7-2021

What Makes a Lie? A Novel Study Investigating the Difference Between Deliberate Lies and Honest Mistakes

Andrew Christopher Provenzano
University of Arkansas, Fayetteville

What Makes a Lie? A Novel Study Investigating the Difference Between Deliberate Lies and Honest Mistakes

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Psychology

by

Andrew Christopher Provenzano
University of North Florida
Bachelor of Arts in Psychology, 2013
University of North Florida
Master of Arts in Psychology, 2015

July 2021
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

_____
James Lampinen, Ph.D
Dissertation Director

_____
Doug Behrend, PhD
Committee Member

_____
Lindsay Ham-Holm, Ph.D
Committee Member

Abstract

Previous research on lie detection suggests that people use cues to deception to make a true-false judgment about a statement. However, no prior research has investigated what factors cause others to classify known false statements as an intentional lie or an "honest" mistake. This dissertation reports two studies that sought to answer this question. Experiment 1 consisted of a diary study where participants reported false statements and then described their reason for classifying the statements as a deliberate lie or honest mistake. In Experiment 2, participants completed a semi-structured interview where they described various false statements and why they classified the statement as a deliberate lie or honest mistake. I found that participants used different kinds of cues to classify honest mistakes and deliberate lies. *Perceived memorability* and *lack of motive* were the most common rationale participants gave for classifying a false statement as an honest mistake. *Motive, reputation,* and *the presence of non-verbal cues* were most associated with deliberate lies. This dissertation suggests that people classify deliberate lies and honest mistakes using different criteria.

Acknowledgments

I would like to thank Dr. James Lampinen for his help and support on this project. Jim, I am so grateful for all the support you have provided me over the years. This past year has been such a challenge, but you made sure I felt supported throughout it all. Thanks are also extended to my other committee members, Dr. Doug Behrend and Lindsay Ham-Holm. I appreciate the feedback that they provided me throughout this project.

I would like to thank the members of the Law and Memory Processes for all the support and encouragement throughout the years. Blake, Kara, and Brittany, thank you for welcoming me when I arrived at the University of Arkansas. Your support as the senior graduate students in the lab made the difficult transition of starting a PhD. program so much more manageable. Nia and Amber (LL Cool Lab), thank you for being two of the best friends (and lab mates) I could ask for. You are both such good-hearted and sincere friends, and that means the world to me. I will truly miss seeing you all the time. Thank you to the many undergraduate research assistants I had during my time at the University of Arkansas. Special thanks to Gigi Rashed, Madison Hawley, Mady Fancher, Megan Bolejack, Macy Valentine, Meredith Bart, and Madison Howard for your help on this project and for listening to all my practice my job talks.

Special thanks to my mother, Carolyn Provenzano, and my great-aunt Florence Fox. You are two of the most important people in my life. Every day I wish both of you were still here to celebrate this achievement. I would not be where I am today without both of your love and support when I was growing up. To my niece, Amelia, this past year has been challenging, but your entry into the world a few months ago was truly a bright spot for me.

Fino Alla Fine.

Table of Contents

Introduction

Imagine that a person tells you something, which later turns out to be untrue. There are two possible interpretations of this situation. On the one hand, the person may have intentionally set out to deceive you. The person may be motivated by greed, embarrassment, or a desire to avoid responsibility for a myriad of other reasons. There is a long history of psychological science research dealing with deception detection, and this research goes back more than a century .(Ekman & Friesen, 1969; Zuckerman, 1981). On the other hand, the person making the false statement, may honestly believe that the information is accurate. They may simply be honestly mistaken or perhaps confused. There is also a long history in psychological science dealing with honest errors in memory, perception, and inference-making that goes back more than a century (Tourangeau, 1999).

These two pieces of literature have never been juxtaposed against each other. This gap in our understanding is highly consequential. Deception detection research deals exclusively with understanding how people distinguish between knowingly false statements and statements that are true. False memory research deals with how people differentiate between statements that are honest mistakes and statements that are true. However, what is left out is understanding how people distinguish between false statements that are deliberately untruthful and statements that are honest mistakes. This distinction occurs commonly in everyday life and legal contexts but to my knowledge this has not been studied by scientists.

Consider the following example. In 1998, Vice President Al Gore was investigated for potentially violating a law prohibiting someone from raising campaign money on government property (Suro, 1998). Gore had apparently made fundraising calls from his office. During the

investigation, Gore told investigators that he did not think the law applied because all the calls involved raising soft money, not hard money. Soft money is money earmarked for political parties or political action committees to serve generic purposes like increasing voter turnout. Hard money, on the other hand, goes directly to the campaign for the purpose of getting a candidate elected. It later turned out that some of the money Gore raised was indeed hard money, money that went directly to the campaign. How are we to interpret Gore's actions. Was he lying to federal investigators (itself a crime)? Was he uninformed? Was he honestly mistaken? Had he been so overwhelmed by the number of calls that he honestly forgot that some of the calls generated hard money?

In the realm of the criminal justice system, the ability to classify statements as truths or lies is imperative because they can have life-changing implications (e.g., innocent suspect convicted, guilty suspect avoiding justice). Occasionally we encounter clear and obvious false information. The false statement may have been motivated by an intent to deceive, but may also be due to some kind of "honest" mistake (e.g., Memory error). Prosecutions for the crime of perjury hinge on this distinction (Baran & Ruby, 1997). The courts have generally recognized that some false statements under oath occur due to the limitations of memory, limitations of perception, misunderstandings or miscommunication (i.e., honest mistakes). Making a false statement under oath is only a crime if it was an intentional falsehood (i.e., deliberate lie).

This dissertation reports the results of two studies that aimed to classify what factors cause people to classify a known false statement as either an honest mistake or a deliberate lie. First, I begin with a review of the classic and modern theories of lie detection. Next, I discuss the occurrence of memory errors using the source-monitoring framework. Then, I review how the

recent work involving alibi generation may inform us on the processes underlying how known false statements are judged as a deliberate lie or an "honest" mistake. Finally, I consider the factors that I predict cause individuals to classify a false statement as an honest mistake or deliberate lie before describing the current studies.

Lie Detection

Traditional lie detection. Over the previous half-century, various theories sought to explain how non-verbal and verbal cues predicted deception. Ekman and Friesen (1969) proposed the first major cue-based theory, which broke up cues into either *leakage cues* or *deception cues*. Leakage cues refer to a kind of non-verbal behavior that may be more readily revealed in situations when one is trying to hide one's real feelings. For example, a leakage cue may be a brief micro-expression (e.g., smile, smirk) that occurs when trying to suppress an undesired emotion. Emotions that leak out while deceiving include fear of getting caught, guilt, and duping delight (pleasure at succeeding with one's lies).

Ekman and Friesen's (1969) theory has been refined over time, including a shift in where the leakage cues occur. Initially, Ekman hypothesized that liars primarily displayed leakage cues in their hands and feet because deceivers actively tried to control their facial expressions. However, Ekman later revised the theory and emphasized the micro-expressions of the face (Ekman, 1985). While still highly influential, Ekman's theory is critiqued for its emphasis on emotional leakage (Levine, 2014).

The 2nd popular classical theory of deception detection was proposed by Zuckerman et al. (1981), which used a four-factor model to predict deceptive behavior. Zuckerman's model focused on the following four factors to explain deceptive behavior, (a) generalized arousal, (b) guilt and other emotions, (c) cognitive aspects, and (d) liars' attempts to control verbal and non-

verbal cues to appear honest. Zuckerman proposed that liars who experience generalized arousal display greater pupil dilation, increased blinking, more frequent speech disturbances, and speak at a higher pitch. Additionally, when fearful or feeling guilt, liars maintain less eye contact, fidget more, and are evasive while answering questions. Zuckerman's theory proposed that lying is more cognitively taxing than telling the truth. Given this, liars have longer response latencies, more speech hesitations, more significant pupil dilation, and the use of fewer illustrators (e.g., hand movements) than truth-tellers. Zuckerman's thoughts regarding the mentally taxing nature of lying would influence more modern deception work, including Vrij's Cognitive Lie Detection Approach.

The third influential theory of deception detection is DePaulo's self-presentational view. Unlike Ekman's theory, DePaulo's research focused on everyday lies, not high-stakes lies. According to DePaulo (1992), people actively try to regulate their non-verbal behaviors based on self-presentational goals. According to DePaulo, liars and truth-tellers aim to convince others they are truthful. Thus, liars and truth-tellers engage in deliberate behaviors to be perceived as trustworthy.

Over the previous two decades, there has been a steady decrease in research investigating nonverbal cues to deception (Verschuere et al., 2018). In the following section, we will describe more recent theories of deception detection.

Cognitive Lie Detection Approach. The cognitive lie detection (CLD) approach aims to move beyond criteria-based approaches, such as SVA and CBCA, and expand beyond the classical theories. Unlike the classical approaches (e.g., Ekman, Zuckerman) to lie detection, CLD places the observer in a more active interviewing role. The critical element of CLD is the belief that lying requires more cognitive effort than describing a real episodic memory (Vrij et

al., 2010). Support for CLD comes from evidence that in an interview setting, lying is more mentally taxing than truth-telling (Christ et al., 2009). Therefore, researchers using CLD aim to increase the cognitive load during the retrieval task to magnify the differences between truth-tellers and liars.

The cognitive lie detection approach consists of (i) imposing cognitive load, (ii)asking interviewees for more information, (iii) asking unexpected questions. Within the lie detection literature, cognitive load increases by asking interviewees to tell their stories in reverse order (e.g., Evans et al., 2013; Vrij et al., 2012), carry out two tasks at once while maintaining eye contact during the interview. The 2nd approach that predicts deception is to ask interviewees for additional information about the event in question (Vrij, 2016). To achieve this, interviewers may use the cognitive interview or have interviewees draw an image of the event in question. CLD proposes responses to these questions will be less detailed, take longer to generate, and contain less information. A rationale for this is liars are cautious about providing additional details out of fear that they might say something incriminating. The 3rd approach to CLD is to have the interviewer ask unexpected questions during the interview. This method is to counter a common tactic by liars, including preparing their statements beforehand and practicing answering anticipated questions.

According to a meta-analysis conducted by Vrij et al. (2017), the cognitive lie detection (71%) approach predicts veracity significantly more than the standard (non-interviewing) approach (56%). Additionally, the cognitive lie detection approach (65%) revealed significantly more cognitive cues than the standard approach (30%). More specifically, CLD elicited more detail, plausibility, and consistency cues than the standard approach.

While highly influential, this approach has several flaws. Unlike what CLD proposes, lying is not always more mentally taxing than telling the truth (Sporer, 2016). Second, CLD is unproven regarding intentions to lie and or stigmatized individuals. Finally, CLD lacks a robust theoretical model that emphasizes specific cognitive processes in lying (Sporer, 2016).

Truth-Default Theory. Truth-Default Theory (TDT) is a theory of deception and deception detection that operates under the assumption that when humans interact with one another, they are typically honest (Levine, 2014). Therefore, humans tend to believe others unless a trigger to deception is activated. TDT incorporates the idea that people are typically "truth-biased" into its framework (McCornack & Parks, 1986; Zuckerman et al., 1981). TDT incorporates several free-standing models including, the Few Prolific Liars Model (Serota et al., 2010; Serota & Levine, 2015), Deception Motives Model (Levine, Kim, & Hamel, 2010), and the Park–Levine Probability Model (Park & Levine, 2001). Unlike the pre-TDT viewpoints, TDT does not view *truth-bias* as maladaptive and a flawed decision rule (Buller & Burgoon, 1996; O'Sullivan, 2003). According to TDT, operating on a truth-default assumption allows for more efficient communication (Levine, 2014). This idea is supported by both the *base rate of lying* and the *veracity effect*. In typical research paradigms, there is a 50/50 chance of being exposed to a truthful or false statement, and truth-bias impacts the direction of these errors (see veracity effect). In more naturalistic settings, base rates of receiving an honest message are much higher (Levine, 2014; Serota & Levine, 2015). When this is taken into account, truth bias enhances deception detection accuracy in non-research situations (Levine & Serota, 2015). Furthermore, there is evidence that deception detection accuracy is contingent on truth-lie base rates (Levine, 1999). The Park-Levine model (Park, Levine, 2017) is a well-supported formula used to predict accuracy rates based on different truth-lie base rates (Levine et al., 2014; Levine

et al., 2006). The Park-Levine model makes the case that truth-bias is advantageous in most everyday circumstances because the probability of being told the truth outweighs that of a lie. Therefore, truth-bias is only detrimental in circumstances where the risks of being lied to dramatically outweigh the likelihood of encountering a truthful message.

