig et al. (2006) and Luke et al. (2016) did not explore SUE-C/E concepts during training, but rather the mechanisms underlying SUE with a focus on late disclosure. As such, future research should aim to examine the effects of training SUE-C/E specifically to law enforcement officers. On this note, it is important to recognize that SUE-C/E is a more complex method compared to late disclosure, therefore presumably requiring more time and resources to train. In instances in which trainings might be limited in time and resources, the training of late disclosure is still valuable. Indeed, the present research shows that late disclosure still renders differences between guilty and innocent interviewees in terms of deceptive cues (albeit to a smaller extent than does SUE-C/E), and enhances observers' culpability judgment accuracy.

It is also important to note that further development and refinement of the SUE technique should be continually pursued. For instance, there is a need to explore the effectiveness of different approaches interrogators can take when confronting statement-evidence inconsistencies (Granhag & Luke, 2018). Most recently, Luke & Granhag (2020) have begun to tackle this need by testing two variations (selective and reactive techniques) of the Shift-of-Strategy (SoS) approach, which is an extension of SUE-C and SUE-C/E variations. The researchers found that participants questioned with the reactive technique (responding to any inconsistency) provided more new information than did participants who were directly questioned with no specific evidence disclosure technique employed; yet, the selective technique (responding to only severe inconsistencies) was ineffective in shifting participants to becoming more forthcoming. Although these findings are insightful, they do highlight the need to conduct further research on how

interrogators should confront statement-evidence inconsistencies, as the framing of confrontations does seemingly have effects on interrogation outcomes. Furthermore, continued research in this domain can help inform the development of law enforcement SUE trainings, as it is important for researchers to be able to provide officers with concrete recommendations on how to implement confrontations in the most effective manner (Granhag & Luke, 2018).

Additionally, it is crucial that more research be conducted regarding the implications that different SUE techniques have for observers' interrogation judgments. One avenue for such future studies concerns the camera focus used for interrogation recordings. For instance, participants in Study 2 only viewed equal-focus interrogations (i.e., both the interrogator and the suspect were in the frame). Past research has shown that when viewing suspect-focused interrogations, observers view the suspect as guiltier than they do when viewing the same interrogation from an equal-focus (see Lassiter, Ware, Lindberg, & Ratcliff, 2010, for a review). It would be interesting to examine whether suspects questioned with SUE techniques are also susceptible to this bias—a research question that future studies should address.

Lastly, future research should examine how the implementation of SUE impacts prosecutors' decisions (e.g., whether to charge a suspect, whether to offer a plea deal, what sentencing to seek). The SUE technique is presumably conducive to helping prosecutors build effective cases, as SUE functions to elicit actionable information from suspects, which can have implications for the specific case outcomes prosecutors seek. Additionally, because SUE has been demonstrated to enhance the difference between guilty and innocent suspects (both in the present study and in past research), investigators

implementing SUE may ultimately aid the prosecution in making more accurate decisions about whether to charge suspects. Future studies should examine these questions, thereby informing the field on whether the benefits of SUE extend beyond the interrogation room.

V. CONCLUSION

The SUE technique is a promising evidence-based interrogation method that can help interrogators achieve their goals of eliciting information from suspects and better distinguish liars from truth-tellers. Findings from the present study further supported the SUE technique's efficacy, as well as offered novel insight on the benefits of employing SUE-C/E. Although researchers can confidently recommend the use of the SUE technique to law enforcement, there is still much work that needs to be accomplished in terms of refining the details surrounding SUE's implementation. By working together on SUE development and training studies, researchers and law enforcement can achieve this goal and ultimately make strides in improving interrogation methods.

LIST OF REFERENCES

Averill, J.R. (1973). Personal control over aversive stimuli and its relationship to stress. *Psychological Bulletin, 80*, 286-303. doi: 10.1037/h0034845

Bond Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality & Social Psychology Review*, *10*, 214-234. doi: 10.1207/s15327957pspr1003_2

Carlson, N. R., Buskist, W., & Martin, G. N. (2000). *Psychology: The science of behaviour*. Harlow, UK: Allyn & Bacon.