The Adaptive Lie Detector Theory (ALIED). Adaptive Lie Detector Theory (ALIED) is based upon the adaptive decision-making perspective and seeks to explain how people judge veracity (Street, 2015). According to ALLIED, we rely on a combination of individuating information and contextual information to make veracity judgments. The ALLIED perspective uses contextual information to explain both truth-bias and lie-bias. It is argued that truth-bias occurs in situations where individuating cues are weak. When this happens, people rely on generalized contextual information, such as that, in general,
people are truthful in everyday conversation (Street et al., 2016). However, a lie-bias would occur if the situation or person in question suggests that there would be an increased likelihood to deceive (Street & Richardson, 2015). For example, if CCTV footage of a suspect at the crime scene were available, we rely on this individuating information instead of the more general contextual information. Since most of the time individuating information is unavailable, an educated guess about general information is used. Additionally, according to ALLIED, good liars can exploit others by building trust and exploiting it. ALLIED is supported by the findings suggesting the less diagnostic the cues, the more the participants used context-general information (specifically, the base-rates of lying) to assess veracity (Street et al., 2016).

Information Manipulation Theory-2. The original Information Manipulation Theory (IMT; McCornack, 1992) aimed to move beyond the typical bald-faced lies (BFLs) and bald-faced truths (BFTs) dichotomy that was predominant among most classic deception researchers.

Information Manipulation Theory accounts for messages somewhere between BFLs and BFTs. According to IMT, messages are considered deceptive if they contain one or more violations of Grice's maxims quantity, quality, relevance, and manner.

Information Manipulation Theory- 2 (IMT2) moved beyond IMT and is a deceptive message production theory. At its core, IMT2 uses Grice's (1989) Cooperative Principle as a foundation. During interactions, messages consist of four maxims, Quantity (the amount of information provided), Quality (The veracity of the information provided), Manner (How information is uttered), and Relation (Relevance of information shared). Unlike traditional approaches to discourse production, IMT2 does not propose that a decision to make a BFL or a BFT comes before the message is initially produced. IMT2 proposes an opportunistic problem-solving approach where people adjust on the fly" in an attempt to reach their desired end goal. IMT2 consists of three theoretical propositions, including (1) intentional states (I.S.), (2) cognitive load (CL), and (3) information manipulation (I.M.).

According to I.S., deception will only occur when considered the most efficient way to achieve one's goal. Additionally, the intention to deceive may arise or decay at any time point during the production of speech. According to the CL proposition, deception is not always more cognitively effortful than truth-telling. Instead, the relative load of truthful and deceptive messages depends on the information activated from memory and contextual factors surrounding the problem (McCornack, 1997; Walczyk, 2013). Additionally, deceptive messages proven successful are easier to repeat (less load-inducing) because they are easier to construct and automatic. According to I.M. propositions, Quality and Quantity violations occur when the most accessible information from memory is considered harmful to achieving one's goal.

Additionally, Quantity violations should be the most common form of deception because one can edit out information from memory that they do not want to be disclosed. The least common type of deception should be Relational violations and should only occur when providing any information that would seem dangerous to achieving one's end goal. According to I.M., when there is a violation of quality, the information given is not "made up" but drawn from real memories. Finally, speakers show behavioral cues to deception when there are no plausible alternatives from memory to construct a deceptive message. In these circumstances, the speaker may take longer than usual to speak, have more pauses, and mention irrelevant information.

Verbal Cues to Deception. Given the limited findings supporting non-verbal cues of deception, a better method for detecting deception may be verbal cues to deception. When investigating the effectiveness of verbal cues of deception, researchers commonly apply one of three verbal veracity tools: Scientific Content Analysis (SCAN), Statement Validity Analysis (SVA), and Reality Monitoring (R.M.).

Scientific Content Analysis (SCAN) The Scientific Content Analysis (SCAN) is a criteria-based verbal veracity tool that is among the most commonly used by criminal investigators (Vrij, 2008). SCAN is used worldwide by federal agencies, law enforcement, and private investigators in many countries, including the United States, Canada, Australia, Belgium, and Israel. According to SCAN, truth-tellers and liars differ in the kind of language displayed in their statements (Sapir, 2005). When using the SCAN procedure, the interviewer asks the examinee to write down as much information as possible about the critical period from the event in question (Sapir, 2005). According to SCAN, truthful written statements consist of 20% of information leading up to the event, 50% of the information about the main event, and 30% of the information should describe what happened following the event. The greater the deviation

from this structure, the more likely it is that the statement is deceptive. Criticisms of SCAN include its ill-defined list of criteria, low inter-rater reliability, and lack of a standardized scoring system (Bogaard et al., 2016). Across several studies, there have been 28 different distinct criteria used in SCAN (Bockstaele, 2008a), 16 criteria (Bockstaele, 2008b, Smith, 2001), and in more recent years, suggestions for 12 criteria version of SCAN (Bogaard, Meijer, Vrij, 2016).

Statement Validity Analysis. Statement Validity Analysis (SVA) is a deception detection technique initially used in both Sweden (Trankell, 1963) and Germany (Undeutsch, 1967) to assess the truthfulness of an individual's account. Originally, SVA was used to determine the credibility of child witnesses' statements about sexual assault allegations (Ruby & Brigham, 1997). SVA is used as evidence in America (Ruby & Brigham, 1997) and some Western European countries, including Germany, the Netherlands, Spain, and Sweden (Vrij, 2008). SVA consists of a combination of several components, including (i) a case-file analysis, (ii) a semi-structured interview, (iii) a Criteria-Based Content Analysis (CBCA), and (iv) a Validity Checklist. A case-file analysis provides the SVA expert with essential information about the witness (e.g., age, cognitive ability), any previous witness statements, and other important case characteristics. The SVA expert conducts a semi-structured interview that is investigative and allows for content to be analyzed via CBCA. Below I will discuss the CBCA and the validity checklist.

Criteria-Based Content Analysis. Criteria-based content analysis (CBCA) is a verbal veracity tool based on the Undeutch (1967) hypothesis, which states that the content of real and imagined memories should fundamentally differ from each other (Steller, 1989). Fabricated statements are thought to contain fewer details and are less vivid than truthful statements based

on one's real memory (Undeutch, 1967). CBCA includes 19 content-based criteria that trained evaluators judge the presence or absence of using a validity checklist. The CBCA consist of two factors: cognitive (1-13) and motivational (14-18) (Köhnken, 1989, 2004). Cognitive factors, including *contextual details, the quantity of details, and descriptions of interactions*, are considered difficult to fabricate, and the presence of them indicates genuine experiences.

Additionally, motivational factors, including questioning one's memory, are considered predictive of genuine experiences. Innocent suspects are thought to be less concerned about impression management than guilty suspects because they are thought to operate according to the just-world fallacy. Therefore, liars avoid adding any information in their statements that they believe is undesirable out of fear of damaging their image (Köhnken, 1999).

According to a meta-analysis by Vrij, 2008 the CBCA discriminates between real and imagined events at above chance levels. Across 20 separate studies, 16 times truth-tellers had higher CBCA scores than liars (Vrij, 2008). However, not all the criteria used on the CBCA are diagnostic measures of deception. In more recent studies, only 15 of the CBCA criteria were diagnostic measures of deception (Amado et al., 2015; Amado et al., 2016).

CBCA is predictive of deception and has high inter-rater reliability on most criteria (Hauch et al., 2017). However, CBCA lacks a strong theoretical foundation because it was constructed via a bottom-up intuitive approach (Sporer, 1997).

Reality Monitoring Approach. An alternative criteria-based approach to deception detection is the reality monitoring approach (R.M.). Like CBCA, R.M. proposes that actual memories differ in fundamental ways compared to imagined events. According to the R.M. approach, memories contain both 'external' and 'internal' features. External features are

considered more indicative of real memory and include features such as perceptual information, sensory processes, and contextual information. On the other hand, internal features are based on reasoning, imagination, and thought processes. Statements about imagined events are believed to describe feelings during the time of the event and contain more cognitive operations, such as, "I must have had my coat on, as it was freezing that night" (Johnson et al., 1993; Johnson & Raye, 1981, 1998).

The R.M. approach contains benefits over some of the previously mentioned approaches. Unlike CBCA, R.M. includes indicators for both perceived actual and imagined events. Furthermore, R.M. helps distinguish between a truthful and false statement 60—70% of the time (Masip et al., 2005; Vrij, 2008).

While the vast majority of R.M. studies have involved deception detection of previous events, there is a recent interest in distinguishing true and false statements about intended events (Granhag & Mac Giolla, 2014). These studies suggest that specific criteria, including *cognitive operations*, *clarity*, re*constructability*, and *realism*, predicted the intention to deceive. As a whole, the R.M. criteria failed to predict true and false intentions (Giolla et al., 2019). Even though both R.M. and CBCA have above-chance lie detection accuracy, the use within the field has been limited (Bogaard et al., 2016).

Memory Errors

A false statement made by a person may result from a deliberate attempt to deceive, as described above. However, a false statement may also be a function of an honest but imperfect memory system. Previously I described theories of lie detection that primarily focused on the distinction between honest intended accurate statements and dishonest intended false statements. However, the source monitoring approach builds upon what was discussed earlier using the

reality monitoring approach and deal with the distinction between honest/true and honest /false statements. According to the Source Monitoring Framework (SMF), our memories about episodic content do not get stored with a specific source label. Instead, stored memories contain memory characteristics, including perceptual details (e.g., lighting, color), affective, spatial, semantic, and temporal information used to remember the source of a retrieved memory (Johnson et al., 1993). For instance, if auditory information in memory matches, then you will likely associate that statement with that individual (Ferguson et al., 1992). However, even if the event in question is correctly recalled, the assigned source to that memory can be incorrect. For example, you might accurately recall hearing about Frederick Bartlett's classic War of the Ghosts study. You might remember sitting in a large lecture hall with white walls and the person who told you had a New York accent. Using this information, you might infer that your cognitive psychology professor was the source of this memory because they have a New York accent. However, a source-monitoring error may occur if an extroverted student with a similar accent is in that class. These kinds of source-monitoring errors are more common when two sources share similar characteristics (e.g., high-pitched voice, brown hair) than when sources are dissimilar to one another (e.g., child versus middle-aged man).

When discussing S.M., sources are broken up into either *internal sources* (e.g., dream, imagination) or external sources (e.g., my doctor, teacher, and dad), thus creating three S.M. decisions. The first S.M. decision occurs when you decide between two different external sources. For instance, you must decide if you heard advice regarding plumbing from your dad (who is not handy) or your plumber. The second S.M. decision is when a decision is made between two internal sources. An example of this decision is deciding whether a memory is derived from an internal thought or a dream. Finally, the third S.M. decision determines whether

the memory came from an internal source (e.g., imagination) and an external source (e.g., newspaper). This S.M. decision is similar to the questions raised by the previously described reality monitoring. (Johnson, 1981). Unlike reality monitoring, source monitoring is not limited to dichotomous distinctions (real or imagined), and the "source" component has unlimited options. (Johnson, Hashtroudi, & Lindsay, 1993). In these circumstances, memories of imagined events are less vivid and contain less perceptual, temporal, and spatial information than real events (Johnson et al., 1993). However, you might attribute a memory of an event to be your imagination if it was from a place you know you never visited (Johnson, 2000).

Since source attributions are made rapidly using a match-to-average- characteristic heuristic, we may unknowingly provide a false account even when we intend to tell accurate information. One factor contributing to S.M. errors is the amount of attention paid toward actively monitoring the memory source. These judgment processes are influenced by factors including, but not limited to, attention, social context, and type of task (Johnson, 2000). If the intended goal is to entertain, we may not direct the required attention to self-monitor accurately. For example, the criterion used to recall the author of an article you recently read is laxer when trying to impress non-academic friends than when talking to scholars at an academic conference. Other factors that influence S.M. errors' frequency include ease of generating information (Johnson & Raye, 1981), quality of initially encoded memory, and temporal distance between the encoding of memory and retrieval of the source. Newer memories are thought to be more vivid than older memories and contain more detailed memory characteristics.