Carver, C. S., & Scheier, M. F. (2012). A model of behavioral self-regulation. In P. M. Van Lange, A. W. Kruglanski, E. Higgins (Eds.), *Handbook of theories of social psychology (vol. 1)* (pp. 505–525). Thousand Oaks, CA: Sage Publications Ltd.

Dando, C. J., & Bull, R. (2011). Maximising opportunities to detect verbal deception: Training police officers to interview tactically. *Journal of Investigative Psychology & Offender Profiling*, *8*, 189-202. doi: 10.1002/jip.145

Dando, C. J., Bull, R., Ormerod, T. C., & Sandham, A. L. (2015). Helping to sort the liars from the truth-tellers: The gradual revelation of information during investigative interviews. *Legal & Criminological Psychology, 20*, 114-128. doi: 10.1111/lcrp.12016

Deeb, H., Vrij, A., Hope, L., Mann, S., Granhag, P. A., & Strömwall, L. A. (2018). Police officers' perceptions of statement inconsistency. *Criminal Justice and Behavior, 45*, 644-665. doi: 10.1177/0093854818758808

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, *111*, 203-243. doi: 10.1037/0033-2909.111.2.203

Drizin, S. A., & Leo, R. A. (2004). The problem of false confessions in the post DNA world. *North Carolina Law Review, 82*, 891-1007.

Evans, J. R., Meissner, C. A., Ross, A. B., Houston, K. A., Russano, M. B., & Horgan, A. J. (2013). Obtaining guilty knowledge in human intelligence interrogations: Comparing accusatorial and information-gathering approaches with a novel experimental paradigm. *Journal of Applied Research in Memory and Cognition*, *2*, 83-88. doi: 10.1016/j.jarmac.2013.03.002

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods, 39*, 175-191. doi: 10.3758/bf0319316

Gilovich, T., Savitsky, K., & Medvec, V. H. (1998). The illusion of transparency: Biased assessments of others' ability read one's emotional states. *Journal of Personality & Social Psychology, 75*, 332-346. doi: 10.1037/0022-3514.75.2.332

Granhag, P. A. & Hartwig, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime & Law, 14*, 189-200. doi: 10.1080/10683160701645181

Granhag, P. A. & Hartwig, M. (2015). The strategic use of evidence technique: A conceptual overview. In P. A. Granhag, A. & Vrij, B. Verschuere (Eds.), *Detecting deception: Current challenges and cognitive approaches* (pp. 231-251). Hoboken, NJ: Wiley.

Granhag, P. A., Hartwig, M., Mac Giolla, E., & Clemens, F. (2015). Suspects' verbal counter-interrogation strategies: Towards an integrative model. In P. A. Granhag, A. & Vrij, B. Verschuere (Eds.), *Detective deception: Current challenges and cognitive approaches* (pp. 293-313). Hoboken, NJ: Wiley.

Granhag, P. A. & Luke, T. J. (2018). How to interview to elicit concealed information: Introducing the shift-of-strategy (SoS) approach. In J. P. Rosenfeld (Ed.), *Detecting concealed information and deception: Recent developments* (pp. 271-295). London: Academic Press.

Granhag, P. A., Strömwall, L. A., Willén, R. M., & Hartwig, M. (2013). Eliciting cues to deception by tactical disclosure of evidence: The first test of the Evidence Framing Matrix. *Legal & Criminological Psychology, 18*, 341-355. doi:10.1111/j.2044-8333.2012.02047.x

Gudjonsson, G. H. (2003). *The psychology of interrogations and confessions: A handbook*. Chichester: John Wiley & Sons Ltd.

Hafer, C. L., & Sutton, R. (2016). Belief in a just world. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 145-160). New York, NY: Springer.