Additionally, similar to event memory, the temporal distance between encoding and retrieving the source makes it more difficult to pinpoint the event's specific timeframe (Neisser 1982). Finally, failure to recognize the memory source occurs in similar ways that people fail to

recognize stimuli during a perception task. When the source is unfamiliar, or the perception details are degraded, there is an increase in S.M. errors (Mitchell, 2009)

Accuracy on an S.M. task is increased when the evaluated statement comes from an expected schema-consistent source (Mather et al., 1999; Ehrenberg & Klauer, 2005). However, there is debate whether people are better at remembering schema-consistent sources. When deciding between two sources, people are biased to guess a source that fits their expectations of who would make that statement (Bayen et al., 2000). Additionally, the MPT model suggests that people use newly learned information to make source attribution for information they could not remember. In a study by Ehrenberg and Klaur (2005), participants learned either 75% positive/25% negative or 25% positive/75% negative traits about two new sources. When asked to make source attributions, participants were more likely to respond with the "good guy" if the unremembered trait is positive. This problem is compounded when the original information is poorly encoded because people are more likely to guess the source in those situations.

Source-monitoring errors can have serious real-world consequences partly due to our vulnerability to suggestion. Misleading suggestions about previously-witness events cause people to make more S.M. errors than those who were not biased (Loftus, 1979). Additionally, in a classic study by Hyman, Husband, and Billings (1995), participants recalled a false childhood experience after being told to think about that suggested false event in-between interviews. Finally, S.M. errors can lead to eyewitness identification errors during situations where memory misattribution takes place.

The Three Distinctions of Veracity Judgments

Statements made by an individual may be true or may be false. When the statement is made, the person may intentionally convey true or false information or unintentionally convey

true or false information (see Figure 1). The traditional lie detection (e.g., truth-default theory) and lie production (IMT2) literature primarily described the distinction between intentionally truthful and intentionally false statements (top two circles). Meanwhile, the traditional source monitoring framework primarily focuses on describing how people distinguish between intentionally truthful and unintentionally false statements (top left circle versus bottom right circle). However, even with these distinctions accounted for, there is a critical piece to fully understanding how people evaluate statement credibility. Specifically, there are cases where one might know that a statement is false but may be unsure whether the false statement is intentional or unintentional (top right circle versus bottom right circle). One example of this situation is perjury trials, where known false statements must be evaluated for intent and materiality (Tiersma, 1989). Currently, only research on alibi generation even indirectly investigates this distinction, and even this research is only tangentially related. In the following section, I will discuss the alibi generation and evaluation research findings. Additionally, I propose a testable set of propositions based on lie detection theories that we will test through a series of empirical studies.

Alibis. Within the legal system, when suspects are accused of crimes, they often have an opportunity to provide an alibi. In legal settings, an alibi refers to "a defense that places the defendant at the relevant time of the crime in a different place than the scene involved and so removed from there as to render it impossible for one to be the guilty party" (Nolan, 1990). More simply, alibis provide the suspect with an opportunity to prove that they were at a different location when the crime occurred and thus dismissed as a suspect (Olson & Wells, 2004).

Courts and prosecutors believe that generating alibis should be relatively easy to construct (Williams v. Florida, 1970). However, research suggests that generating a convincing alibi can prove challenging even under seemly ideal circumstances (Olson & Wells, 2004). Strong alibis increase the perceived credibility and honesty of the suspect. However, weak alibis are met with skepticism and may increase the suspect's perceived guiltiness (Olson and Wells, 2012). In approximately two-thirds of cases that involved a DNA exoneration, innocent suspects' alibis were not believed, and "the prosecution used weak alibis" as incriminating evidence (Olsen & Wells, 2004). Given these findings, it is important to understand how alibis are generated, evaluated and how memory errors influence alibis.

How Do People Know When a False Statement is a Lie?

One question that has not been thoroughly investigated is how people classify known false statements as an intentional lie or an "honest" mistake. This question is important because it helps us better understand a missing piece to effective interpersonal communication. Additionally, this distinction has critical legal implications, including how jurors evaluate witness creditability and perjury evidence. This research aims to test a framework that explains this unanswered research question. The framework is based on the deception detection, deception production, source monitoring, and alibi literature described above.

Non-verbal Cues. Traditional lie detection literature suggests that non-verbal cues (e.g., averting their gaze) are relied upon when judging statement veracity (Global Deception Research Team, 2006). While non-verbal cues to deception are weakly correlated with actual deceit, they may provide insight into how people classify known false statements as a deliberate lie or an honest mistake (e.g., memory error). I anticipate that in interactions where little-to-no non-verbal

cues to deception (e.g., averting gaze) are present, people are more likely to judge a known false statement as an honest mistake. My rationale for this prediction is based upon findings supporting truth-default theory (TDT). According to TDT, communication between people is believed to be honest unless a specific "trigger" to deception is activated (Levine, 2014). Therefore, I anticipate that if non-verbal cues to deception accompany a blatantly false statement, the suspicion threshold will be crossed, and the statement will be classified as a deliberate lie. This is a bit of an extension to TDT because the statements I am examining are known falsehoods. So rather than assuming truth in the absence of a specific "trigger" the implication is that what people are assuming is a lack of intentionality in the absence of the trigger.

Motivational. A 2nd factor that I predict influences the classification of known false statements is the perceived motivation behind the false statement. I base this prediction on findings in the alibi generation literature that suggests *person corroborating evidence* influences an alibi's credibility and believability (Culhane & Hosch, 2004). Sworn testimony by a family member (e.g., motivated other) is regarded as less credible and believable compared to that of a stranger (nonmotivated other). Furthermore, 1/3rd of jury eligible participants believed that individuals who have close social relationships would be more likely to lie (Marion & Burke, 2013).

Given these findings, I predict false statements that are about a close friend or family member are more likely to be perceived to be deceptive than ones made about an acquaintance or stranger. However, this effect may be influenced by how self-serving the false statement was. Suppose a false statement was made about a family member, and that statement protected the family member from a controversial situation. In that case, that statement might be more likely to

be judged as a deliberate lie. Both IMT2 and ALLIED suggest that deceptive communication is typically done as a last resort. Turner et al. (1975) proposed a set of motivations for everyday deceptive behavior, including to a) to save face, (b) to manage relationships, (c) to exploit, (d) to avoid tension or conflict, and (e) to control situations. Given this, I expect people will classify statements as deliberate lies when the evaluator perceives that the person has one of these motives (or perhaps some other motive) for intentionally lying. However, if the false statements do not benefit the speaker or their close family or friends, that statement will be more likely to be judged as an honest mistake.

Memorability of statement. A 3rd factor that I predict will influence how false statements are evaluated is the false statement's perceived memorability. Previous research suggests that people use a *metacognitive strategy* to reject or accept a previous event's occurrence. Events are confidently rejected if (1) the event in question is deemed memorable, and (2) they fail to retrieve a feature about the event (Strack & Bless, 1994). Metacognitive awareness of others' memory may influence how false statements are perceived. I predict if a false statement is perceived as memorable, it will be judged as a deliberate lie. If a false statement is a type that people judge to be less memorable, people should be more likely to judge it to be an honest mistake. A limitation of this approach may be the faulty belief that laypeople have regarding memory. In a recent public opinion survey, 46.6% of laypeople believe memory is permanent and does not change, and over 60% believe that memory operates like a video camera (Simons and Chabris, 2011).

Reputational. The 4th factor that I propose influences how people classify deliberate lies and honest mistakes is based on the ALLIED perspective of deception detection. According to

ALLIED, when there is no individuating information available, veracity judgments are based upon contextual information. Known false statements may be evaluated similarly. If a usually trustworthy source makes a false statement, I propose that statement will be judged as an honest mistake. However, if your cousin has a history of lying, you are more likely to judge their statement as a deliberate lie. It is important to note that these proposed findings might interact with the presence of non-verbal cues to deception.

The Current Study

My dissertation aims to determine the factors that are associated with classifying false statements as honest mistakes or deliberate lies. This dissertation contains two experiments with varying methodology that seek to answer this question. Experiment 1 is a diary study where participants monitored their daily conversations for the occurrence of false statements. Participants then classified these statements as either deliberate lies or an honest mistakes. Furthermore, I was interested in determining why participants made that classification through a series of follow-up questions. The second experiment was a semi-structured interview study where participants described experiences where they encountered someone telling them a deliberate lie and an honest mistake. These experiments are primarily exploratory, but I proposed hypotheses based upon the literature reviewed above.

H1: Participants will perceive the presence of non-verbal cues to deception as a trigger to deceptive behavior. Therefore, non-verbal cues to deception will be mentioned in a greater proportion of deliberate lies than honest mistakes.

H2: Participants will use perceived motive to classify false statements. I expect participants will be more likely to mention motive while describing deliberate lies than honest

mistakes. I also expect participants will mention that the transgressor of a false statement did not have motivation to lie while describing honest mistakes.

H3: Participants will use information regarding the perceived memorability of the false statement in order to classify the statement as an honest mistake or a deliberate lie. Difficult to remember statements will be more likely to be classified as an honest mistake. Meanwhile, easy to remember statements will be more likely to be classified as a deliberate lie.

H4: Participants will use information regarding the perceived reputation of the transgressor to classify false statements as an honest mistake or a deliberate lie. False statements made by those with a history of lying will be evaluated as a deliberate lie. Furthermore, false statements made by those with no history of lying will be evaluated as an honest mistake. It is also possible that more generic information about having a disreputable character may be associated with judging a false statement to be a deliberate lie.

RQ: My primary research question is to identify the reasons that statements are classified as either an honest mistake or deliberate lie.

Experiment 1

Method

Participants. Thirty-eight University of Arkansas undergraduates (Age: *M* = 21.87 years; SD = 2.34) participated in the following study in exchange for extra credit. The majority of participants were female (n = 37, 97.4 %). The majority of participants were Caucasian (68.4 %), followed by Latinx (15.8%), Asian-Pacific Islander (7.9 %), and Black (7.9 %).

Survey Day One. Participants received a Qualtrics survey that explained the purpose of this study is to better understand how people interpret false statements in everyday communication. More specifically, we asked participants to monitor their daily conversations for 25 days. Each day, participants reported if anyone made a false statement to them within the past 24 hours. If they indicated that a false statement was made, participants answered follow-up questions, including whether they believed that the statement was a deliberate lie or an honest mistake. To explain to participants what are counted as a deliberate lie, we instructed participants "*For a statement to be a deliberate lie, the statement must be explicitly false, the person making the statement must know its false at the time they make it, and the person making the statement must intend to deceive another person."* To explain to participants what should be counted as an honest mistake, we instructed participants *"For a statement to be an honest mistake, the statement must be explicitly false, but there was no intention to deceive and the person who said it did not realize it was false at the time they said it. For instance, someone may simply make a memory error or be confused about the facts or may have just misspoken".*

Additionally, participants read examples of a deliberate lie and an honest mistake. After reading the examples, participants described the difference between a deliberate lie and an honest mistake in their own words. After that, participants were told, "Over the next 25 days, we will be asking you to monitor your conversations for both deliberate lies and honest mistakes. We will send you a survey each night and ask you a series of questions about your previous day". Participants answered a demographic questionnaire and provided an email address to receive the follow-up survey.

Follow-up Daily Survey. Participants completed a daily Qualtrics survey each night until they have responded on 25 different days. Each night participants answered the following

question, "*In the past 24 hours did someone say something to you that you believe is false (whether or not you realized it was false at the time)? The statement could be either a deliberate lie or an honest mistake. So, say "Yes" if any false statement was made to you in the past 24 hours."*. If participants respond yes, they described the false statement in as much detail as possible. After describing the false statement, participants answered a series of follow-up questions to gather additional information about the false statement (see Table 1). After answering the follow-up questions, participants were asked, "if there were any other false statements in the past 24 hours that they would like to report?". If the participant answered "yes", the procedure described above was repeated. After responding to the survey for 25 days, participants were thanked and debriefed about the nature of the study. We continued to follow up with participants until they completed 25 days of surveys or participants no longer responded. That is, if the participant forgot to respond to the survey one day, we just continued data collection with that participant until all 25 days' worth of data was collected. Thus, some participants completed the experiment after 25 days, but some participants took somewhat longer to complete the experiment. On average participants completed the survey within 29 days (median = 26 days).

Coding. All responses were coded independently by three pairs of trained coders. Once coding was complete, we calculated intercoder reliability (Kappa), and I resolved all disagreements with a discussion between the trained coders. To start the coding process, I developed a coding scheme for four theory-based coding categories related to my primary hypotheses. Before the coding proceeded, I identified two additional categories, Verbal Cues, and Failed Prediction. Therefore, the primary coding categories are as follows, 1. *Non-verbal cues* (e.g., averting their gaze, increased movement of feet and legs); 2. Verbal Cues (e.g., admits

lack of memory, spontaneous corrections), 3. *Memorability* (whether the false statement was perceived to be easy or difficult to remember); 4. *Motivational or Lack of Motivation (e.g., Self-Serving);* 5. *Reputational* (e.g., (a) Prior lying/truthfulness, (b) general good / bad character); and 6. Epistemic Paucity/Failed Prediction -- made a statement, but it's not the kind of thing that a person could know for sure (e.g., it's going to rain tonight; Trump is going to win the election) (See Table 1 for examples).