Hartwig, M., & Bond, C. F., Jr. (2011). Why do lie-catchers fail? A lens model meta-analysis of human lie judgments. *Psychological Bulletin, 137*, 643-659. doi: 10.1037/a0023589

Hartwig, M., Granhag, P. A., & Luke, T. (2014). Strategic use of evidence during investigative interviews: The state of the science. In D. C. Raskin, C. R. Honts, & J. C. Kircher (Eds.), *Credibility assessment: Scientific research and applications* (pp. 1-36). Oxford, UK: Academic Press (Imprint of Elsevier).

Hartwig, M., Granhag, P. A., & Strömwall, L. A. (2007). Guilty and innocent suspects' strategies during police interrogations. *Psychology, Crime & Law*, *13*, 213-227. doi: 10.1080/10683160600750264

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Kronkvist, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law & Human Behavior, 30*, 603-619. doi: 10.1007/s10979-006-9053-9

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Vrij, A. (2005). Detecting deception via strategic disclosure of evidence. *Law & Human Behavior, 29*, 469-484. doi: 10.1007/s10979-005-5521-x

Hartwig, M., Granhag, P. A., Stromwall, L., Wolf, A. G., Vrij, A., & af Hjelmsäter, E. R. (2011). Detecting deception in suspects: Verbal cues as a function of interview strategy. *Psychology, Crime & Law*, *17*, 643-656. doi: 10.1080/10683160903446982

Houston, K. A., Meissner, C. A., & Evans, J. R. (2014). Psychological processes underlying true and false confessions. In R. Bull (Ed.), *Investigative Interviewing* (pp. 19-34). New York: Springer.

Inbau, F., Reid, J., Buckley, J., & Jayne, B. (2013). *Criminal interrogation and confessions (5th ed)*. Jones & Bartlett Publishers: Burlington, MA.

Johnson, J.E. (1984). Psychological interventions and coping with surgery. In A. Baum, S.E. Taylor & J.E. Singerd (Eds.), *Handbook of psychology and health: Social psychological aspects of health (vol. 4)* (pp. 167-188). Hillsdale, NJ: Erlbaum.

Jordan, S., Hartwig, M., Wallace, B., Dawson, E., & Xhihani, A. (2012). Early versus late disclosure of evidence: Effects on verbal cues to deception, confessions, and lie catchers' accuracy. *Journal of Investigative Psychology and Offender Profiling, 9*, 1-12. doi: 10.1002/jip.1350

Kassin, S. M. (2005). On the psychology of confessions: Does innocence put innocents at risk? *American Psychologist, 60*, 215-228. doi: 10.1037/0003-066X.60.3.215

Kassin, S. M., Drizin, S. A., Grisso, T., Gudjonsson, G. H., Leo, R. A., & Redlich, A. D. (2010). Police-induced confessions, Risk factors and recommendations. *Law & Human Behavior*, *34*, 3-38. doi: 10.1007/s10979-009-9188-6

Kassin, S. M. & Kiechel, K. L. (1996). The social psychology of false confessions: compliance, internalization, and confabulation. *Psychological Science, 7*(3), 125-128. doi: 10.1111/j.1467-9280.1996.tb00344.x

Kassin, S. M., Leo, R. A., Meissner, C. A., Richman, K. D., Colwell, L. H., Leach, A. M., & La Fon, D. (2007). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law & Human Behavior*, *31*, 381-400. doi: 10.1007/s10979-006-9073-5

Kassin, S. M., & Neumann, K. (1997). On the power of confession evidence: An experimental test of the fundamental difference hypothesis. *Law and human Behavior*, *21*, 469-484. doi: 10.1023/A:1024871622490

Kelly, C. E., Miller, J. C., Redlich, A. D., & Kleinman, S. M. (2013). A taxonomy of interrogation methods. *Psychology, Public Policy, & Law*, *19*, 165-178. doi: 10.1037/a0030310

Kelly, C. E., Redlich, A. D., & Miller, J. C. (2015). Examining the meso-level domains of the interrogation taxonomy. *Psychology, Public Policy & Law, 21*, 179-191. doi: 10.1037/law0000034

Lassiter, G. D., Ware, L. J., Lindberg, M. J., & Ratcliff, J. J. (2010). Videotaping custodial interrogations: Toward a scientifically based policy. In G. D. Lassiter & C. A. Meissner (Eds.), *Police interrogations and false confessions: Current research, practice, and policy recommendations* (p. 143–160). American Psychological Association: Washington, D.C.