Table 1

Primary coding categories and examples based on participants responses

Categories	Examples
Non-verbal Cues	1. I know she didn't want to talk about it, wouldn't make eye contact, and she always has a bit of a different inflection in her voice when lying
	2. I could tell by her body language that she was lying and hiding something. She knows she isn't allowed to have the full tub of Nutella in her room, so she was trying to hide it.
Verbal Cues	1. I knew it was false because shortly after his proclamation, he said he was kidding.
	2. She told me the truth like 2 seconds later to correct herself.
Motivational	1. He was trying to be hurtful, had the intent of misleading me so that I would not like her as much, but she helped more than he did. Probably wanted to make himself look better by comparison.
	2. There's no reason for her to lie about that specific fact, she just had the wrong information.

Table 1 Cont.

Primary coding categories and examples based on participants responses

Reputational	1. I knew it was a deliberate lie because I saw her texts, and she has lied in the past. 2. Historically speaking, this individual is likely to not pay others back even after promising to.
Memorability	1. I thought it was an honest mistake because she simply forgot that she made the appointment with the advisor for that time. Therefore, she thought she did not know who called her. 2. It was just a forgetful mistake.
Failed Prediction	1. I honestly believe this person truly thought Trump would win again. As this person also explained that they were so certain he would that they weren't concerned about voting. Clearly this isn't true because no one can predict that. 2. The speaker had checked all their tracking links for their incoming packages and truly believed that no packages would arrive for them until Sunday. Upon receiving the packages, they realized that their statement had been false.

Additionally, each primary coding category contained several subcategories that were coded for. To code the data, coders read participant's responses about what caused them to classify a false statement as a deliberate lie or an honest mistake. After reading the participant's response, the coders marked 0 (No) and 1 (Yes) on a separate data sheet to identify the presence or absence of the six primary coding categories and the respective subcategories.

Results

Reliability of Coding Taxonomy. To assess inter-rater reliability, I computed Cohen's kappa statistic using SPSS version 27. For all three pairs of coders there was moderate agreement (pair 1: $\kappa = 0.400$, $p < 0.000$; pair 2: $\kappa = 0.754$, $p < 0.000$; pair 3: $\kappa = 0.698$, $p < 0.001$). According to Cohen's system for interpreting Kappa, the kappa for pair 1 indicates a fair amount of agreement between the coders, and the kappas for pair 2 and 3 indicate substantial agreement

between coders. [1]After computing kappa, I sought clarification from coders on their disagreements and resolved any differences.

Likelihood of Reporting False Statements. In this experiment, participants monitored their everyday conversations for false statements for twenty-five days. The vast majority of participants (n = 35, 92%) responded to the daily survey at least twenty times. There were 898 responses, and participants mentioned false statements on 163 days (15%). On average each participant reported 4.3 false statements (median = 4; range 0 – 12). Of the 163 false statements that were described, participants interpreted 81 to be honest mistakes (49.7%) and 82 to be deliberate lies (50.3%). I was interested in whether participants changed their response patterns as the diary study persisted. A correlation suggested that participants were less likely to report false statements as the diary study persisted, $r = -.161$, $p < .001$. To further investigate this finding, I split study days into quartiles and conducted a one-way ANOVA to test any differences. Participants reported more false statements during the first quartile (days 1-6: $M = 30.16$) than the second (days 7-14: $M = 13.27$), third (days 14-20: $M = 11.48$), or fourth quartiles (days 21 – survey end: $M = 14.85$), $F(3, 890) = 12.73$, $p < .001$. There was no other difference in response patterns. This effect may reflect fatigue with the process of keeping a daily diary.

Perceived Trustworthiness. I was interested in whether participants adjusted their future trust depending on whether participants perceived a deliberate lie or an honest mistake. I used the survey item, "*After this interaction, how likely are you to trust the person who made the false statement?"* as the dependent variable on the following analysis. An independent samples t-test suggested that participants are more likely to trust those who made an honest mistake ($M = 4.16$,

[1] "Cohen suggested the Kappa result be interpreted as follows: values ≤ 0 as indicating no agreement and 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41– 0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement." (McHugh, 2012)

SD = 1.08, *SE* = .12) than those who told a deliberate lie (*M* = 3.26, *SD* = 1.35, *SE* = .15), t (161) = - 4.704, p < .001.

Perceived Seriousness. I was interested in whether participants perceived the seriousness of a deliberate lie or an honest mistake differently. I used the survey item**,** "How serious do you view the severity of the false statement*?"* as the dependent variable on the following analysis. An independent samples t-test suggested that participants view deliberate lies as more serious (*M* = 1.95, SD = 1.05, SE = .11) than honest mistakes (M = 1.37, SD = .64, SE = .07), t (161) = 4.247, p < .001.

Certainty. I was interested in whether participants' certainty that the statement was false statement was different for a deliberate lie as opposed to honest mistakes. I used the survey item, "How certain are you that the statement was false*?"* as the dependent variable on the following analysis. An independent samples t-test suggested that there was no significant difference in the certainty of a deliberate lie (*M* = 1.95, *SD* = 1.05, *SE* = .11) and honest mistake (*M* = 1.37, *SD* = .64, *SE* = .07), t (161) = 4.247, p < .001.

Primary Analyses. To test my four primary hypotheses, I was interested in whether non-verbal cues, memorability, reputation cues, and motivation cues differentiated deliberate lies and honest mistakes. To investigate this, I analyzed my data using Fisher's exact test (Bower, 2003) to look at associations between these categories and whether it occurred more often in honest mistakes or deliberate lies. I used Fisher's exact test because it is appropriate for small cell sizes and is a more conservative test statistic that minimizes type one error.

Non-verbal Cues. My first hypothesis was that participants would describe *non-verbal cues to deception* more often for deliberate lies than honest mistakes. The presence of *non-verbal cues to deception* was significantly more associated with deliberate lies (*M* = .20, *SD* = .40) than

honest mistakes (M = .01, SD = .11), p < .001. I followed up this analysis by investigating if there were specific non-verbal cues most associated with deliberate lies. I found that *voice change* was associated with deliberate lies (M = .07, SD = .26) more than an honest mistake (M = .0, SD = .0), p = .028. No other non-verbal cues had a significant association (See table 2 for the full list of categories and analyses)

Table 2
Experiment One. Descriptive and test statistics for Non-verbal cues and the appropriate subcategories

Categories	N	*M*	*SD*	Exact
Non-verbal Cues	17	.10	.31	p < .001
Deliberate Lie	16	.20	.40	
Honest Mistake	1	.01	.11	
Pauses	1	.01	.08	p = 1.00
Deliberate Lie	1	.01	.11	
Honest Mistake	0			
Poor eye contact	2	.01	.11	p = .497
Deliberate Lie	2	.02	.16	
Honest Mistake	0			
Nervous looking	4	.03	.16	p = .620
Deliberate Lie	3	.04	.19	
Honest Mistake	1	.01	.11	
Voice Change	6	.04	.19	p = .028
Deliberate Lie	6	.07	.26	
Honest Mistake	0	0	0	
Rate of Speaking	0	0	0	
Deliberate Lie	0	0	0	
Honest Mistake	0	0	0	

Motivational. My second hypothesis was that participants would be more likely to describe motivational factors when describing deliberate lies than honest mistakes. Motivational cues were significantly associated with deliberate lies (M = .50, SD = .56) more than honest mistakes (M = .27, SD = .48), p = .003. Participants were more likely to mention motivation to lie when describing a deliberate lie (M = .38, SD = .49) than an honest mistake (M = .03, SD =

.16). Also, participants were more likely to mention no motivation to lie when describing an honest mistake (M = .13, SD = .34) than a deliberate lie (M = .01, SD = .11). Finally, participants were more likely to mention the statement was self-serving when describing a deliberate lie (M = .17, SD = .38) than an honest mistake (M = .0, SD = .0). There were no other types of motivational cues that were associated with honest mistakes or deliberate lies (See table 3 for the full list of categories and analyses).

Table 3

Experiment One. Descriptive and test statistics for Motivation cues and the appropriate subcategories

Categories	N	M	SD	Exact
Motivation to lie (general)	63	.31	.46	p = .003
Deliberate Lie	41	.50	.50	
Honest Mistake	22	.27	.33	
Motivated to lie?	33	.20	.40	p < .001
Deliberate Lie	31	.38	.49	
Honest Mistake	2	.03	.16	
Not motivated to lie?	21	.13	.34	p < .001
Deliberate Lie	1	.01	.11	
Honest Mistake	20	.25	.43	
Self-serving	14	.09	.28	p < .001
Deliberate Lie	14	.17	.38	
Honest Mistake	0	0	0	
Detrimental to self	2	.01	.11	p = .497
Deliberate Lie	2	.02	.16	
Honest Mistake	0	0	0	

Memorability. My third hypothesis was that participants would use information about perceived memorability to classify statements as deliberate lies and honest mistakes. Perceived memorability was significantly associated more with honest mistakes (M = .28, SD = .45) than deliberate lies (M = .07, SD = .26), p < .001. More specifically, participants associated difficult to remember statements with honest mistake (M = .14, SD = .35) more than a deliberate lie (M =

.0, *SD* = .0). There were no other memorability cues associated with honest mistakes or deliberate lies (See table 4 for the full list of categories and analyses)

Table 4
Experiment One. Descriptive and test statistics for Memorability cues and the appropriate subcategories

Categories	N	M	SD	Exact
Perceived Memorability	29	.18	.38	$p < .001$
Deliberate Lie	6	.07	.26	
Honest Mistake	23	.28	.45	
Difficult to remember	11	.07	.25	$p < .001$
Deliberate Lie	0	0	0	
Honest Mistake	11	.14	.35	
Easy to remember	5	.03	.17	$p = .367$
Deliberate Lie	4	.05	.22	
Honest Mistake	1	.01	.11	

Reputation. My fourth hypothesis was that participants would use information about the perceived reputation of the transgressor to classify statements as deliberate lies and honest mistakes. Overall, Reputation cues were only marginally more likely to be mentioned for deliberate lies (*M* = .17, *SD* = .33) than honest mistakes (*M* = .07, *SD* = .26), *p* = .093. However, participants were more likely to mention the history of prior lying when describing a deliberate lie (*M* = .12, *SD* = .33) than an honest mistake (*M* = .01, *SD* = .11). No other reputational cues were associated with honest mistakes or deliberate lies (See table 5 for the full list of categories and analyses)

Table 5
Experiment One. Descriptive and test statistics for Reputational cues and the appropriate subcategories

Categories	N	M	SD	Exact
Reputational	20	.12	.33	$p = .093$
Deliberate Lie	14	.17	.33	
Honest Mistake	6	.07	.26	
Prior Lying	11	.07	.25	$p = .009$
Deliberate Lie	10	.12	.33	
Honest Mistake	1	.01	.11	
Prior Truthfulness	2	.01	.11	$p = 1.00$
Deliberate Lie	1	.01	.11	
Honest Mistake	1	.01	.11	
Good Character	1	.01	.08	$p = .497$
Deliberate Lie	0			
Honest Mistake	1	.01	.11	
Bad Character	3	.02	.14	$p = .245$
Deliberate Lie	3	.04	.19	
Honest Mistake	0	0	0	

Exploratory Analyses. Based on my preliminary coding, I identified the presence of two additional categories, failed predictions and verbal cues. I aimed to test whether these categories helped differentiate deliberate lies and honest mistakes.

Failed Predictions. A failed prediction is a false statement in which the person made a claim about a future event, and that future event did not occur (e.g., "It is going to rain tonight"). Because the future is never perfectly predictable, it is reasonable that people would treat these false statements as honest mistakes rather than deliberate lies. Failed predictions were significantly associated with honest mistakes ($M = .32$, $SD = .47$) more than deliberate lies ($M = .04$, $SD = .19$), $p < .001$.

Verbal Cues. The presence of Verbal cues was not significantly associated with deliberate lies ($M = .09$, $SD = .28$) or honest mistakes ($M = .10$, $SD = .31$), $p = .454$.