Leo, R. A. (1996). Inside the interrogation room. *Journal of Criminal Law and Criminology, 86*, 266-303.

Leo, R. A. (2008). *Police interrogation and American justice*. Cambridge, MA: Harvard University Press.

Lerner, M. J. (1980). *The belief in a just world*. New York: Plenum.

Levine, T. (2018). Ecological validity and deception detection research design. *Communication Methods & Measures, 12*, 45-54. doi: 10.1080/19312458.2017.1411471

Luke, T. J., Dawson, E., Hartwig, M., & Granhag, P. A. (2014). How awareness of possible evidence induces forthcoming counter-interrogation strategies. *Applied Cognitive Psychology, 28*, 876-882. doi: 10.1002/acp.3019

Luke, T. J., & Granhag, P. A. (2020). *The Shift-of-Strategy (SoS) approach: Using evidence strategically to influence suspects' counter-interrogation strategies.* PsyArXiv. doi: 10.31234/osf.io/wncb5

Luke, T. J., Hartwig, M., Brimbal, L., Chan, G., Jordan, S., Joseph, E., Osborne, J., & Granhag, P. A. (2013). Interviewing to elicit cues to deception: Improving Strategic Use of Evidence with general-to-specific framing of evidence. *Journal of Police & Criminal Psychology, 28*, 54-62. doi: 10.1007/s11896-012-9113-7

Luke, T. J., Hartwig, M., Joseph, E., Brimbal, L., Chan, G., Dawson, E., Joradn, S., Donovan, P., & Granhag, P. A. (2016). Training in the Strategic Use of Evidence technique: Improving deception detection accuracy of American law enforcement officers. *Journal of Police & Criminal Psychology*, *31*, 270-278. doi: 10.1007/s11896-015-9187-0

May, L., Granhag, P. A., & Tekin, S. (2017). Interviewing suspects in denial: On how different evidence disclosure modes affect the elicitation of new critical information. *Frontiers in Psychology*, *8*, 1-11. doi: 10.3389/fpsyg.2017.01154

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia medica, 22*, 276-282.

Moston, S., Stephenson, G. M., & Williamson, T. M. (1992). The effects of case characteristics on suspect behaviour during police questioning. *The British Journal of Criminology*, *32*, 23-40. doi: 10.1093/oxfordjournals.bjc.a048178

Ofshe, R. J., & Leo, R. A. (1997). The decision to confess falsely: Rational choice and irrational action. *Denver University Law Review, 74*, 979-1121.

Perillo, J. T. & Kassin, S. M. (2011). Inside interrogation: The lie, the bluff, and false confessions. *Law & Human Behavior*, *35*, 327-337. doi: 10.1007/s10979-010-9244-2

Russano, M. B., Meissner, C. A., Narchet, F. M., & Kassin, S. M. (2005). Investigating true and false confessions within a novel experimental paradigm. *Psychological Science, 16*, 481-486. doi: 10.1111/j.0956-7976.2005.01560.x

Read, J. M., Powell, M. B., Kebbell, M. R., & Milne, R. (2009). Investigative interviewing of suspected sex offenders: A review of what constitutes best practice. *International Journal of Police Science & Management, 11*, 442-459. doi: 10.1350/ijps.2009.11.4.143

Sellers, S., & Kebbell, M. R. (2009). When should evidence be disclosed in an interview with a suspect? An experiment with mock-suspects. *Journal of Investigative Psychology & Offender Profiling*, *6*, 151-160. doi: 10.1002/jip.95

Sorochinski, M., Hartwig, M., Osborne, J., Wilkins, E., Marsh, J., Kazakov, D., & Granhag, P. A. (2014). Interviewing to detect deception: When to disclose the evidence? *Journal of Police & Criminal Psychology*, *29*, 87-94. doi: 10.1007/s11896-013-9121-2