Furthermore, no specific verbal cues were associated with honest mistakes or deliberate lies (See table 6 for full analyses)

Table 6
Experiment One. Descriptive and test statistics for Verbal cues and the appropriate subcategories

Categories	N	M	SD	Exact
Verbal Cues	17	.10	.31	$p = .454$
Deliberate Lie	7	.09	.28	
Honest Mistake	10	.12	.33	
Admits lack of memory	6	.04	.19	$p = .117$
Deliberate Lie	1	.01	.11	
Honest Mistake	5	.06	.24	
Spontaneous Corrections	8	.05	.22	$p = .495$
Deliberate Lie	3	.04	.19	
Honest Mistake	5	.06	.24	

Experiment 2

Results from Experiment 1 revealed that participants used different cues to differentiate between deliberate lies and honest mistakes. Participants associated *perceived memorability* with honest mistakes and both *non-verbal cues* and *motivation* with deliberate lies. In Experiment 2, I attempted to replicate these using a novel format. Second, I designed Experiment 2 to allow myself to investigate whether any of the hypothesized predictors influenced the initial decision and/or final decision about statement veracity. Finally, in Experiment 2, participants are asked to describe false statements from any point of their lives. Potentially false statements that are most memorable may differ in nature compared to those that were described in Experiment 1.

Method

Participants. Fifty University of Arkansas undergraduates (Age: $M = 20.56$ years; $SD = 1.92$) participated in the following study in exchange for extra credit or fulfilling a course

requirement. The majority of participants were female (n = 36,72 %). The majority of participants were Caucasian (76%), followed by Latinx (14%), Asian-Pacific Islander (6%), and Black (5.4%).

Procedure. Participants completed a semi-structured interview via Zoom, talking to a researcher about their life experiences. The interviewer, a psychology undergraduate research assistant, began the session by sending a Qualtrics survey containing a demographic questionnaire. After completing the questionnaire, the research assistant asked for the participant's consent to record the interview and subsequently began the task. After this, the interviewer read the participant the following instructions:

In the following task, we will ask you a series of questions regarding deliberate lies and honest mistakes. For a statement to be a deliberate lie, the statement must be explicitly false, the person making the statement must know it's false at the time they make it, and person making the statement must intend to deceive another person. An example of a deliberate lie includes, "Employee tells their boss that they are missing work because they are sick. Really, they are not sick but just wanted to go to a baseball game." For a statement to be an honest mistake, the statement must be explicitly false, but there was no intention to deceive, and the person who said it did not realize it was false at the time they said it. For instance, someone may simply make a memory error or be confused about the facts or may have just misspoken. An example of an honest mistake includes, "Jim tells his spouse that he never met Jenny. It turns out that Jim and Jenny went to school together, but Jim just honestly forgot."

Following this, the interviewer asked participants to "please describe the difference between a deliberate lie and an honest mistake?". If participants could not describe the difference between a deliberate lie and an honest mistake, the research assistant described the difference for

the participant and re-asked the question. The interviewer then asked participants a series of questions regarding false statements and deliberate lies. The first question asked by the interviewer was, "Please describe in as much detail as possible a false statement that was made to you that you believed was a deliberate lie.". After answering this question, participants were asked follow-up questions to gather additional information about the described event. The follow-up questions were, 1.*"How certain are you that the statement was false?",* 2.*"What convinced you that the statement was false?"*, 3. "How seriously do you view the severity of the false statement?", 4. *"How difficult do you think it was for the other person who told the false statement to recall the correct information?",* and 5. "After this interaction, how likely are you to trust the person who made the false statement?". These follow-up questions were asked after each of the following questions. The 2nd question asked by the interviewer was, "*Please describe in as much detail as possible a false statement that was made to you that you believed was an honest mistake.".* After this, the interviewer asked, *"Was there a time where you believed a false statement was a deliberate lie, but later turned out to be an honest mistake? If so, please describe in as much detail as possible"*. The final question asked by the interviewer was, *"Was there a time where you believed a false statement was an honest mistake, but later turned out to be a deliberate lie? If so, please describe in as much detail as possible."*. After the completion of the interview, the interviewer thanked and debriefed the participants. On average the interview took approximately 15 minutes.

Coding. All videos were coded independently by trained coders (blind to the hypotheses) using the same coding scheme developed for experiment one. Once coding was complete, I calculated intercoder reliability (Kappa), and I resolved all disagreements with a discussion between the trained coder.

Results

Reliability of Coding Taxonomy. To assess inter-rater reliability, I computed Cohen's kappa statistic using SPSS version 27. Two pairs of coders there had fair agreement (pair 1: $\kappa = 0.339$, $p < 0.000$; pair 2: $\kappa = 0.249$, $p < 0.000$; pair 3: $\kappa = 0.615$, $p < 0.001$) (Landis, Koch, 1977). Due to the relatively low kappa, I calculated the overall agreement rate for all three coders. Overall, the agreement between the coders were high (pair 1 = 94.8%; pair 2 = 96.2%; pair 3= 98.11%). One explanation for the paradox is the relatively low prevalence rate of the coded behaviors (Viera and Garret, 2005)

After computing the reliability ratings, I sought clarification from the coders on disagreements and resolved any differences.

Primary Analyses. To test the four primary hypotheses, I conducted four logistic regressions using non-verbal cues, memorability, reputational cues, and motivation cues as separate outcome variables. Additionally, I conducted two exploratory logistic regressions using verbal cues and failed predictions as different outcome variables. In the regression model, initial statement belief (Deliberate lie, Honest mistake), final statement belief (Deliberate lie, Honest mistake), and the interaction term were predictor variables. The reference category for both predictor variables was deliberate lies.

Non-verbal Cues. My first hypothesis was that participants would describe non-verbal cues more often for deliberate lies than honest mistakes. To test this, I conducted a logistic regression on the first dependent variable, the proportion of non-verbal cues reported. A test of the full model was significant, $\chi^2 = 11.45$, df = 3, $p = .010$. SPSS took six iterations to converge on a solution. The goodness of fit test was not significant, p = 1.00. The effect of final statement belief was significant, b = - 1.43, S.E. = .68, Wald $\chi^2(1) = 5.51$, $p = .019$, OR = .20, CI 95%

[.053 - .77]. Deliberate lies (*M* = .18, *SE* = .04) were more likely to include non-verbal cues to deception than statements about honest mistakes (*M* = .05, *SE* =.02). Both initial statement belief and the interaction term were not significant predictors, $p > .05$.

Motivation. My second hypothesis was that participants would be more likely to describe motivational factors when describing deliberate lies than honest mistakes. To test this, I conducted a logistic regression on the second dependent variable, the proportion of motivational cues reported. A test of the full model was significant, $\chi_2 = 28.79$, df = 3, $p < .001$. SPSS took five iterations to converge on a solution. The goodness of fit test was not significant, p = 1.00. The effect of final statement belief was significant, b = - 2.09 S.E. = .67, Wald $\chi_2(1) = 9.83$, $p = .002$, OR = .12, CI 95% [.03 - .46]. Participants were more likely to mention motivational cues while describing statements perceived as deliberate lies (*M* = .37, *SE* = .05) than statements about honest mistakes (*M* = .07, *SE* =.03). Both initial statement belief and the interaction term were not significant predictors, $p > .05$.

Memorability. My third hypothesis was that participants would use information about perceived memorability to classify statements as deliberate lies and honest mistakes. To test this, I conducted a logistic regression on the third dependent variable, the proportion of memorability cues reported. A test of the full model was significant, $\chi_2 = 17.71$, df = 3, $p < .001$ SPSS took 7 iterations to converge on a solution. The goodness of fit test was not significant, p = 1.00. The effect of final statement belief was marginally significant, b = 1.90 S.E. = 1.10, Wald $\chi_2(1) = 2.98$, $p = .084$, OR = 6.68, CI 95% [.78 – 57.70]. Statements describing honest mistakes (*M* = .21, *SE* = .04) were marginally more likely to include perceived memorability than statements about deliberate lies (*M* = .07, *SE* =.04). The effect of initial statement belief was marginally

significant, b = 1.90 S.E. = 1.10, Wald $\chi_2(1) = 2.98$, p = .084, OR = 6.68, CI 95% [.78 – 57.70]. Statements describing honest mistakes (*M* = .21, *SE* = .04) were marginally more likely to include perceived memorability than statements about deliberate lies (*M* = .07, *SE* =.03). To further test my hypothesis, I conducted a logistic regression on the dependent variable, the proportion of difficult to remember memory cues reported. When I included the interaction term, the model failed to converge. Therefore, my regression model included initial statement belief (Deliberate lie, Honest mistake) and final statement belief (Deliberate lie, Honest mistake). A test of the full model was significant, $\chi_2 = 21.07$, df = 2, p < .001 SPSS took 7 iterations to converge on a solution. The goodness of fit test was not significant, p = 1.00. The effect of final statement belief was significant, b = 2.23 S.E. = .76, Wald $\chi_2(1) = 8.27$, *p* = .004, OR = 9.30, CI 95% [2.03 – 42.52]). Statements describing honest mistakes (*M* = .15, *SE* = .04) were more likely to include perceived memorability than statements about deliberate lies (*M* = .02, *SE* =.01). The effect of initial statement belief was significant, b = 1.75 S.E. = .67, Wald $\chi_2(1) = 6.87$, *p* = .009, OR = 5.73, CI 95% [1.55 – 21.12]. Statements describing honest mistakes (*M* = .14, *SE* = .04) were more likely to include difficult to remember memory cues than statements about deliberate lies (*M* = .03, *SE* =.02). There were not enough reported instances where participants described easy to remember cues to conduct further analyses.

Reputation. My fourth hypothesis was that participants would use information about the perceived reputation of the transgressor to classify statements as deliberate lies and honest mistakes. To test this, I conducted a logistic regression on the fourth dependent variable, the proportion of reputational cues reported. A test of the full model was not significant, $\chi_2 = 3.389$, df = 3, *p* = .335. SPSS took 7 iterations to converge on a solution. The goodness of fit test was not significant, *p* = 1.00. Initial statement belief, final statement belief, and the interaction were

not significant predictors, $p > .05$. To further investigate this effect, I conducted a logistic regression on the proportion of prior history of lying reported. A test of the full model was not significant, $\chi_2 = 6.040$, df = 3, $p = .110$ SPSS was unable to converge on a solution. The goodness of fit test was not significant, $p = 1.00$. Initial statement belief, final statement belief, and the interaction were not significant predictors, $p > .05$. Since the regression was unable to converge, I tested the model without the interaction term. A test of the full model was not significant, $\chi_2 = 1.184$, df = 2, $p = .553$. SPSS took 6 iterations to converge on a solution. The goodness of fit test was not significant, $p = .16$. Initial statement belief and final statement belief were not significant predictors, $p > .05$

Verbal Cues. I conducted a logistic regression on the fifth dependent variable, the proportion of verbal cues reported. A test of the full model was not significant, $\chi_2 = 4.34$, df = 3, p = .227 SPSS took 6 iterations to converge on a solution. The goodness of fit test was not significant, p = 1.00. The interaction was marginally significant, b = 2.12 S.E. = 1.11, Wald $\chi_2(1) = 3.66$, p = .056, OR = 8.35, CI 95% [.95 - 73.44]. Initial statement belief and final statement belief were not significant predictors, p > .05.

Failed Prediction. A test of the full model was not significant, $\chi_2 = 19.638$, df = 3, $p < .001$ SPSS took 20 iterations and was unable to converge on a solution. The goodness of fit test was not significant, p = 1.00. Initial belief, end belief, and the interaction were not significant.

Table 7

Experiment Two. Descriptive statistics for primary coding categories

Categories	*N*	*M*
Non-verbal Cues		
Deliberate Lie	12	.24
Honest Mistake	2	.04
Deliberate lie to Honest Mistake	3	.06
Honest Mistake to Deliberate Lie	6	.12
Verbal Cues		
Deliberate Lie	7	.14
Honest Mistake	6	.12
Deliberate lie to Honest Mistake	2	.04
Honest Mistake to Deliberate Lie	3	.06
Perceived Memorability		
Deliberate Lie	1	.07
Honest Mistake	15	.28
Deliberate lie to Honest Mistake	6	.12
Honest Mistake to Deliberate Lie	6	.12
Motivation to lie (general)		
Deliberate Lie	17	.34
Honest Mistake	4	.08
Deliberate lie to Honest Mistake	3	.06
Honest Mistake to Deliberate Lie	20	.40
Reputational		
Deliberate Lie	4	.08
Honest Mistake	1	.02
Deliberate lie to Honest Mistake	5	.10
Honest Mistake to Deliberate Lie	4	.08
Failed Prediction		
Deliberate Lie	0	.00
Honest Mistake	8	.16
Deliberate lie to Honest Mistake	1	.02
Honest Mistake to Deliberate Lie	0	.00

General Discussion

In this study, I was interested in how people determine whether a false statement should be classified as an honest mistake or a deliberate lie. The research question asked for this dissertation was novel and serves as an important contribution to the growing literature on the broader field of deception detection. Previous research involving deception detection has

exclusively focused on the lie vs. truthful statement dichotomy. In this dissertation, I reviewed the deception detection, alibi generation, and source monitoring literature. I then identified a framework based on an initial set of variables that I hypothesized would impact the classification of deliberates lies vs. honest mistakes. I developed an exploratory coding scheme containing six primary categories and several subcategories to test this proposed framework. I used this coding scheme to classify participants' responses across two experiments.

The first experiment I conducted was a diary study that had participants monitor their everyday conversations for false statements. Participants who reported hearing a false statement were asked a series of questions, including being asked to describe their rationale for deciding whether the false statement was an honest mistake or a deliberate lie. Trained research assistants coded the participants' responses for the presence of the six primary categories and their respective subcategories.

In Experiment 2, participants completed a semi-structured interview with a trained research assistant about previous times that the participant was lied to. Interviewers asked participants to describe events when someone told them deliberate lies and honest mistakes. Interviewers asked participants to explain the rationale for why they decided that the statement was false. Afterward, trained research assistants watched video recordings of the interviews and coded the participants' responses for the presence of the six primary categories and their respective subcategories

Both experiments revealed that participants used different cues to classify statements as a deliberate lie or an honest mistake. In the general discussion, I will discuss how these findings reconciled with my primary hypotheses. Additionally, I will discuss the general pattern of data, including how participants emphasized different categories for both kinds of false statements. I

will describe the theoretical implications of the findings. Finally, I will discuss the limitations of dissertations and the future directions for research investigating the honest mistake vs. deliberate lie dichotomy.

Non-verbal Cues. Previous research investigating lie detection suggests that people use non-verbal cues when determining if communication is deceptive (Global Deception Team, 2006). Since these findings were confined to the true-false dichotomy, I was interested in testing whether participants used the same cues to make decisions about the intent of false statements. I hypothesized participants would indicate the presence of non-verbal cues when describing deliberate lies more than describing honest mistakes. Results from both Experiment 1 and Experiment 2 supported this hypothesis. Participants described non-verbal cues as their reason why the false statement was a deliberate lie more than they did to describe the false statement was an honest mistake. Interestedly, there was not one specific type of non-verbal cue that was mentioned by participants in more than 7% of deliberate lies. Participants were often vague while describing non-verbal cues, and when more specific, the specific cue varied from participant to participant. This result was initially surprising because I anticipated that there would be a clear pattern to the non-verbal cues reported as deceptive.

However, after looking closer at the literature, I found a rationale for why specific non-verbal cues are sparsely reported. As stated above, past research investigating the truth/lie dichotomy suggests that non-verbal cues are believed by others to indicate lying. However, there is little agreement regarding specific non-verbal cues being associated with deceptive behavior (Global Deception Team, 2006). Specific non-verbal cues, except for gaze aversion were believed to indicate lying by under 40% of participants. Therefore, the pattern of responding that I had in my study aligns with the previous literature that invested the truth-lie distinction. Overall, the result

from both experiments provides evidence that non-verbal cues are used when people make judgments about deliberate lies but are rarely used when classifying a statement as an honest mistake. This finding may have significant legal implications regarding cases of perjury. In these situations, a juror may only look for non-verbal cues associated with deception. Therefore, signs of nervousness including but not limited to looking away may cause an increase in perceived guiltiness. In the future direction section, I will discuss how non-verbal cues can be investigated further when paired with motive and perceived memorability.

Motivation. Previous research investigating motive has primarily been studied in the context of alibi generation. According to this literature, alibis are considered less credible and less believable when the alibi contains a perceived motive to protect oneself or others. In this dissertation, I was interested in whether the perceived motive was used to classify known false statements as a deliberate lie or an honest mistake. I hypothesized that participants would use information about the perceived motive when classifying false statements. In Experiment 1, participants described *motive* as the reason for their classification for approximately 40% of all false statements. While describing deliberate lies, roughly half of all participants described *motivation to lie* as the reason why they made that classification. Compared to all categories coded for during this dissertation, *motivation to lie* was the category that was mentioned by the most participants. This finding was replicated in Experiment 2.

Additionally, *motive* was also associated with deliberate lies. In Experiment 1, participants described the transgressor of the false statement as *not having motivation to lie* in over 20% of honest mistakes. However, Experiment 2 did not replicate this finding. Potentially participants are less likely to recall details about motive when describing honest mistakes that occurred in the distant past.

Taken together, the findings from both experiments suggest *motive* plays an influential role is how people classify known false statements. Motive influenced both the perception of deliberate lies and honest mistakes. In particular, deliberate lies were commonly considered *self-serving.* An informal investigation of the data suggests that self-serving statements were often used as a way to save face and to manage relationships. These reasons are consistent with Turner et al. (1975) proposed a set of motivations for everyday deceptive behavior. In my future directions, I will discuss ways researchers can test whether motive mediates the relationship that other categories have on the classification of false statements.

Memorability. Memory errors are considered a special kind of honest mistake that involves an unintended failure to recall information. These circumstances may be explained by a source-monitoring mistake (Johnson et al. 1993). In my dissertation, I investigated whether information regarding the perceived memorability of false statements is used to decide if the statement was a memory error (honest mistake) or a deliberate lie. I predicted that participants would be likely to describe honest mistakes as *difficult to remember*. Additionally, I anticipated participants would describe the deliberate lies as *easy to remember*. In both experiments, the results only partially supported my hypothesis. Participants mentioned an honest mistake being *difficult to remember* 14% of the time. However, participants rarely classified a statement as a deliberate lie because a statement was judged as easy to remember. Overall, this finding suggests that participants will make use of perceived memorability only when the false statement is deemed as *difficult to remember*. One challenge related to perceived memorability is that most people believe that memory operates as a video recorder (Simons and Chabris, 2011). Even so, this finding suggests that memorability of a perceived false statement does influence whether a statement is classified as an honest mistake. In my future directions I will discuss what

implications this has for future research investigating the honest mistake vs. deliberate lie distinction.

Reputation. According to traditional lie detection research, people use the base rate of prior lying to judge whether a statement is likely to be true or false (Street, 2014). For instance, people may expect politicians to lie to them, but not a Catholic nun. In this study, I hypothesized that the perceived reputation of the transgressor of the false statement would influence whether it was classified as a deliberate lie or an honest mistake. The results from my dissertation only partially supported this hypothesis. If the participant classified the statement as a deliberate lie, they were more likely to mention that the transgressor has a previous *history of lying.* However, when participants classified the statement as an honest mistake, they did not use previous *history of truthful behavior* to classify the statement. Furthermore, general characteristics about individuals, such as having *good* or *bad character,* did not influence the categorization of false statements. Potentially, this result is explained because people operate under a truth-default state (Levine, 2014). Therefore, only information that makes people question the truth-default state would be deemed relevant.

Overall Categorization. In addition to comparing the factors that differentiate how people compare deliberate lies and honest mistakes, I wanted to determine which of my proposed categories are most prevalent when describing deliberate lies and honest mistakes. Participants described their rationale for categorizing a false statement as either a deliberate lie or an honest mistake. Afterward, my team of research assistants coded the participants' responses. After investigating the data from both experiments, I identified two distinct patterns.
Participants were most likely to mention *motive, non-verbal cues, history of lying* as the reason they classified a false statement as a deliberate lie. Meanwhile, participants used information

regarding the *perceived memorability* and *lack of motive* to classify a statement as an honest mistake. Overall, these two patterns of responding are relatively unique. Across the four categories that I hypothesized influences the categorization of false statements, only *motive* influences the categorization of both deliberate lies and honest mistakes. In the next section, I will discuss the theoretical ramifications of the findings discussed above and then discuss future research to further investigate the two patterns described above.

Theoretical Implication

The findings from this study have important theoretical ramifications. First, it provides support for a subset of theories I discussed during the introduction. Second, these findings serve as a building block to further refine a theoretical model that can disentangle how people classify deliberate lies and false statements.

Deliberate Lies. According to truth-default theory, communication operates under the assumption that others are inherently honest unless a "trigger" is activated (Levine, 2014). In experiment one, participants reported relatively few false statements during the 25-day diary study. This finding is aligned with the general idea that people are "truth-biased" when evaluating the veracity of statements. Additional evidence for truth-default theory is that only 49% of the described false statements were initially perceived as a false statement. Of these false statements, approximately 65% were classified as deliberate lies, while 35% were classified as honest mistakes. One potential distinguishing factor for whether a deliberate lie is perceived right away is the presence of non-verbal cues. In experiment one, participants described the presence of non-verbal cues in 23.1% of deliberate lies that were known right away. Meanwhile, participants described the presence of non-verbal cues in 13.3% of deliberate lies that were discovered after the fact. A similar pattern of results occurred in Experiment 2 when I compared

statements that participants knew were deliberate lies to deliberate lies participants originally thought were honest mistakes. Taken together, this pattern of results demonstrates further evidence for TDT. Non-verbal cues operated as a "trigger" that caused participants to be more likely to conclude the false statement was deliberate.

Another finding that has theoretical ramifications is that participants rely on using *motive* more often to classify false statements as deliberate lies than honest mistakes. Across both experiments' participants mentioned *motive* over 40% of the time as their reason that they classified a statement as a deliberate lie. One explanation for this finding is that participants in this study were sensitive to situational factors associated with lying. This finding suggests that when a statement is classified as a false statement, people use Turner et al. (1975) proposed a set of motivations for everyday deceptive behavior, including to a) to save face, (b) to manage relationships, (c) to exploit, (d) to avoid tension or conflict, and (e) to control situations as a way to classify a statement as a deliberate lie or honest mistake.

Honest Mistakes. Unlike deliberate lies, *perceived memorability* was an important predictor of honest mistakes. When a statement was assumed to be *difficult to remember,* participants classified the false statement as an honest mistake. Additionally, in Experiment 2 participants were more likely to mention perceived memorability when discussing statements participants *always knew were honest mistakes* (28%) than statements participants *initially thought was a deliberate lie* that later turned out to be an honest mistake. This finding suggests that participants often think about the memorability of a false statement once it is encountered. Potentially these findings are limited to people who believe in the fallibility of memory. Previous research suggests that over 60% believe that memory operates like a video camera (Simons &

Chabris, 2011). Therefore, *perceived memorability* may be more influential if participants are aware of the fallibility of one's memory.

The findings discussed above have important theoretical ramifications. These findings suggest that people may focus on distinctly different cues to classify deliberate lies and honest mistakes. When determining whether a statement is a deliberate lie, participants primarily described the *motive of the transgressor* or focused on classical cues to deception such as the presence of *non-verbal cues*. Meanwhile, honest mistakes were classified as a result of either a *memory error* or the *lack of motive*. These findings are the first of their kind to determine how people classify deliberate lies and honest mistakes. In my future directions, I will describe ways to further investigate the pattern of responding described above.

Limitations.

There are several limitations to this dissertation that limit some of the conclusions I can make. First, this dissertation was completed the covid-19 pandemic. A subset of participants was likely practicing proper social distancing while participating in the diary study. In this case, I anticipate that participants communicate less with others and thus, this study may underestimate the number of false statements reported on an everyday basis. Another limitation cause by the covid-19 pandemic is that the interview study was completed remotely using Zoom. Participants might have been less comfortable sharing details with my interviewers than they would in a face-to-face situation.

There are some limitations regarding the coding used in the study. First, not all the coder pairs had high inter-rater reliability. In the future, I will use a single pair of coders for all statements. Finally, this study did not manipulate the presence or absence of the categories associated with the false statements. Because of this, I cannot tell the likelihood that participants

are about to identify any perceived categories associated with false statements. In the section below, I will discuss future directions for this line of research that aims to be better able to make causal conclusions.

Future Directions and Conclusions.