Strömwall, L. A., Hartwig, M., & Granhag, P. A. To act truthfully: Nonverbal behaviour and strategies during a police interrogation, *Psychology, Crime & Law, 12*, 207-219. doi: 10.1080/10683160512331331328

Swanner, J. K., Meissner, C. A., Atkinson, D. J., & Dianiska, R. E. (2016). Developing evidence-based approaches to interrogation. *Journal of Applied Research in Memory & Cognition, 5*, 295-301. doi: 10.1016/j.jarmac.2016.07.001

Tekin, S., Granhag, P. A., Strömwall, L., & Vrij, A. (2016). How to make perpetrators in denial disclose more information about their crimes. *Psychology, Crime & Law, 22*, 561-580. doi: 10.1080/1068316X.2016.1168425

Tekin, S., Granhag, P. A., Strömwall, L., Giolla, E. M., Vrij, A., & Hartwig, M. (2015). Interviewing strategically to elicit admissions from guilty suspects. *Law & Human Behavior, 39*, 244-252. doi: 10.1037/lhb0000131

Vredeveldt, A., van Koppen, P. J., & Granhag, P. A. (2014). The inconsistent suspect: A systematic review of different types of consistency in truth tellers and liars. In R. Bull (Ed.), *Investigative interviewing* (pp. 183-207). New York, NY: Springer.

Vrij, A. (2003). We will protect your wife and child, but only if you confess. In van Koppen, P.J. & Penrod, S.D. (Eds.) *Adversarial versus Inquisitorial Justice: Psychological Perspectives on Criminal Justice Systems* (pp. 55-79). New York: Kluwer Academic.

Vrij, A., & Granhag, P. A. (2012). Eliciting cues to deception and truth: What matters are the questions asked. *Journal of Applied Research in Memory & Cognition*, *1*, 110-117. doi: 10.1016/j.jarmac.2012.02.004

Vrij, A., Hope, L., & Fisher, R. P. (2014). Eliciting reliable information in investigative interviews. *Policy Insights form the Behavioral & Brain Sciences, 1*, 129-136. doi: 10.1177/2372732214548592

Vrij, A., Leal, S., & Fisher, R. P. (2018). Verbal deception and the model statement as a lie detection tool. *Frontiers in Psychiatry*, *9*, 492. doi: 10.3389/fpsyt.2018.00492

Vrij, A., Meissner, C. A., Fisher, R. P., Kassin, S. M., Morgan III, C. A., & Kleinman, S. M. (2017). Psychological Perspectives on Interrogation. *Perspectives on Psychological Science, 12*, 927-955. doi: 10.1177/1745691617706515

Wald, M., Ayres, R., Hess, D. W., Schantz, M., & Whitebread, II, C. H. (1967). Interrogations in New Haven: The impact of Miranda. *Yale Law Journal*, *76*, 1519-1648.

Walsh, D., & Bull, R. (2015). Interviewing suspects: Examining the association between skills, questioning, evidence disclosure, and interview outcomes. *Psychology, Crime & Law, 21*, 661–680. doi: 10.1080/1068316X.2015.1028544

VITA

AMELIA MINDTHOFF

2012	B.A., Psychology University of California, Los Angeles Los Angeles, California
2015	M.A., Psychology California State University, Fullerton Fullerton, CA

PUBLICATIONS AND PRESENTATIONS

Blandón-Gitlin, I., & Mindthoff, A. (2018). Do video recordings help jurors recognize coercive influences in interrogations? In C. J. Najdowski and M. C. Stevenson (Eds.), *Criminal juries in the 21st century: Contemporary issues, psychological science, and the law* (pp. 195-220). New York, NY: Oxford University Press.

Goldfarb, D., & Mindthoff, A. (2020). Often but not always: When does age at the time of event predict memory for sexual violence? In J. D. Pozzulo, E. Pics, & C. Sheahan (Eds.), *Memory and Sexual Misconduct: Psychological Research for Criminal Justice*. Taylor Francis.