The two studies discussed above provided evidence that a unique pattern of cues is used to classify a statement as either a deliberate lie or an honest mistake. In the future, I would like to conduct experimental research where participants make forced-choice honest mistake/deliberate judgments while manipulating cues to honest mistake/deliberate judgments in novel scenarios. In this study, I identified the importance of non-verbal cues, motive, and memorability on classifying false statements. Deliberate lies were most associated with *motive* and *non-verbal cues*. In contrast, honest mistakes were associated with a *lack of motive* and *memorability*. In a future study, I can manipulate *motive, non-verbal cues,* and *memorability* and test whether participants use the same criteria described by participants in this study. Additionally, this approach will allow me to investigate causality and the strength of each predictor. Potentially, variables such as *memorability* only influence the statement classification in the absence of triggers such as *motive* and *non-verbal cues*. Afterwards, I want to extend my research to investigate if a similar pattern of results occurs for sentencing perjury cases. In conclusion, this study was the first of its kind to investigate the honest mistake vs. deliberate lie distinction. I discovered that people use different cues to classify honest mistakes and deliberate lies.

References

Akehurst, L., Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology*, *10*(6), 461-471.

Alowibdi, J. S., Buy, U. A., Philip, S. Y., Ghani, S., & Mokbel, M. (2015). Deception detection in Twitter. *Social network analysis and mining*, *5*(1), 32.

Amado, B. G., Arce, R., & Fariña, F. (2015). Undeutsch hypothesis and Criteria Based Content Analysis: A meta-analytic review. *The European Journal of Psychology Applied to Legal Context*, *7*(1), 3-12.

Amado, B. G., Arce, R., Farina, F., & Vilarino, M. (2016). Criteria-Based Content Analysis (CBCA) reality criteria in adults: A meta-analytic review. *International Journal of Clinical and Health Psychology*, *16*(2), 201-210.

Baran, K. K., & Ruby, R. I. (1997). Perjury. *American Criminal Law Review*, *35*, 1035-1060.

Bayen, U. J., Nakamura, G. V., Dupuis, S. E., & Yang, C. L. (2000). The use of schematic knowledge about sources in source monitoring. *Memory & Cognition*, *28*(3), 480-500.

Bok, S. (1999). *Lying: Moral choice in public and private life*. Vintage.

Bond Jr, C. F., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and social psychology Review*, *10*(3), 214-234.

Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Scientific Content Analysis (SCAN) cannot distinguish between truthful and fabricated accounts of a negative Event. *Frontiers in psychology*, *7*, 243.

Bogaard, G., Meijer, E. H., Vrij, A., & Merckelbach, H. (2016). Strong, but wrong: Lay people's and police officers' beliefs about verbal and nonverbal cues to deception. *PloS one*, *11*(6), e0156615.

Bower, K. M. (2003, August). When to use Fisher's exact test. In *American Society for Quality, Six Sigma Forum Magazine*(Vol. 2, No. 4, pp. 35-37).

Burgoon, J. K., Blair, J. P., & Strom, R. E. (2008). Cognitive biases and nonverbal cue availability in detecting deception. *Human Communication Research*, *34*(4), 572-599.

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory & cognition*, *20*(3), 277-290.

Burke, T. M., Turtle, J. W., & Olson, E. A. (2007). Alibis in criminal investigations and trials.

Carson, T. L. (2006). The definition of lying. *Noûs*, *40*(2), 284-306.

Christ, S. E., Van Essen, D. C., Watson, J. M., Brubaker, L. E., & McDermott, K. B. (2009). The contributions of prefrontal cortex and executive control to deception: evidence from activation likelihood estimate meta-analyses. *Cerebral cortex*, *19*(7), 1557-1566.

Culhane, S. E., & Hosch, H. M. (2004). An Aiibi Witness' Influence on Mock Jurors' Verdicts 1. *Journal of Applied Social Psychology*, *34*(8), 1604-1616.

Culhane, S. E. (2005). *Changing your alibi: Current law enforcement, future law enforcement, and layperson beliefs and behaviors*. The University of Texas at El Paso.

Culhane, S. E., Hosch, H. M., & Kehn, A. (2008). Alibi generation: Data from U.S. Hispanics and U.S. non-Hispanic whites. *Journal of Ethnicity in Criminal Justice*, *6*(3), 177-199.

Culhane, S. E., & Hosch, H. M. (2012). Changed alibis: Current law enforcement, future law enforcement, and layperson reactions. *Criminal Justice and Behavior*, *39*(7), 958-977.

DePaulo, B. M., & Pfeifer, R. L. (1986). On-the-Job Experience and Skill at Detecting Deception 1. *Journal of Applied Social Psychology*, *16*(3), 249-267.

DePaulo, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological bulletin*, *111*(2), 203.

DePaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of personality and social psychology*, *70*(5), 979.

DePaulo, B. M., Charlton, K., Cooper, H., Lindsay, J. J., & Muhlenbruck, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review*, *1*(4), 346-357.

DePaulo, B. M., & Kashy, D. A. (1998). Everyday lies in close and casual relationships. *Journal of personality and social psychology*, *74*(1), 63.

Dysart, J. E., & Strange, D. (2012). Beliefs about alibis and alibi investigations: A survey of law enforcement. *Psychology, Crime & Law*, *18*(1), 11-25.

Ehrenberg, K., & Klauer, K. C. (2005). Flexible use of source information: Processing components of the inconsistency effect in person memory. *Journal of Experimental Social Psychology*, *41*(4), 369-387.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*(1), 88-106.

Ekman, P., & O'Sullivan, M. (1991). Who can catch a liar?. *American psychologist*, *46*(9), 913.

Ekman, P., O'Sullivan, M., & Frank, M. G. (1999). A few can catch a liar. *Psychological science*, *10*(3), 263-266.

Ekman, P. (2011) *Testimony to the Subcommittee on Investigations and Oversight Hearing – Behavioral Science and Security: Evaluating TSA's SPOT Program*

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *Psychology of learning and motivation*, *44*, 145-200.

Evans, J. R., Michael, S. W., Meissner, C. A., & Brandon, S. E. (2013). Validating a new assessment method for deception detection: Introducing a Psychologically Based Credibility Assessment Tool. *Journal of Applied Research in Memory and Cognition*, *2*(1), 33-41.

Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and aging*, *7*(3), 443.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. Mcgraw-Hill Book Company.

Global Deception Research Team. (2006). A world of lies. *Journal of cross-cultural psychology*, *37*(1), 60-74.

Gordon, A. K., & Miller, A. G. (2000). Perspective differences in the construal of lies: Is deception in the eye of the beholder?. *Personality and Social Psychology Bulletin*, *26*(1), 46-55.

Grice, H. P. (1989). *Studies in the Way of Words*. Harvard University Press.

Hancock, J. T., Thom-Santelli, J., & Ritchie, T. (2004, April). Deception and design: The impact of communication technology on lying behavior. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 129-134).

Hauch, V., Sporer, S. L., Masip, J., & Blandón-Gitlin, I. (2017). Can credibility criteria be assessed reliably? A meta-analysis of criteria-based content analysis. *Psychological Assessment*, *29*(6), 819.

Hyman Jr, I. E., Husband, T. H., & Billings, F. J. (1995). False memories of childhood experiences. *Applied cognitive psychology*, *9*(3), 181-197.

Inbau, F. E., Reid, J. E., & Buckley, J. P. Jayne (2001). *Criminal interrogation and confessions*.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological review*, *88*(1), 67.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological bulletin*, *114*(1), 3.

Johnson, M. K., & Raye, C. L. (1998). False memories and confabulation. *Trends in cognitive sciences*, *2*(4), 137-145.

Kassam, K. S., Gilbert, D. T., Swencionis, J. K., & Wilson, T. D. (2009). Misconceptions of memory: the Scooter Libby effect. *Psychological Science*.

Kassin, S. M., & Norwick, R. J. (2004). Why people waive their Miranda rights: The power of innocence. *Law and Human Behavior*, *28*(2), 211-221.

Kelly, A., Carroll, M., & Mazzoni, G. (2002). Metamemory and reality monitoring. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *16*(4), 407-428.

Kim, H. Y. (2017). Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative dentistry & endodontics*, *42*(2), 152.

Köhnken, G. (1989). Behavioral correlates of statement credibility: Theories, paradigms, and results. In *Criminal behavior and the justice system* (pp. 271-289). Springer, Berlin, Heidelberg.

Köhnken, G. (1999). Statement validity assessment. *preconference programme of applied courses "Assessing credibility ", organised by the European Association of Psychology and Law, Dublin, Ireland*.

Köhnken, G. (2004). Statement validity analysis and the 'detection of the truth'. *The detection of deception in forensic contexts*, 41-63.

Kraut, R. E., & Poe, D. B. (1980). Behavioral roots of person perception: The deception judgments of customs inspectors and laymen. *Journal of Personality and Social Psychology*, *39*(5), 784.

Lampinen, J. M., Faries, J. M., Neuschatz, J. S., & Toglia, M. P. (2000). Recollections of things schematic: The influence of scripts on recollective experience. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *14*(6), 543-554.

Leins, D. A., & Charman, S. D. (2016). Schema reliance and innocent alibi generation. *Legal and Criminological Psychology*, *21*(1), 111-126

Lewis, M., & Saarni, C. (Eds.). (1993). *Lying and deception in everyday life*. Guilford press.

Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: development and validation of an interview faking behavior scale. *Journal of applied psychology*, *92*(6), 1638.

Levine, T. R., Park, H. S., & McCornack, S. A. (1999). Accuracy in detecting truths and lies: Documenting the "veracity effect". *Communications Monographs*, *66*(2), 125-144.

Levine, T. R., Kim, R. K., Sun Park, H., & Hughes, M. (2006). Deception detection accuracy is a predictable linear function of message veracity base-rate: A formal test of Park and Levine's probability model. *Communication Monographs*, *73*(3), 243-260.

Levine, T. R., Kim, R. K., & Hamel, L. M. (2010). People lie for a reason: Three experiments documenting the principle of veracity. *Communication Research Reports*, *27*(4), 271-285.

Levine, T. R. (2014). Truth-default theory (TDT) a theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*(4), 378-392.

Levine, T. R., Clare, D. D., Green, T., Serota, K. B., & Park, H. S. (2014). The effects of truth-lie base rate on interactive deception detection accuracy. *Human Communication Research*, *40*(3), 350-372.

Loftus, E. F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition*, *7*(5), 368-374.

Mac Giolla, E., Ask, K., Granhag, P. A., & Karlsson, A. (2019). Can reality monitoring criteria distinguish between true and false intentions?. *Journal of Applied Research in Memory and Cognition*, *8*(1), 92-97.

Mann, S., Vrij, A., & Bull, R. (2004). Detecting True Lies: Police Officers' Ability to Detect Suspects' Lies. *Journal of applied psychology*, *89*(1), 137.

Marion, S. B., & Burke, T. M. (2013). False alibi corroboration: Witnesses lie for suspects who seem innocent, whether they like them or not. *Law and Human Behavior*, *37*(2), 136.

Masip, J., Garrido, E., & Herrero, C. (2004). Defining deception. *Anales de psicologia*.

Masip, J., Sporer, S. L., Garrido, E., & Herrero, C. (2005). The detection of deception with the reality monitoring approach: A review of the empirical evidence. *Psychology, Crime & Law*, *11*(1), 99-122.

Masip, J., Blandón-Gitlin, I., de la Riva, C., & Herrero, C. (2016). An empirical test of the decision to lie component of the Activation-Decision-Construction-Action Theory (ADCAT). *Acta Psychologica*, *169*, 45-55.

Mather, M., Johnson, M. K., & De Leonardis, D. M. (1999). Stereotype reliance in source monitoring: Age differences and neuropsychological test correlates. *Cognitive Neuropsychology*, *16*(3-5), 437-458.

McCornack, S. A., & Parks, M. R. (1986). Deception detection and relationship development: The other side of trust. *Annals of the International Communication Association*, *9*(1), 377-389.

McCornack, S. A. (1992). Information manipulation theory. *Communications Monographs*, *59*(1), 1-16.

McCornack, S. A. (1997). The generation of deceptive messages: Laying the groundwork for a viable theory of interpersonal deception.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, *22*(3), 276-282.

Meissner, C. A., & Kassin, S. M. (2002). "He's guilty!": Investigator bias in judgments of truth and deception. *Law and human behavior*, *26*(5), 469-480.

Mitchell, K. J., & Johnson, M. K. (2009). Source monitoring 15 years later: what have we learned from fMRI about the neural mechanisms of source memory?. *Psychological bulletin*, *135*(4), 638.

Morris, M. R., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012, February). Tweeting is believing? Understanding microblog credibility perceptions. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*(pp. 441-450).

Olson, E. A., & Wells, G. L. (2004). What makes a good alibi? A proposed taxonomy. *Law and Human Behavior*, *28*(2), 157-176.

Olson, E. A., & Charman, S. D. (2012). 'But can you prove it?'–examining the quality of innocent suspects' alibis. *Psychology, Crime & Law*, *18*(5), 453-471.

O'Sullivan, M. (2003). The fundamental attribution error in detecting deception: The boy-who-cried-wolf effect. *Personality and Social Psychology Bulletin*, *29*(10), 1316-1327.

O'Sullivan, M., & Ekman, P. (2004). 12 The wizards of deception detection. *The detection of deception in forensic contexts*, 269.

Park, H.S, & Levine, T. (2001). A probability model of accuracy in deception detection experiments. *Communication Monographs*, *68*(2), 201-210.

Park, H. S., & Levine, T. R. (2017). The effects of truth–lie base-rates on deception detection accuracy in Korea. *Asian Journal of Communication*, *27*(5), 554-562.

Patterson, J., & Kim, P. (1991). *The day America told the truth: What people really believe about everything that really matters*. Prentice Hall.

Porter, S., Woodworth, M., & Birt, A. R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and human behavior*, *24*(6), 643-658.

Ritchie, Hannah. 2016. "Read all about it: The Biggest Fake News Stories of 2016." http://www.cnbc.com/2016/12/30/read-all-about-it-the-biggest-fake-news-stories-of-2016.html.

Ruby, C. L., & Brigham, J. C. (1997). The usefulness of the criteria-based content analysis technique in distinguishing between truthful and fabricated allegations: A critical review. *Psychology, Public Policy, and Law*, *3*(4), 705.

Schacter, D. L. (1999). The seven sins of memory: insights from psychology and cognitive neuroscience. *American psychologist*, *54*(3), 182.

Serota, K. B., Levine, T. R., & Boster, F. J. (2010). The prevalence of lying in America: Three studies of self-reported lies. *Human Communication Research*, *36*(1), 2-25.

Silverman, Craig. 2016. "Here are 50 of the Biggest Fake News Hits on Facebook from 2016." BuzzFeed, https://www.buzzfeed.com/craigsilverman/top-fake-news-of-2016?utm_term=. iqzA7qdxG#.ihMA0kVRx.

Serota, K. B., & Levine, T. R. (2015). A few prolific liars: Variation in the prevalence of lying. *Journal of Language and Social Psychology*, *34*(2), 138-157.

Silverman, C., & Singer-Vine, J. (2016). Most Americans who see fake news believe it, new survey says. *BuzzFeed News*, *6*.

Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative survey of the U.S. population. *PloS one*, *6*(8), e22757.

Sorensen, R. (2007). Bald-faced lies! Lying without the intent to deceive. *Pacific Philosophical Quarterly*, *88*(2), 251-264.

Sporer, S. L. (1997). The less travelled road to truth: Verbal cues in deception detection in accounts of fabricated and self-experienced events. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *11*(5), 373-397.

Sporer, S. L. (2016). Deception and cognitive load: Expanding our horizon with a working memory model. *Frontiers in Psychology*, *7*, 420.

Steller, M. (1989). Recent developments in statement analysis. In *Credibility assessment* (pp. 135-154). Springer, Dordrecht.

Strack, F., & Bless, H. (1994). Memory for nonoccurrences: Metacognitive and presuppositional strategies. *Journal of Memory and Language*, *33*(2), 203.

Strange, D., Dysart, J., & Loftus, E. F. (2014). Why errors in alibis are not necessarily evidence of guilt. *Zeitschrift für Psychologie*, *222*(2), 82.

Street, C. N. (2015). ALIED: Humans as adaptive lie detectors. *Journal of Applied Research in Memory and Cognition*, *4*(4), 335-343.

Street, C. N., Bischof, W. F., Vadillo, M. A., & Kingstone, A. (2016). Inferring others' hidden thoughts: Smart guesses in a low diagnostic world. *Journal of Behavioral Decision Making*, *29*(5), 539-549.

Stokke, A. (2013). Lying and asserting. *The Journal of philosophy*, *110*(1), 33-60.

Suroo, R. (1998, August 30). New Gore Probe Centers on Specific Issue of whether He Lied. *The Washington Post*. Retrieved from https://www.washingtonpost.com/archive/politics/1998/08/30/new-gore-probe-centers-on-specific-issue-of-whether-he-lied/099ad8da-52f0-476e-a505-25a1edc4b456/

Tiersma, P. M. (1989). The language of perjury: literal truth, ambiguity, and the false statement requirement. *S. Cal. L. Rev.*, *63*, 373.

Tourangeau, R. (1999). Remembering what happened: Memory errors and survey reports. In *The science of self-report* (pp. 41-60). Psychology Press.

Trankell, A. (1963). *Vittnespsykologins Arbedsmetoder*. Stockholm: Liber.

Turner, R. E., Edgley, C., & Olmstead, G. (1975). Information control in conversations: Honesty is not always the best policy. *Kansas Journal of Sociology*, 69-89.

Undeutsch, U. (1967). Beurteilung der Glaubhaftigkeit von Aussagen [Veracity assess- ment of statements]. In U. Undeutsch (Ed.), *Handbuch der Psychologie: Vol. 11. Forensische Psychologie* (pp. 26–181). Go¨ttingen, Germany: Hogrefe

Verschuere, B., Köbis, N. C., Bereby-Meyer, Y., Rand, D., & Shalvi, S. (2018). Taxing the brain to uncover lying? Meta-analyzing the effect of imposing cognitive load on the reaction-time costs of lying. *Journal of applied research in memory and cognition*, *7*(3), 462-469.

Vrij, A., & Semin, G. R. (1996). Lie experts' beliefs about nonverbal indicators of deception. *Journal of nonverbal behavior*, *20*(1), 65-80.

Vrij, A. (2005). Criteria-Based Content Analysis: A Qualitative Review of the First 37 Studies. *Psychology, Public Policy, and Law*, *11*(1), 3.

Vrij, A. (2008). Detecting lies and deceit: Pitfalls and opportunities (2nd ed.). Chichester, England: John Wiley and Sons

Vrij, A. (2008). Nonverbal dominance versus verbal accuracy in lie detection: A plea to change police practice. *Criminal Justice and Behavior*, *35*(10), 1323-1336.

Vrij, A., Fisher, R., Mann, S., & Leal, S. (2008). A cognitive load approach to lie detection. *Journal of Investigative Psychology and Offender Profiling*, *5*(1-2), 39-43.

Vrij, A., Mann, S., Leal, S., & Fisher, R. (2010). 'Look into my eyes': Can an instruction to maintain eye contact facilitate lie detection?. *Psychology, Crime & Law*, *16*(4), 327-348.

Vrij, A., Leal, S., Mann, S., & Fisher, R. (2012). Imposing cognitive load to elicit cues to deceit: Inducing the reverse order technique naturally. *Psychology, Crime & Law*, *18*(6), 579-594.

Vrij, A., & Fisher, R. P. (2016). Which lie detection tools are ready for use in the criminal justice system?. *Journal of Applied Research in Memory and Cognition*, *5*(3), 302-307.

Vrij, A., Fisher, R. P., & Blank, H. (2017). A cognitive approach to lie detection: A meta-analysis. *Legal and Criminological Psychology*, *22*(1), 1-21.

Wagenaar, A. C., Maybee, R. G., & Sullivan, K. P. (1988). Mandatory seat belt laws in eight states: A time-series evaluation. *Journal of Safety Research*, *19*(2), 51-70.

Walbert, D. F. (1970). The effect of jury size on the probability of conviction: an evaluation of Williams v. Florida. *Case W. Res. L. Rev.*, *22*, 529.

Walczyk, J. J., Igou, F. D., Dixon, L. P., & Tcholakian, T. (2013). Advancing lie detection by inducing cognitive load on liars: a review of relevant theories and techniques guided by lessons from polygraph-based approaches. *Frontiers in psychology*, *4*, 14.

Walczyk, J. J., Harris, L. L., Duck, T. K., & Mulay, D. (2014). A social-cognitive framework for understanding serious lies: Activation-decision-construction-action theory. *New Ideas in Psychology*, *34*, 22-36.

Wells, G. L., Memon, A., & Penrod, S. D. (2006). Eyewitness evidence: Improving its probative value. *Psychological science in the public interest*, *7*(2), 45-75.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In *Advances in experimental social psychology* (Vol. 14, pp. 1-59). Academic Press.

Zubiaga, A., Liakata, M., Procter, R., Bontcheva, K., & Tolmie, P. (2015). Towards detecting rumours in social media. *arXiv preprint arXiv:1504.04712*.

Zuckerman, M., Koestner, R., & Driver, R. (1981). Beliefs about cues associated with deception. *Journal of Nonverbal Behavior*, *6*(2), 105-114.

Appendix

Appendix A: Supplementary Figure

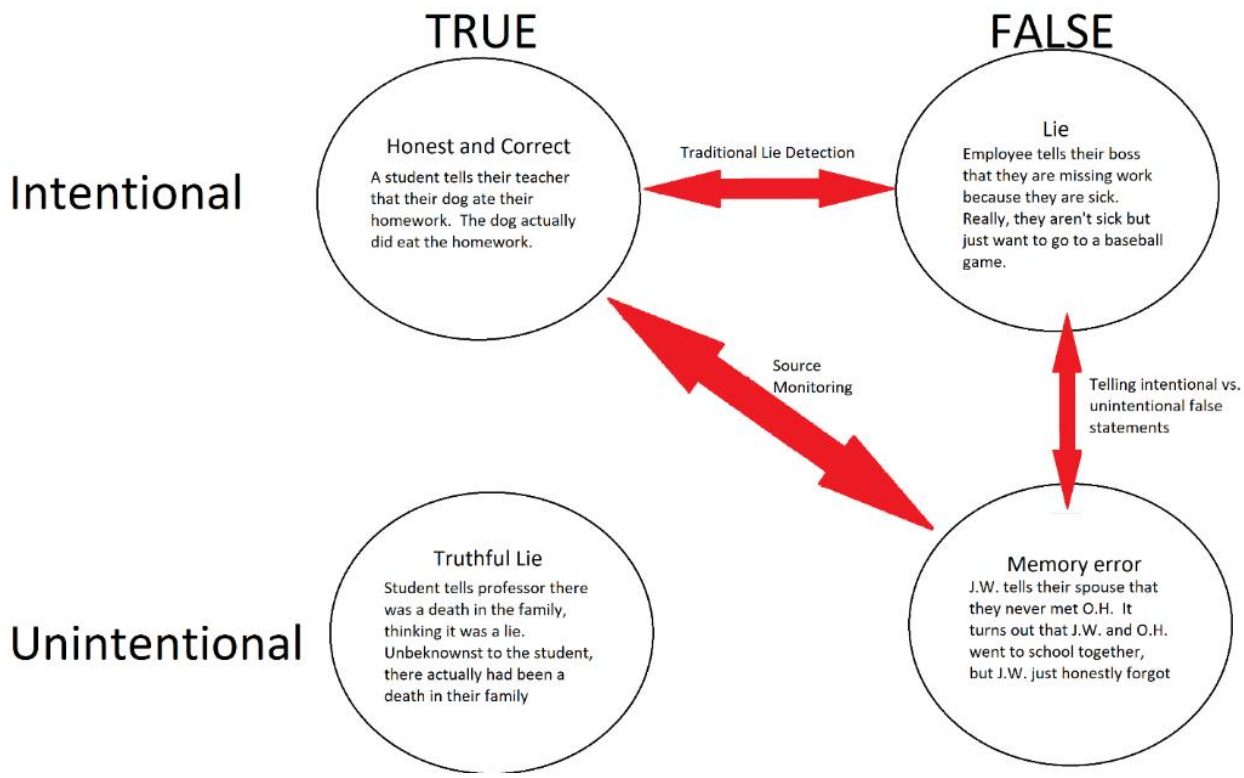

Figure 1. The distinction between types of true and false statements

Appendix B: Institutional Review Board Approval

UNIVERSITY OF ARKANSAS

To:	James M Lampinen MEMH 305
From:	Douglas J Adams, Chair IRB Expedited Review
Date:	10/27/2020
Action:	**Exemption Granted**
Action Date:	10/27/2020
Protocol #:	2008280144
Study Title:	Evaluation of Statements

The above-referenced protocol has been determined to be exempt.

If you wish to make any modifications in the approved protocol that may affect the level of risk to your participants, you must seek approval prior to implementing those changes. All modifications must provide sufficient detail to assess the impact of the change.

If you have any questions or need any assistance from the IRB, please contact the IRB Coordinator at 109 MLKG Building, 5-2208, or irb@ uark.edu.

cc: Andrew C Provenzano, Investigator