Mindthoff, A., Malloy, L. C., & Höhs, J. (2020). Mock jurors' perceptions and case decisions following a juvenile interrogation: Investigating the role of interested adults and confession type. *Law & Human Behavior*. Advance online publication.

Mindthoff, A., Goldfarb, D., & Behre, K. (2020) How social science can help us understand why family courts discount women's testimony in intimate partner violence cases. *Family Court Review, 53*, 243-264.

Mindthoff, A., Evans, J. R., Perez, G., Woestehoff, S. A., Olaguez, A. P., … & Stocks, E. L. (2020). Juror perceptions of intoxicated suspects' interrogation-related behaviors. *Criminal Justice & Behavior, 47*, 222-246.

Mindthoff, A., Pena, M. M., & Roger, C. D. (2020, June). Mock Jurors' Perceptions of Confession and Forensic Evidence. Paper talk accepted to presented at the Virtual 9th Annual Forensic Science Symposium, Miami, FL.

Mindthoff, A., Evans, J. R., & Luke, T. (2020, March). *Examining the Strategic Use of Evidence Using a Psychologically Realistic Interrogation Paradigm*. Paper talk presented at the American Psychology-Law Society Annual Conference, New Orleans, LA.

Mindthoff, A., & Pena, M. M. (2020, March). *"I Knew He Confessed": Mock Jurors' Decisions in Light of Unreliable Confession & Forensic Evidence*. Paper talk presented at the American Psychology-Law Society Annual Conference, New Orleans, LA.

Mindthoff, A., Hagsand, A. V., Schreiber Compo, N., & Evans, J. R. (2019). Does alcohol loosen the tongue? Intoxicated persons' willingness to report transgressions or criminal behavior carried out by themselves or others. *Applied Cognitive Psychology, 33*, 414-425. doi: 10.1002/acp.3480

Mindthoff, A., Stephens, D. P., & Madhivanan, P. (2019). Using bystander programs to tackle college sexual violence: A critical evaluation of a meta-analysis. *Journal of American College Health, 67*, 85-87. doi: 10.1080/07448481.2018.1467423

Mindthoff, A., Höhs, J., & Malloy, L. C. (2019, March). *The role of allied adults in juvenile interrogations: How do potential jurors evaluate parents and attorneys?* Poster presented at the biennial meeting of the Society for Research in Child Development, Baltimore, MD.

Mindthoff, A., Höhs, J., & Malloy, L. C. (2019, March). Perspectives from the jury box: Having an allied adult present during a juvenile interrogation. In L.C. Malloy, *Do parents know best? The role of parents in juvenile interrogation and plea bargaining contexts*. Symposium presented at the annual meeting of the American Psychology-Law Society, Portland, OR.

Mindthoff, A., Evans, J. R., Schreiber Compo, S., Wolfs, A., Hagsand, A. V., & Goldstein, N. E. (2019, March). *The effects of alcohol on individuals' Miranda rights comprehension.* Paper talk presented at the American Psychology-Law Society Annual Conference, Portland, OR.

Mindthoff, A., Evans, J. R., Perez, G., Woestehoff, S. A., Olaguez, A. P., … & Woody, W. D. (2018). A survey of potential jurors' perceptions of interrogations and confessions. *Psychology, Public Policy & Law*, *24*, 430-448.

Pena, M. M., Klemfuss, J. Z., Loftus, E. F., & Mindthoff, A. (2017). The effects of exposure to differing amounts of misinformation and source credibility perception on source monitoring and memory accuracy. *Psychology of Consciousness: Theory, Research, & Practice, 4*, 337-347.

Polanco, K., Mindthoff, A., & Malloy, L. C. (2020, March). *The Influence of Juvenile Interrogation Conditions on Jurors' Perceptions of Defendant Culpability and Potential to Rehabilitate*. Poster presented at the American Psychology-Law Society Annual Conference, New Orleans, LA.